Lecture Notes in Electrical Engineering 178

David Jin
Sally Lin
Editors

Advances in Mechanical and Electronic Engineering

Volume 3

Springer

For further volumes:
http://www.springer.com/series/7818

David Jin and Sally Lin (Eds.)

Advances in Mechanical and Electronic Engineering

Volume 3

Editors
David Jin
Wuhan Section of ISER Association
Wuhan
China

Sally Lin
Wuhan Section of ISER Association
Wuhan
China

ISSN 1876-1100
ISBN 978-3-642-31527-5
DOI 10.1007/978-3-642-31528-2
Springer Heidelberg New York Dordrecht London

e-ISSN 1876-1119
e-ISBN 978-3-642-31528-2

Library of Congress Control Number: 2012940780

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

In the proceeding of Volume 3 of ICMEE2012, you can learn much more knowledge about Electronic Engineering and Electronic Communication; Electronic Engineering and Electronic Image Processing all around the world. The main role of the proceeding is to be used as an exchange pillar for researchers who are working in the mentioned field. In order to meet high standard of Springer, the organization committee has made their efforts to do the following things. Firstly, poor quality paper has been refused after reviewing course by anonymous referee experts. Secondly, periodically review meetings have been held around the reviewers about five times for exchanging reviewing suggestions. Finally, the conference organization had several preliminary sessions before the conference. Through efforts of different people and departments, the conference will be successful and fruitful.

During the organization course, we have got help from different people, different departments, different institutions. Here, we would like to show our first sincere thanks to publishers of Springer, LNEE series for their kind and enthusiastic help and best support for our conference.

In a word, it is the different team efforts that they make our conference be successful on June 23–24, 2012, Hefei, China.We hope that all of participants can give us good suggestions to improve our working efficiency and service in the future. And we also hope to get your supporting all the way. Next year, In 2013, we look forward to seeing all of you at ICMEE2013.

April 2012 ICMEE2012 Committee

Committee

Honor Chairs

Prof. Chen Bin	Beijing Normal University, China
Prof. Hu Chen	Peking University, China
Chunhua Tan	Beijing Normal University, China
Helen Zhang	University of Munich, China

Program Committee Chairs

Xiong Huang	International Science & Education Researcher Association, China
LiDing	International Science & Education Researcher Association, China
Zhihua Xu	International Science & Education Researcher Association, China

Organizing Chair

ZongMing Tu	Beijing Gireida Education Co.Ltd, China
Jijun Wang	Beijing Spon Technology Research Institution, China
Quanxiang	Beijing Prophet Science and Education Research Center, China

Publication Chair

Song Lin	International Science & Education Researcher Association, China
Xionghuang	International Science & Education Researcher Association, China

International Committees

Sally Wang	Beijing Normal University, China
LiLi	Dongguan University of Technology, China
BingXiao	Anhui University, China

Z.L. Wang	Wuhan University, China
Moon Seho	Hoseo University, Korea
Kongel Arearak	Suranaree University of Technology, Thailand
Zhihua Xu	International Science& Education Researcher Association, China

Co-sponsored by

International Science& Education Researcher Association, China
VIP Information Conference Center, China
Beijing Gireda Research Center, China

Reviewers of ICMEE2012

Z.P. Lv	Huazhong University of Science and Technology
Q. Huang	Huazhong University of Science and Technology
Helen Li	Yangtze University
Sara He	Wuhan Textile University
Jack Ma	Wuhan Textile University
George Liu	Huaxia College Wuhan Polytechnic University
Hanley Wang	Wuchang University of Technology
Diana Yu	Huazhong University of Science and Technology
Anna Tian	Wuchang University of Technology
Fitch Chen	Zhongshan University
David Bai	Nanjing University of Technology
Y. Li	South China Normal University
Harry Song	Guangzhou Univeristy
Lida Cai	Jinan University
Kelly Huang	Jinan University
Zelle Guo	Guangzhou Medical College
Gelen Huang	Guangzhou University
David Miao	Tongji University
Charles Wei	Nanjing University of Technology
Carl Wu	Jiangsu University of Science and Technology
Senon Gao	Jiangsu University of Science and Technology
X.H. Zhan	Nanjing University of Aeronautics
Tab Li	Dalian University of Technology (City College)
J.G. Cao	Beijing University of Science and Technology
Gabriel Liu	Southwest University
Garry Li	Zhengzhou University
Aaron Ma	North China Electric Power University
Torry Yu	Shenyang Polytechnic University
Navy Hu	Qingdao University of Science and Technology
Jacob Shen	Hebei University of Engineering

Contents

Analysis of Attacks on Data Aggregation Protocol in Electronic Communication Technologies

Shuai Fu[1], Chang-guang Wang[2], and Jianfeng Ma[1]

[1] School of Computer Science and Technology, Xidian University, Xi'an 710071, China
[2] College of Information Technology, Hebei Normal University,
Shijiazhuang 050016, China
ljhfs0803@126.com, wangcg@mail.hebtu.edu.cn

Abstract. As a kind of brand-new electronic technology, the data aggregation protocol is confronted with many difficulties by taking safety issues as representative. In this paper, we first analyze the well-known threats against data aggregation in WSNs. In order to study the performance of the common 6 kinds of data aggregation protocols in defending against such attacks, we categorize them into two parts: tree-based and cluster-based data aggregation protocols. On this basis, we compare their performance in defending against these attacks and point out which aspects of them should be improved for higher security.

Keywords: data aggregation, wireless sensor networks, security.

1 Introduction

Sensor networks are the new paradigm for the future electronic communication. Wireless nodes are deployed in large numbers inside or close to the inspected phenomenon to form a wireless sensor network (WSN). Sensor networks may consist of many different types of sensors, such as seismic, thermal, and radar, which can monitor the environment or system by the measurement of physical parameters such as temperature, pressure, relative humidity noise levels, etc [1]. Recently, WSN networks have been used in many promising applications including wildlife monitoring, human health monitoring, or intelligent transportation systems, etc.

Since the sensor nodes are energy-constrained, data aggregation technology is adopted to help the electronic communication between the BS and sensors energy-efficient. As the sensor nodes may be deployed in remote and hostile environments, attackers may control a sensor node and change the aggregation result or inject false data into the WSNs easily. Thus, the secure data aggregation becomes an important research field in electronic communication technologies. In this paper, we identify the threats and vulnerabilities to secure data aggregation for sensor networks by defining the main attack issues. Our objective is to provide a deeper understanding of the performance of current security data aggregation protocols in WSNs.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 1–6.
springerlink.com

2 Types of Attacks on WSN Aggregation

Many factors may lead to its being attacked in electronic communication technologies. WSNs are vulnerable to various types of attacks due to the nature of the transmission medium (broadcast). In this section, these attacks that might affect the aggregation in the WSN are discussed.

2.1 Sybil Attack

Sybil Attack is where the attacker is able to disguise itself as a valid sensor node in the network and present more than one identity within it. It can affect aggregation schemes in different ways [2]. Firstly, an adversary may create multiple identities in order for selecting a malicious node to be the aggregator. Secondly, the aggregated result may be affected if the adversary is able to generate multiple entries with different readings. Thirdly, an adversary can launch a Sybil attack and generate n or more witness identities to make the base station accept the aggregation results.

2.2 Wormhole Attack

In a wormhole attack, it always needs two compromised nodes to colluding and conducting attack together. An adversary tunnels messages received in one part of the network over a low-latency link and replays them in a different part to make a fake that these two parts are very close. Normally, the two colluding malicious nodes are distant, in which case one is close to the Sink node while the other is far away from it. The two compromised nodes communicate directly through relaying packets along an out-of-bound channel available only to the attackers.

2.3 Selective Forwarding Attack

A compromised node might discard all or part of the information packets, because it is up to the adversary that is controlling the compromised node to decide whether forward the received messages or not. Considering the aggregation scenario, any compromised intermediate nodes have the ability to launch the selective forwarding attack and subsequently this influence the aggregation values.

2.4 Denial of Service Attack

A Denial of service attack(DoS attack), which is sometimes called jamming, is the result of attacks on the WSN by transmitting radio signals that interfere with the radio frequencies used by the WSN. It prevents the WSN from functioning correctly in a timely manner. The increase of the adversary ability may lead to a larger portion of the network affected.

2.5 Stealthy Attack

Stealthy attack is one of the most common attacks caused the security issue in data aggregation. The goal of this attack is to inject false data into the network without revealing its existence. In a data aggregation scenario, sink nodes make critical

decision based on all the data, so an injected false data value will lead to a false aggregation result and totally change the final decision of the BS.

2.6 Replay Attack

Replay attack means the attacker repeatedly sends the previous sensing information without even understanding its content and replays them later on to mislead the aggregator. This attack will affect the freshness of sensor data and consequently the aggregation results will be alerted. To avoid the replay attack, the method of attaching a time specific tag with each packet transmitting in the network can be adopted.

3 Data Aggregation Protocols Analysis

To reduce energy consumption and eliminate data redundancy, many systems perform in-network data aggregation in the process of electronic communication. The typical protocols can be categorized into two parts: the tree-based and cluster-based data aggregation protocols.

3.1 Tree-Based Data Aggregation Protocols

3.1.1 Directed Diffusion

Directed diffusion [3] is a data-centric and application-aware routing algorithm for drawing information out of a sensor network. It is a reactive protocol which takes places in three phases: (a) interest dissemination; (b) gradient setup, and (c) path reinforcement and forwarding.

Possible attacks: Due to the robust nature of flooding, an adversary could attack a data flow with one of four goals once sources begin to generate data events: (1) suppression, which is an instance of denial-of-service. The easiest way to launch it is to spoof negative reinforcements.(2) Cloning: an adversary can replay the interest flooded from a base station with herself listed as a base station, thus realizing eavesdropping. (3) Path influence: An adversary can influence the path taken by a data flow by spoofing positive and negative reinforcements and bogus data events. (4) Selective forwarding: an adversary can gain full control of the flow by using the above attacks and modify and selectively forward packets of his choosing.

3.1.2 The Power-Efficient Data-Gathering Protocol for Sensor Information Systems (PEGASIS)

The basic idea of PEGASIS [4] is that nodes need only communicate with their closest neighbors, and they take turns in transmitting to the sink. It has two main objectives. First, increase the lifetime of each node by using collaborative techniques. Second, allow only local coordination between nodes that are close together.

Possible attacks: As nodes take turns in communicating with the base station in PEGASIS, it is susceptible to the Sybil attack. To locate the closest neighbor node, each node uses the signal strength to measure the distance to all neighboring nodes, so

every node is likely to choose the adversary as its closest neighbor node due to the large signal strength of the advertisement.

3.1.3 The protocol of Energy-Aware Distributed Aggregation Tree (EADAT)

The EADAT protocol is based on an energy-aware distributed heuristic. The algorithm is initiated by the sink which broadcasts a control message having five fields: ID, parent, power, status, and hop count. To maintain the tree, a residual power threshold Pth is associated with each sensor.

Possible attacks: Since residual power is used to distributively schedule the local broadcasts among neighboring sensors, an adversary may grasp the channel and broadcast the control packets by advertising a large amount of residual power. Then the adversary may launch a selective forwarding attack.

3.2 Cluster-Based Data Aggregation Protocols

3.2.1 The Low-Energy Adaptive Clustering Hierarchy (LEACH) protocol

The LEACH protocol is distributed and sensor nodes organize themselves into clusters for data fusion [5]. In this protocol, there are two main phases: The setup phase, which involves the organization of the network into clusters and the selection of cluster heads, and the steady-state phase, which involves data aggregation at the cluster heads and data transmission to the sink.

Possible attacks: Since nodes choose a cluster-head based on received signal strength, an adversary can send a powerful advertisement to all nodes in the network to attract them for choosing him as their cluster-head. Then the adversary can mount a selective forwarding attack on the entire network. As LEACH randomly selects a few sensor nodes as cluster heads and rotates this role to evenly distribute the energy load, they are vulnerable to the Sybil attack.

3.2.2 The Geographic Adaptive Fidelity (GAF) Protocol

GAF[6] places nodes into virtual “grid squares”, and for each particular grid area, a representative node acts as the leader to transmit the data to other nodes. Nodes are in one of three states: discovery, active and sleeping. Discovery nodes probe the network to determine the neighbors in the grid while active nodes participate in routing. Sleeping nodes have their radio turned off.

Possible attacks: It is noted that nodes in the discovery or active state that receive a discovery message from a higher ranking node will change to sleeping. An adversary can mount a selective forwarding attack or choose to ignore incoming packets completely by periodically broadcasting high ranking discovery messages. Furthermore, by broadcasting a high ranking discovery message from a non-existent node in each grid, an attacker can also target individual grids with the Sybil attack and a HELLO flood attack,. Done frequently enough, the adversary can ensure the entire network remains sleeping.

3.2.3 The Rumor Routing Protocol

Rumor routing uses a set of long-lived agents to create paths that are directed toward the events they encounter. Agents carry a list consisting of events, a TTL field, etc. The agent has a lifetime of a certain number of hops, after which it dies. When a query agent arrives at a node previously traversed by a event agent that satisfies the query, the route from a base station to a source is established.

Possible attacks: As the establishment of routes is entirely dependent on nodes properly handling agents, a Dos attack may be mounted by an adversary through removing event information carried by the agent or refusing to forward agents entirely. A smarter adversary may create a wormhole between a node near a base station and a node near a potential source, and then using the Sybil attack to maximize each node' s chance of being chosen as the initial destination of a event or query agent. Queries are immediately matched with events via the wormhole, and the adversary can then selectively forward events of her choosing. In addition, an adversary can also mount a selective forwarding attack by forwarding multiple copies of a received agent to create a sinkhole.

4 Comparison

According to the analysis above, we can give the results in a more intuitive way. From Table 1 we can see that, on the one hand, the ability of each protocol to defend against attacks is not necessarily linked with the type of it. The characteristics of the protocol itself determine the loophole. For instance, in the tree-based protocols, the Directed diffusion may encounter with three kinds of attacks while the EADAT does only one. On the other hand, it is obvious that the Sybil attack and the selective forwarding attack are the most common, and almost every protocol cannot defend against them. The essence of the problem is that the adversaries gain the power of receiving or forwarding data. In general, the selective forwarding attacks are used by combining other attacks such as Wormhole attack or Sybil attack. Through the analysis we realize that how to prevent an adversary to obtain the control power is the main task in the future research work.

Table 1. Summarization of the security of data aggregation protocols ("*" represents the attacks that the protocol may encounter)

Attack / Protocol	Sybil attack	Wormhole attack	Selective Forwarding attack	Dos attack	Stealthy attack	Replay attack
DF			*	*		*
PEGASIS	*		*			
EADAT			*			
LEACH	*		*			
GAF	*		*			
Rumor	*	*	*	*		

5 Conclusion

In this article we analyze the performance of typical data aggregation protocols of being against common attacks. As security issues pose a great challenge to WSNs in electronic communication technologies, analyses of the main threats to data aggregation protocols provide the foundation for further study. According to the comparison between them, we point out that trying to prevent an adversary to gain the power of receiving or forwarding data is the key problem for obtaining higher security.

Acknowledgements. This research was supported by the Natural Science Foundation of Hebei Education Department under No.2010257, the Science Foundation of Hebei Normal University.

References

1. Estrin, D., Govindan, R., Heidemann, J., Kumar, S.: Next century challenges: Scalable coordination in sensor networks. In: Proc. International Conf. Mobile Computing Networking, pp. 263–270 (1999)
2. Alzaid, H., Foo, E., Nieto, J.G.: Secure Data Aggregation in Wireless Sensor Network: a survey. In: Proceedings of the 6th Australasian Information Security Conference (2008)
3. Intanagonwiwat, C., Govindan, R., Estrin, D., Heidemann, J., Silva, F.: Directed diffusion for wireless sensor networking. IEEE/ACM Transactions on Networking 11, 2–16 (2003)
4. Lindsey, S., Raghavendra, C.: PEGASIS: Power-Efficient Gathering in Sensor Information Systems. In: IEEE Aerospace Conf. Proc., vol. 3(9–16), pp. 1125–1130. IEEE Press (2002)
5. Heinzelman, W.R.: Application-Specific Protocol Architectures for Wireless Networks. Ph.D. thesis, Massachusetts Institute of Technology (2000)

Mixed Collaborative Detection Based on the SNR

ZhenChao Wang and Chao Ma

College of Electronic and Information Engineering, HeBei University,
071000, BaoDing, China
{Wangzhenchaohbdx,mc1987mc}@163.com

Abstract. In cognitive radio(CR) system, the signal-to-noise ratio(SNR) of sensing signal can influence the result of the detection. This paper presents a collaborative algorithm for the spectrum sensing based on SNR. It groups the cognitive users (unauthorized users) according to their SNR, and synthesizes the detection results of grouped users using "*OR* "guideline or "*AND* " guideline respectively. The fusion center fuses the results of each group according to "*OR* " guideline finally. The algorithm improves the detection probability, because it has the characteristics of every cognitive user testing conclusion. The simulation results prove the validity of the algorithm.

Keywords: Cognitive radio, Spectrum detection, SNR, "*OR* " guideline, "*AND* " guideline, Mixed collaborative detection.

1 Introduction

With the development of communication technology, spectrum resource becomes more and more scarce, and the spectrum utilization of authorized users is very low[1].The appearance of CR technology provides the possibility to solve this problem. The cognitive users detect the spectrum and use the idle channel. This is cognitive radio. The accurate detection is the basis and premise of CR. Spectrum detection includes single-node detection and collaborative detection. Single-node detection is difficult to guarantee the credibility because of shadow effects, multi-path effects, SNR and other factors. While in collaborative detection, there will be more than one node detecting the band at the same time, and the final conclusion includes the detection results of every node. So it has become the main technology of CR.

2 Single-Node Detection

The single-node detection includes the matched filter detection, energy detection and cyclostationary detection. If the priori information of primary user(authorized user) is known first, the matched filter detection works. So it is lack of flexibility. Cyclostationary detection is achieved by using the stationarity of the signal, it can distinguish

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 7–12.
springerlink.com

the signal and noise. But the calculation is large. Energy detection needn't to know the prior information and it is easy to realize. So it is the most commonly detection algorithm. In this paper, we analyze collaborative detection algorithm based on the energy detection.

Energy detection can be seen as a binary hypothesis[2]:

$$\begin{cases} X(t) = n(t) & H_0 \\ X(t) = hs(t) + n(t) & H_1 \end{cases} . \tag{1}$$

Where $s(t)$ is a primary signal; $n(t)$ is the additive white Gaussian noise (AWGN); h is the channel gain; H_0 and H_1 represents there exiting the primary user or not respectively.

We make $m = TW$, which stands for the product of observation time and bandwidth. The detection probability P_d and the false alarm probability P_f can be expressed as Eq. 2 and Eq. 3[3]:

$$P_d = P_r\left(r > \lambda \middle| H_1\right) = Q_m\left(\sqrt{2\gamma}, \sqrt{\lambda}\right) . \tag{2}$$

$$P_f = P_r\left(r > \lambda \middle| H_0\right) = \frac{\Gamma\left(m, \frac{\lambda}{2}\right)}{\Gamma(m)} . \tag{3}$$

Where λ is the decision threshold; $\Gamma(\ )$ is the gamma function, and $\Gamma(.,.)$ is the upper incomplete gamma function; Q_m is the generalized Marcum function, which is expressed as Eq. 4:

$$Q_m(a,b) = \int_b^{\infty} \frac{x^m}{a^{m-1}} e^{\frac{x^2+a^2}{2}} I_{m-1}(ax)dx . \tag{4}$$

Where $I_{m-1}(.)$ is m-1-order modified Bessel function.

In Figure 1 we simulate P_d and P_f curve using single-node detection in different SNR (- 18 dB, -20 dB, -24 dB, - 28dB and - 30 dB).

From Figure 1, we can see the difference of the performance of the cognitive node in different SNR. With the reduction of SNR, the false alarm probability increases in the case of a certain detection probability. So we should give different trust degrees of every node according to the SNR.

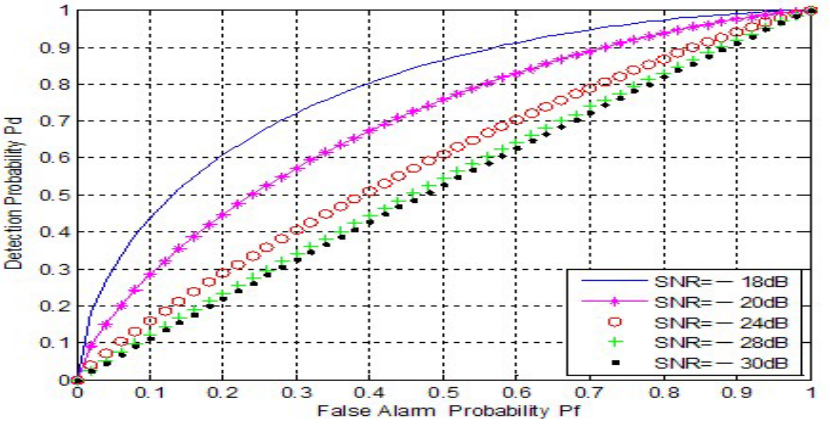


Fig. 1. Energy detection in different SNR

3 Collaborative Detection

Collaborative detection includes hard-merging algorithm and soft-merging algorithm. The decision of every node is sent to the fusion center, and the fusion center makes the final decision according to some criterion. That is hard-merging algorithm. It includes "OR" guideline, "AND" guideline and "K/N" guideline.

The primary user exists if there is only one cognitive user detecting the presence of primary user, that is "OR" guideline. We assume that there are N cognitive users, P_d and P_f can be expressed as Eq.5 and Eq. 6.

$$P_d = 1 - \prod_{i=1}^{K} (1 - P_{di}) \tag{5}$$

$$P_f = 1 - \prod_{i=1}^{K} (1 - P_{fi}) \tag{6}$$

Where P_{di} and P_{fi} represents the detection probability and the false alarm probability of every node respectively.

We can see from the equations above that "OR" guideline can not only improve P_d, but also increase P_f. It can reduce interference to the primary user because of misjudgment, but it makes the cognitive user decrease the probability of occupation of the idle channel.

The primary user exists in the band if all the cognitive users detect the presence of primary user. That is " AND " guideline. We assume that there are N cognitive users, P_d and P_f can be expressed as Eq. 7 and Eq. 8.

$$P_d = \prod_{i=1}^{K} P_{di} \tag{7}$$

$$P_f = \prod_{i=1}^{K} P_{fi} \tag{8}$$

" AND " guideline not only reduces P_f, but also reduces P_d. So it can bring higher spectrum utilization rate, but increase the probability of misjudging for idle state.

We assume that there are N cognitive users, the primary user exists in the band if there are K cognitive users detecting the presence of the primary user. That is" K/N "guideline. We can see that" K/N "guideline is " OR " guideline when $K = 1$, and " K/N " guideline is " AND " guideline when $K = N$.

The energy value of every node or the likelihood ratio is sent to the fusion center, then the fusion center synthesizes the results according to some algorithm. That is soft-merging algorithm. It usually includes the linear weighted algorithm and the likelihood ratio algorithm. It is better on detection performance, but it also increases the requirement of bandwidth and the pressure of the channel.

4 Mixed Collaborative Detection Based on the SNR

Actually, the distance to the primary user is different among cognitive nodes, leading to the difference of SNR that every node receives, which causes different reliability.

Taking the limitation of " OR " guideline and " AND " guideline using by itself and the different SNR of every cognitive node into consideration, this paper proposes mixed cooperative detection based on the SNR.

We assume that every cognitive node has a function of estimating SNR (a variety of techniques can be used to estimate SNR). Every cognitive node sends their own verdict and SNR information to the fusion center. The process is as follows:

Step1: Every cognitive user uses the energy detection for spectrum sensing independently, and obtains their own SNR;

Step 2: Every cognitive user sends their own verdict and SNR information to the fusion center;

Step 3: Fusion center determines the SNR threshold according to the SNR which every cognitive user sends to the center (using the mean of the SNR of all the cognitive users in this paper).Then the fusion center divides cognitive users into two groups according to the comparison the SNR of every cognitive user and the threshold;

Step 4: The users whose SNR are greater than the threshold are as a group. The fusion center uses " OR " guideline in this group and draw a conclusion;

Step 5: The users whose SNR are less than the threshold are as a group. The fusion center uses " AND " guideline in this group and draw a conclusion;

Step 6: The fusion center re-uses " OR " guideline on the judgment results of the two groups and makes the ultimate decision.

We assume that there are N cognitive users and the SNR of K of N are greater than the threshold. So these K users use " OR " guideline. And the other $N-K$ users use " AND " guideline. P_{di} and P_{fi} represents the detection probability and the false alarm probability of every node respectively.

The final detection probability and the final false alarm probability can be expressed as Eq. 9 and Eq. 10.

$$P_d = 1 - \left(1 - C_N^K \left[1 - \prod_{i=1}^{K} \left(1 - P_{di}\right)\right]\right)\left(1 - \prod_{i=1}^{N-K} P_{di}\right) \quad . \tag{9}$$

$$P_f = 1 - \left(1 - C_N^K \left[1 - \prod_{i=1}^{K} \left(1 - P_{fi}\right)\right]\right)\left(1 - \prod_{i=1}^{N-K} P_{fi}\right) \quad . \tag{10}$$

4 Simulation

Based on the process above, we set the SNR values of the cognitive users to -18 dB, -20 dB, -24 dB, -28 dB and -30 dB. According to the provisions of the algorithm, the SNR threshold should be set to -24 dB (the mean of all the SNR values). -24 dB is just at the mean point, we set it less than the SNR threshold. Thus, -18 dB, -20 dB are in the group of “large SNR” using the " OR " guideline.-24 dB, -28 dB and -30 dB are in the group of “small SNR” using the " AND " guideline. Simulation is shown below.

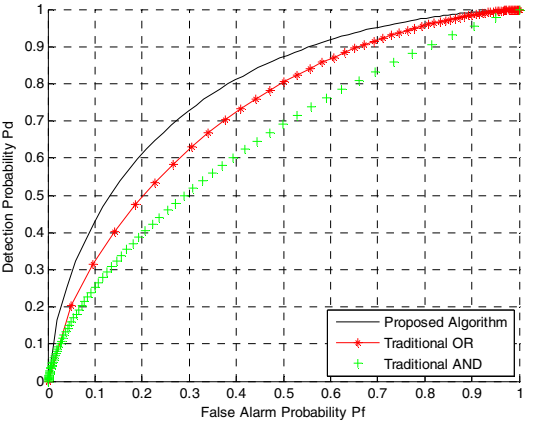


Fig. 2. Comparison proposed method with conventional detection methods

It can be seen from Figure 2 that the algorithm this paper proposes is better than the traditional "*OR*" guideline and "*AND*" guideline. We can get higher detection probability in the case of a certain false alarm probability and achieve better detection performance. This shows that the proposed detection method is feasible and effective in improving the detection probability.

5 Conclusion

Cognitive radio as a new solution to solve the shortage of spectrum resource has been widely concerned in recent years, and spectrum sensing is a very important aspect of cognitive radio. This paper focuses on the collaborative detection algorithm. We introduce mixed collaborative detection based on the SNR on the basis of analyzing "*OR*" guideline and "*AND*" guideline. The algorithm is not only the combination of hard-merging algorithm, but also reflects the idea of the soft-merging algorithm. Namely, the cognitive nodes not only send the local decision, but also the SNR values to the fusion center and the fusion center can process these data better. Compared with traditional collaborative detection method, the proposed method increases the detection probability in the condition of the same false alarm probability.

References

1. Kolodzy, P.: Spectrum policy task force: findings and recommendations. In: International Symposium on Advanced Radio Technologies, ISART, Colorado, USA (2003)
2. Zhou, X., Zhang, H.: Principle and Application of Cognitive Radio. Beijing University of Posts and Telecommunications Press, Beijing (2007)
3. Dighim, F.F., Alouini, M.S., Simon, M.K.: On the energy detection of unknown signals over fading channels. In: IEEE International Conference on Communications, Alaska, pp. 3575–3579 (2005)
4. Wu, S., Zhu, J., Qiu, L., Zhao, M.: SNR-based weighted cooperative spectrum sensing in cognitive radio networks. The Journal of China Universities of Posts and Telecommunications 17, 1–7 (2010)
5. Wang, H., Su, X., Wang, J.: Cooperative Spectrum Sensing Algorithms Based on Correlation Matrix in Cognitive Radio Networks. Tsinghua Science and Technology 16, 386–392 (2011)

The Optimization Design of Main Circuit for Hybrid Active Power Filter

Jingrong Yu, Pan Pan, Mi Dong, Yao Sun, Mei Su, and Weibiao Wu

School of Information Science and Engineering, Central South University,
Changsha 410083, Hunan Province, China
mi.dong@csu.edu.cn

Abstract. The performance of hybrid active power filter (HAPF) with different components and parameters is analyzed, a multi-objective optimization algorithm based on improved particle swarm optimization (PSO) is proposed. To reduce the influence of engineering experiences and redundant constraints, a unified description of performance for HAPF is proposed on the basis of theoretical analysis, and redundant constraints are removed from the optimization algorithm; To avoid premature of PSO algorithm, non-linear time-varying parameters are introduced, and a new global optimal particle selection operator is employed to improve solution diversity. The experimental results show that the proposed optimal design of main circuit for HAPF is of correctness and effectiveness.

Keywords: Hybrid active power filter, Particle swarm optimization algorithm, Multi-objective optimization, Pareto optimum, Injection circuit.

1 Introduction

Due to the increasing application of nonlinear loads, power grid is now heavily polluted by harmonics. There are three solutions to restrain harmonic pollution in power system—Passive Power Filter (PPF), Active Power Filter(APF) and Hybrid Active Power Filter(HAPF)[1]. Although PPF is simple and less expensive, its filtering characteristics are heavily dependent on parameters of power grid, and it may be easily detuned and get resonant. APF can overcome some defects of PPF, and it is much more effective in filtering dynamical harmonics[2]. However, the cost of APF is much higher, and the application of solely APF with large capacity cannot be implemented in large scale[3]. HAPF is a proper combination of PPF and APF, and it makes good tradeoff between filtering performance and the cost of investment. Among various topologies of HAPF, the topology with inject circuit and resonance circuit presented in this article is especially suitable for application of harmonic elimination in high voltage and large capacity field.

For a given topology and control objective, proper combination of parameters for main circuit is essential to assure the performance and reduce the cost of HAPF.

Recently there are some researches about the optimization of main circuit for HAPF[4][5], but the optimization designs presented in those researches are usually

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 13–19.
pringerlink.com

based on many constraints from engineering experiences, which is comparatively coarse, and may cause unnecessarily cost or sacrifice the diversity of optimized solution.

In order to increase the variety of the optimized solution for the design of main circuit, a novel multi-objective optimization algorithm based on improved particle swarm optimization for the design of main circuit of HAPF is proposed. In this optimization algorithm, a unified performance description of HAPF is used as the objective function, and suitable constraints based on theoretical analysis are summarized. Since non-linear time-varying parameters and a new global optimal particle selection operator are introduced to particle swarm optimization (PSO) algorithm, the optimization results of main circuit for HAPF are more various. Furthermore, in the proposed method the influence of engineering experience and redundant constraints could be avoided, and the convergence and diversity of multi-objective optimization algorithm are improved.

2 Description of HAPF Optimization Issue

The main circuit of HAPF consists of inject circuit and active part, its topology is shown in Fig.1. Where e_s is the source voltage of the power system, L_s is the equivalent inductor of lines in power supply side, T_f is coupling transformer, C_f and L_f are capacitor and inductor of output filter for APF. The inject circuit is composed of C_{inj}, R, C_1 and L_1, and C_{inj} is inject capacitor, R, C_1 and L_1 are resistor, capacitor and inductor of resonance circuit. The function of inject capacitor C_{inj} is on one hand to compensate the reactive power of system and share the most part of fundamental harmonic, on the other hand it provides low impedance path for harmonic currents. Active part, which realizes the function of restraining harmonics, consists of output filter, coupling transformer and voltage source inverter. The output current of active part flows into power grid through inject circuit, and output filter is utilized to filter out the switching harmonics caused by power electronic devices.

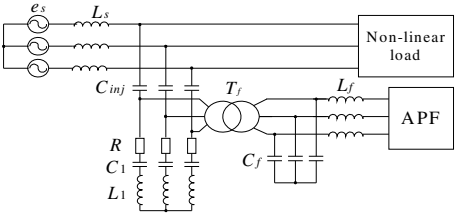


Fig. 1. The main circuit of HAPF

The optimization design of HAPF system mainly includes two parts, inject circuit and active part. Since there is little coupling between the two parts, the optimization design is beginning separately.

The design of inject circuit can be described as multi-objective optimization issue, and its optimization objective is $min\ \{J_{inj}, \sum_{n=2}^{M} k_n (I_{Sh}^n)^2\}$, where J_{inj} is cost of inject

circuit, k_n is weight of nth order harmonic current and i_{Sh}^n is the n^{th} harmonic current after compensation by HAPF.

Constraints of inject circuit include:

① The reactive power compensation should meet the requirement, that is $Q_{\min} < S_{Cinj1} < Q_{\max}$, where S_{Cinj1} is the fundamental capacity of inject capacitor, and Q_{min}, Q_{max} are minimum and maximum value of passive power capacity respectively.

② The resistor R of resonance circuit should meet requirement of quality factor, that is $R = \dfrac{\omega_1 L_1}{Q}$, where ω_1 is angular frequency of fundamental current.

The optimization objective of active part is $min\ \{C_f, L_f, f, \mathrm{S_{CDC}}\}$, and its constraints include:

① $1.2N\omega_1 < \omega_{res} < \omega_s/2$, where ω_{res} is the resonance frequency, and ω_s is angular frequency of switching frequency.

② $\max(|\vec{e}(t)|) < \Delta I_M$, where $\vec{e}$ is current tracking error in active part, and ΔI_M is the maximum ripple current.

③ $\max(|\Delta V_{dc}(t)|) < \Delta V_M$, where ΔV_{dc} is the fluctuation of DC voltage, and ΔV_M is an engineering allowed range.

3 The Improved PSO Algorithm

In the optimization design of main circuit, many objectives are conflict with each other, and those objectives have to be optimized at the same time[6]. Therefore, the design of main circuit is actually a multi-objective optimization issue. Usually, through weighting method, multi-objective optimization can be transformed into single-objective one, but suitable weights have to be obtained through abundant prior knowledge. Recent researches show that the optimization algorithms based on populations such as Particle Swarm Optimization (PSO) are more suitable for such problems[7][8], multi-objective algorithm based on PSO will find the best solution set, namely Pareto best solution by coordinating among each objective function[9]. To improve convergence and diversity of multi-objective algorithm based on PSO, an improved PSO multi-objective algorithm is proposed to solve the problem of optimization design for HAPF main circuit.

In PSO algorithm, the individuals who have better suitability to the environment are transferred to better areas. Firstly, a group of particles were randomly initialized in the solution space; After several iterations, the best solution will be found by these particles and the iteration functions are

$$v_{id}(k+1) = wv_{id}(k) + c_1 r_1 (p_{best} - x_{id}(k)) + c_2 r_2 (g_{best} - x_{id}(k)) \quad (1)$$

$$x_{id}(k+1) = x_{id}(k) + v_{id}(k) \tag{2}$$

Where, w is inertia weight, c_1 and c_2 are accelerating constants, r_1 and r_2 are stochastic function with a range [0,1].

The standard PSO is subject to single-objective optimization problem, but it is able to realize multi-objective optimization by introducing Pareto optimum. To improve convergence and diversity of multi-objective PSO algorithm based on Pareto optimum, the improved algorithm proposed in this article introduces filing operation, congestion estimation, etc.

To improve efficiency and avoid premature of optimization, the improved algorithm takes measures of increasing weight coefficient and perception coefficient, so that it would not stuck into local optimum. Nonlinear time-varying method is employed in the improved algorithm to modify control parameters, the specific operation is

$$w(k) = w_{\min} + (\frac{iter_{\max} - iter}{iter_{\max}})^{\alpha} \cdot (w_{\max} - w_{\min}) \tag{3}$$

$$c_1(k) = c_{1\min} + (\frac{iter_{\max} - iter}{iter_{\max}})^{\beta} \cdot (c_{1\max} - c_{1\min}) \tag{4}$$

$$c_2(k) = c_{2\max} + (\frac{iter_{\max} - iter}{iter_{\max}})^{\gamma} \cdot (c_{2\min} - c_{2\max}) \tag{5}$$

The realization steps of improved PSO algorithm base on Pareto optimum are

① Set the parameters of algorithm, set group scale as M, maximum searching iteration time as $iter_{\max}$; Initialize inertia weight and accelerating constant, randomly initialize primary population P, and $iter = 1$.

② Identify non-inferior solutions in the population P, and file them into global optimal archive A_g; Each particle is allocated with a outside archive $A_i, i \in \{1, M\}$, the local non-inferior solution of each particle is backed up.

③ Evaluate particle density for non-inferior solution in the archive A_g by adaptive grid technique. Evaluation standard is: for the grids (its set is s) whose particle number $N_{grid[i]}$ is more than 0, the opportunity for the particles in the this grid to be chosen as G_{best} is $1/[N_{grid[i]} \sum_{i \subset S} \Sigma(1/N_{grid[i]})]$, thus particles with large density were given a proper punishment to reduce their opportunity of being G_{best}. Moreover, the number of particles in each grid has an upper limit, and particles in the grid with overflow quantity of particles will be deleted randomly.

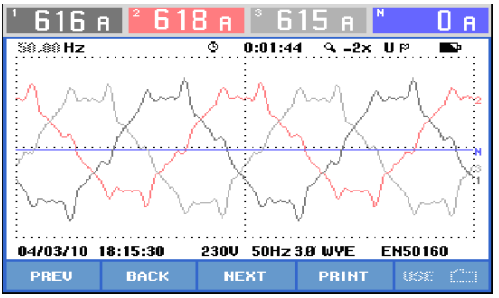
④ Randomly choose a particle from A_i as p_{best} for i^{th} particle, and update particles following iterative formula in the function (1) and (2). In the iterative progress, time-varying coefficients (weights and accelerating coefficients) update following functions (3-5).

⑤ When all the particles finished update for one generation, $iter = iter + 1$, or else turn back to step 3.

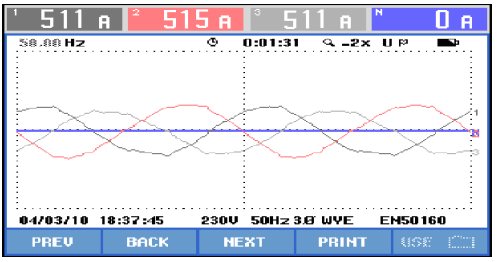
⑥ If convergence condition is achieved, the optimization will finish, or else it will turn back to step 2.

4 Experiment

HAPF system, after its main circuit carefully optimized, is tested in a metallurgy plant in Jiangxi, China. Load currents contain harmonic current and passive power current, where 5^{th} and 7^{th} harmonic currents are dominated. Due to heavy pollution of harmonics, capacitor compensation devices were not able to operate normally, which leads to problem of low power factor, and the average power factor is almost 0.89 per month. Fig.2.a records current waveform (the low voltage side of transformer) before compensation. Fig.2.b is current waveform after putting HAPF into operation.



a. Source current waveform before compensation



b. Source current waveform after compensation

Fig. 2. Compensation results of HAPF

According to Fig.2, current distortion rate of metallurgy plant is greatly decreased by the application of HAPF system. The details of compensation effect are displayed in Table 1, in which total harmonic distortion (THD) is measured by FLUKE 434, power factors are average power factors per month counted by Watt-hour Meter of consumers.

After three months operation of APF system, it exposes considerable harmonic restraint and reactive power compensation effects, and fast dynamic response rate, after all meets the customer's requirements well.

Table 1. Compensation effects under different working conditions

Working condition	RMS value	THD	Power factor
Source current before compensation	616A	14.9%	0.81
Compensation effects	514A	3.5%	0.96

5 Conclusion

The theoretical analysis of operation principle for HAPF system is developed, and a unified performance description for HAPF system is proposed. To improve diversity of optimization of HAPF main circuit, influences of each component parameter to performance are researched, and impacts of redundant constraints and engineering experiences are removed from optimization algorithm. Moreover, Pareto optimum is introduced to solve the multi-objective issue, and a improved PSO algorithm is proposed. The inject circuit and active part of HAPF system are optimized separately with the proposed PSO multi-objective algorithm, and best solution set derived from optimization of HAPF main circuits is applied in practical engineering design. The optimization algorithm proposed in this article would provide more selectable proposals and new idea for engineers, and simulation and experiments demonstrate the validity of the solution.

Acknowledgments. This work was supported by the National Science Foundation Foundation under Grant 61104047 and Research Fund for the Doctoral Program of Higher Education of China under Grant 20100162120014.

References

1. Peng, F.: Applications issues of active power filters. IEEE Industry Applications Magazine 4(5), 212–230 (1998)
2. Chen, G., Lv, H., Qian, Z.: The general principle of active filter andits applicalion. Proceedings of the CSEE 20(9), 17–21 (2000) (in Chinese)
3. Xiao, X., Xu, Y., Liu, H.: Analysis of hybrid power compensator and its experimental study. Automation of Electric Power Systems 26(3), 39–44 (2002) (in Chinese)
4. Yang, H., Ren, S.: A practical series-shunt hybrid active power filter based on fundamental magnetic potential self-balance. IEEE Transactions on Power Delivery 23(4), 2089–2096 (2008)
5. Luo, A., Tang, C., Shuai, Z.K.: A novel three-phase hybrid active power filter with a series resonance circuittuned at the fundamental frequency. Transactions on Industrial Electronics 56(7), 2431–2440 (2009)
6. Shuai, Z., Luo, A., Zhu, W., et al.: Study on a novel hybrid active power filter applied to a high-voltage grid. IEEE Transactions on Power Delivery 24(4), 2344–2352 (2009)

7. Jiang, Y., Liao, D.: Multi-objective optimal design for hybrid active power filter based on composite method of genetic algorithm and particle swarm optimization. In: International Conference on Artificial Intelligence and Computational Intelligence, vol. 2, pp. 550–553 (2009)
8. Del Valle, Y., Venayagamoorthy, G.K., Mohagheghi, S., et al.: Particle Swarm Optimization: Basic Concepts, Variants and Applications in Power Systems. IEEE Transactions on Evolutionary Computation 12(2), 171–195 (2008)
9. Coello, C.A.C., Pulido, G.T., Lechuga, M.S.: Handling multiple objectives with particle swarm optimization. IEEE Transactions on Evolutionary Computation 8(3), 256–279 (2004)

Research of FMCW for Perfect Binary Sequences Pairs

Jie Zhang[1] and Ran Bai[2]

[1] Academy of National Defense Information, Wuhan, 430010, China
uranus.zhang@gmail.com
[2] Hebei Normal University, Shijiazhuang, 050016, China

Abstract. An improved FMCW (frequency modulation continuous wave) radar system is introduced. The system includes perfect binary sequences pairs theory, the method using auto-correlation functions to determine IF (intermediate- frequency) signal, and measure distance and velocity, enable itself get constant development with theoretical progress. The new system only changing program of emission signal can be realized by DSP, which can reduce the fake alert notably so as to be applied more widely.

Keywords: stepped-frequency modulation, multi-target detection, perfect binary sequences pairs.

1 Introduction

Frequency modulation continuous wave (FMCW) radar range finder is a research hotspot in the car anti-collision, the main part of the car anti-collision system is a range finder radar in the car, it's purpose is to measure the real-time distance and speed of opposite car, it will alarm the driver or start the auto brake system in danger to avoid the collision. The FMCW radar works on the 30-100GHz millimeter band, it has less weather influence than laser or IR ranger finder system, it can normally work in rain, frog and dark, it also has low emission frequency, high distance resolution, high time-bandwidth product, high sensitivity and suitable for short range measurement. The time of FMCW radar applying in the car anti-collision system was short, and it's working conditions is full of clutter and noise, there has no mature pattern in it's signal detection and process. The main technology difficulty is fake alarm which need the angle resolute capability, and emission signal with high anti-jam capability, and efficient signal-processing algorithm.[1] We design the radar's emission signal by adapting perfect binary sequences pairs theory to meet the fake alarm in the speed and distance measurement.

2 The Principle of FMCW Radar

The constitution of FMCW radar measurement system is as figure 1. the system is composed of single-chip, transceiver front-end, modulation signal producer, IF amplifier, signal processor and alarm circuit.[6] Modulation signal circuit produce a low frequency triangular wave, the wave is sent to the Voltage Controlled Oscillator (VCO)

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 21–27.
springerlink.com

and it's center frequency f0 is controlled, after processed by the directional coupler, circulator and antenna, the modulated microwave signal radiate electromagnetic wave, when meeting target, the reflect signal enter the mixer after passing the antenna and circulator, and mixed with the emission signal from the directional coupler, the beat frequency was sent to the signal processing circuit after the intermediate frequency amplification, the relatively distance and speed were calculated by the single-chip and signal processing circuit, depends the danger level the system will send alarm signal, the waveform of signal frequency is as figure 2.

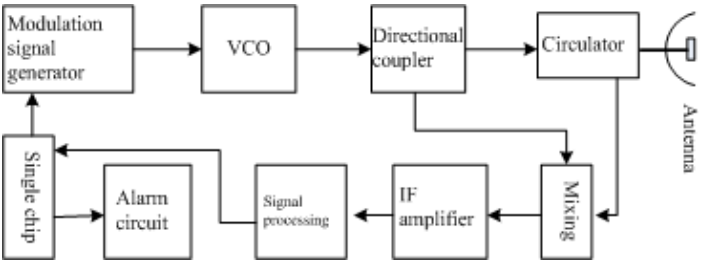


Fig. 1. The construction of FMCW radar measurement system

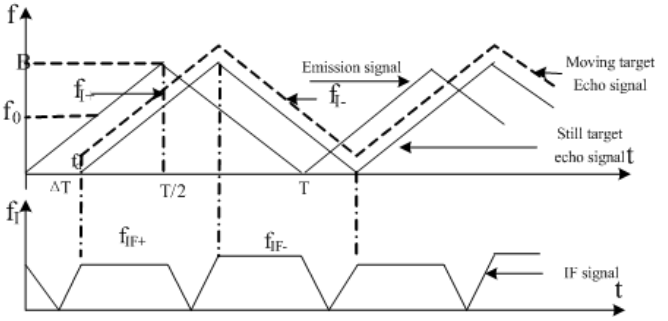


Fig. 2. The waveform of signal frequency

When the radar works on the continuous wave, it's emission signal can be represented as: $s(t) = A\arccos(\omega_0 t + \phi)$, ω_0 is the pulsatance, φ is the initial phase, A is the amplitude of the RF signal, the reflection echo signal is $s_r(t) = Ks(t - t_r) = KA\cos[\omega_0(t - t_r) + \phi]$, t_r is the delay time of the echo signal. As to still target, the distance R between radar and target is a constant, the delaying time of the echo signal is also a constant, say ΔT, $\Delta T = 2R / C$, the phase difference between echo signal and emission signal is ω_0, $\Delta T = 2\pi f_0 \bullet (2R / C) = 4\pi R / \lambda$, which is the phase-lag of the wave traveling between the radar and target. C is the speed of the electromagnetic wave, λ is the wavelength of emission signal.

As to moving target, the distance R changes with time, we assume the relative movement between radar and target is uniform motion, on moment t, the distance R(t) between radar and target is $R(t) = R_0 - Vt$, R_0 is the distance on t=0, V is the radial velocity of target relative to radar. The relative speed V is far less than the speed of the electromagnetic wave C, so the delay time t_r can be approximate written: $t_r = 2R(t)/C = 2(R_0 - Vt)/C$. The high frequency phase difference between echo signal and emission signal is: $\phi = -\omega_0 t_r = -\omega_0 2(R_0 - Vt)/C = -4\pi(R_0 - Vt)/\lambda$, when speed is a constant, we make differentiation to ϕ , we get the frequency difference between echo signal and emission signal: $f_d = (1/2\pi)(d\phi/dt) = 2V/\lambda$, which is the Doppler frequency.

When the radar works on the triangle wave liner FM, as to still target, the IF signal exported by mixer is f_I , $f_I = 4BR/TC$, T is the cycle of modulated triangle wave, B is the bandwidth of the emission signal.[6][7]

As to moving target, the reflected signal includes Doppler frequency f_d , in the increasing potion of the modulated signal,

$$f_{I+} = \frac{4BR}{TC} - f_d \tag{1}$$

In the decreasing potion of the modulated signal,

$$f_{I-} = \frac{4BR}{TC} + f_d \tag{2}$$

By adding and subtracting f_{I+} and f_{I-} , we get

$$R = \frac{(f_{I+} + f_{I-})TC}{8B} \tag{3}$$

$$V = \frac{C}{4f_0}(f_{I-} - f_{I+}) \tag{4}$$

f_0 is the center frequency of the emission signal, concluding from equation R and V, we can calculate the distance and speed via the real-time calculating of f_{I+} and f_{I-}.

3 The Application of Perfect Binary Sequences Pairs Theory in FMCW Radar

3.1 The Problem of FMCW Radar in Multi-target Detection

When we get f_{I+} and f_{I-}. we have two lines with opposite slope in the R-V plane, as to single target, there will be only two lines in the R-V plane, the intersection is the real target's information point, then we could determine the R, V by adopting equation (3) and (4), the system will make an in time reflection. But in most circumstances we

have more than one target, there will be multiple reflected signal in the increasing and decreasing portion of the emission signal, by using above method, we get multiple target information points, we assume there be two targets, as figure 4 shows, 1, 2 are real target information points, 3, 4 are fake target information points, and there are 50 percent fake alarms, this will cause the system lose the alarm ability. Too much fake alarm will not decrease the burden of the driver, on the opposite, it will make them highly nervous. We make the effort to solve this problem by introducing the perfect binary sequences pairs theory.

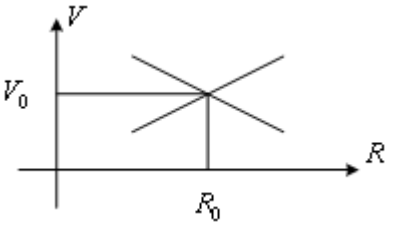


Fig. 3. Single target determination

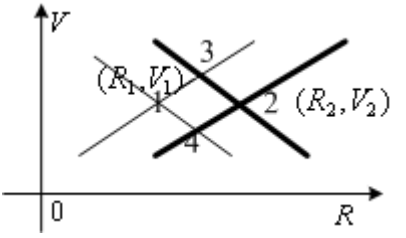


Fig. 4. Two targets determination

3.2 Perfect Binary Sequences Pairs Theory

In the area of radar, spread spectrum communication and mobile communication, we often need to design all kinds of sequences, and then we need to determine whether the sequence is the best signal by using the sequence's cyclic auto-correlation function, which is represented by the inner product of the sequence and its delay sequence's conjugate sequence. The more less the sequence's out phase cyclic auto-correlation function value, the better the sequence. If the value is 0, the sequence is the optimum signal. As to the binary sequence, only one sequence will be used to form optimum signal, which limited the existential space of the optimum signal, then the research which use sequence pairs (X,Y) composed of two sequences to form optimum signal,[2][3] this enhanced the request to the optimum signal in engineering. Before when the sequence is applied in some system, same sequence must be applied in the system's sender and receiver. When using sequence pairs formed by sender's sequence and receiver's local sequence, if it meets some prerequisites (the prerequisite to the optimum signal in engineering), the sequence pairs will form an optimum signal.

The autocorrelaton function of perfect binary sequence pairs is defined as below: assume $X=(x_0,x_1,...,x_{n-1})$ and $Y=(y_0,y_1,...,y_{n-1})$ are two sequences, function

$$R_{(X,Y)}(\tau)=\begin{cases}\sum_{i=0}^{n-1}x_i y_i = F, & \tau = 0(\bmod n)\\ \sum_{i=0}^{n-1}x_i y_{i+\tau}=0, & \tau \neq 0(\bmod n)\end{cases}$$

is the autocorrelaton function of signal X and Y,[2][3] $n+\tau=(n+\tau)\bmod n$, F is an integer not equal to 0.

3.3 The Principle of Multiple Targets Detection

The Improved FMCW radar's emission signal will be multi-slope stepped frequency modulated continuous wave, the signal's period will be triangle wave (Figure 5) [1][7], when apply time frequency modulated to the emission signal, the emission signal is modulated by on sequence(X sequence) of the sequence pairs in the increasing portion of triangle wave, in the decreasing portion, use another sequence(Y sequence) to modulate emission signal. Assume there are two targets, after reflection, mixing and FFT processing, we get four IF signals of two groups (f_{I1+}、f_{I2+}、f_{I1-}、f_{I2-}), and corresponding pulse sequences X_1、X_2、Y_1、Y_2, which are the IF pulse sequences of the emission signal X and Y after reflection delay. Using (3) and (4), we get four groups R and V, two of them are fake alarms, in order to get ride of them, when get a group of R and V, we calculate the X and Y delay sequences' autocorrelaton function corresponding to this group's IF.[4] For example, when we get R and V from f_{I2+}、f_{I1-} we calculate corresponding X_2、Y_1 's autocorrelaton function $R_{(X_2,Y_1)}(\tau)$ of f_{I2+}、f_{I1-}, because X_2、Y_1 are X and Y's delay with different value, which cause the τ in X_2、Y_1 's autocorrelaton function $R_{(X_2,Y_1)}(\tau)$ equals to 0, then we can decide this group's R and V be fake information point.

As above shows, depends on the value of autocorrelaton functions of four groups' $(X_1、Y_1),(X_2、Y_2),(X_1、Y_2),(X_2、Y_1)$, which correspond to the four groups' IF signals $(f_{I1+},f_{I1-}),(f_{I2+},f_{I2-}),(f_{I1-},f_{I2+}),(f_{I1+},f_{I2-})$,we can distinguish the correct same group relation of IF signals, then we get two groups of real R and V, we get the real target information points by get ride of fake points.

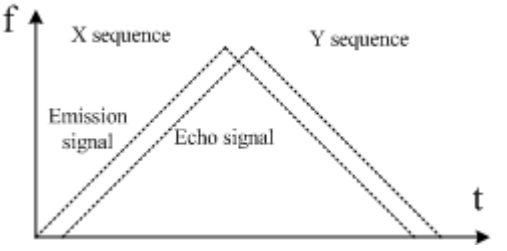


Fig. 5. The emission signal changing with t

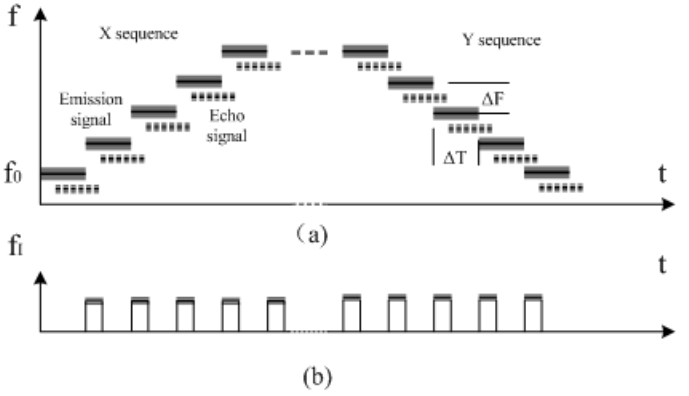


Fig. 6. The emission signal and echo signal

The whole system's flow chart is as Figure7.

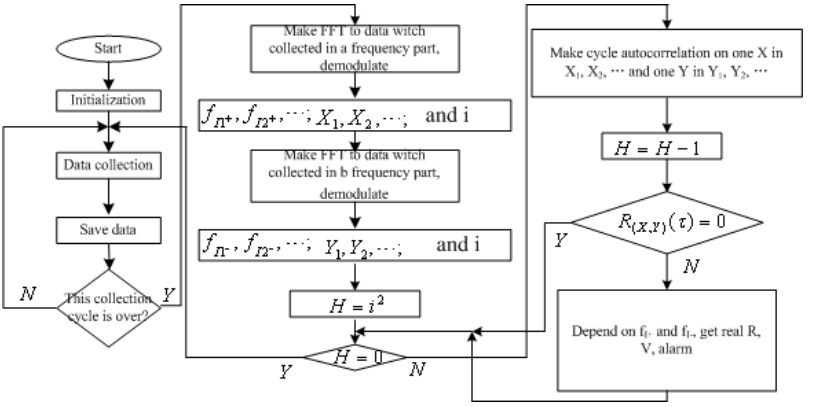


Fig. 7. System's flow chart

4 Conclusion

We get this improved system based on the previous FMCW radar, we only need to improve the emission signal hardware which apply the time sequence modulate. Because the huge improvement of DSP and software radio, we can easily implement the time sequence modulate and DSP by programming. The perfect binary sequence pairs of pseudo-random code has good relativity and easy to get, it can be used widely. This system is easy to implement, and can vastly reduce the probability of fake alarms, which make the FMCW radar better apply in the car anti-collision system.

References

1. Zhang, D.-B., Yu, H.-L.: Design of automobile collision avoidance warning system based on LabVIEW. Computer Engineering and Applications 44(21), 54–56 (2008)
2. Xu, C.-Q., Jin, H.-L.: Families of Almost Perfect Periodic Complementary Binary Sequence Pairs. Systems Engineering and Electronics 25(9), 1086–1089 (2003)
3. Zhao, X., He, W.: The theory of the perfect binary array pairs. Acta Electronica Sinica 27(1), 34–37 (1999)
4. Jia, Y.-G., Xu, C.-Q., Tang, Y.: Study on the uniqueness of perfect binary array pairs. Journal of China Institute of Communications 25(9), 112–117 (2004)
5. Liu, G., Hou, D., Li, K.: Warning algorithm of vehicle collision voidance system. Tsinghua Univ. (Sci.&Tech.) 44(5), 697–700 (2004)
6. Ding, L.-F., Geng, F.-L.: Radar Principle. Xidian press, Xian (1997)
7. Zhang, D.-B., Kang, Y.-Y., Liu, H.-Y.: The design of Automobile Collision Avoidance Warning System Based on LabVIEW. In: The Eighth International Conference on Electronic Measurement and Instruments 2007 (2007) IEEE Catalog Number:07EX1748 P2-51-P2-54

Hardware Design of the Open Channel Ultrasonic Sensor

Xingguo Qiu[1], Yu Liu[1], and Yi Ren[2]

[1] College of Computer Science and Technology, Xi'an University of Science and Technology, Xi'an 710054, China
[2] Xi'an Xinyuan Control Technology Co.,Ltd. Xi'an 710065, China
royy27@163.com

Abstract. The paper designs the ultrasonic sensor according to time difference method. Firstly the basic principles of time difference method are introduced to measure open channel flow. And then the overall design framework of the transit-time ultrasonic sensor is proposed. Lastly specific hardware design is given .The paper focuses on describing design process of transmitting circuit, receiving circuit and communication circuit. Complete the hardware system by using the MSP430 as the control core, with a strong anti-interference ability. Therefore we can achieve accurate measurement of open channel flow in the turbid water through proposed hardware design.

Keywords: Time Difference Method, open channel flux, MSP430, circuit diagram.

1 Introduction

The flow changes of open channel, which is an important mine drainage, is of great significance for monitoring the downhole environment to ensure mine safety. Ultrasonic sensor is more and more used in open channel flow measurement because it can measure velocity and flow more accurately base on acoustic characteristics of the ultrasonic. However, In present stage, the ultrasonic flow meter is of obvious shortcomings. For example, measurement accuracy is not good; Time difference information cannot be extracted in the muddy water because ultrasonic signal noise ratio is lower; Ultrasonic probe is fouled in the sewage so that it cannot be long-term use. In order to improve measurement accuracy of the sensor, the paper strengthen hardware design that the high-performance single-chip is selected as timing control unit and the comparator in the peripheral circuit of single-chip is applied for collecting time difference information. At the same time, square wave pulse generated by single-chip MSP430 is chose as a launch incentive because it is of high accuracy. A band-pass filter produced by OPA603 is chose as a core part of receiver circuit to eliminate the noise and improve sensor performance.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 29–33.
springerlink.com

2 System Hardware Design

The hardware system is consisted of the transmitting circuit and the transmit receive switch, receiver circuit, MCU control circuit and the communication module.

2.1 The Basic Principle of a Transit-Time Ultrasonic Sensor

In order to calculate the speed and volume of the fluid flow, Time Difference Method measures the difference of the downstream propagation time t_1 and countercurrent propagation time t_2 cording to its characteristic that the sound waves spread at different speeds due to the different direction of fluid flow.[1]

2.2 The Transmitting Circuit and the Transmit Receive Switch

The main role of the transmitting circuit of the transmitting circuit is to: provide sufficient power of transmitting pulse, driving ultrasonic transducer. [2]

In this design, the TI company 's UC3708 is used for the power amplifier chip of the transmitting circuit. UC3708 is composed of two groups of independent power amplifying circuit. It features high speed, small delay. When the output current is 500mA, the pressure difference of input and output is 2.5V. It can provide very high output power to meet the demand of the system. It implements low voltage for power supply which is applicable to underground explosion environment.

Integrated photoelectric isolation switch Panasonic AQW214EH is used for the transceiver conversion circuit. Conversion speed of the chip is high. Delay time is 5 nanoseconds. It's high sensitivity. Its front-end LED input current is only 0.9mA. The effect of isolation is good.

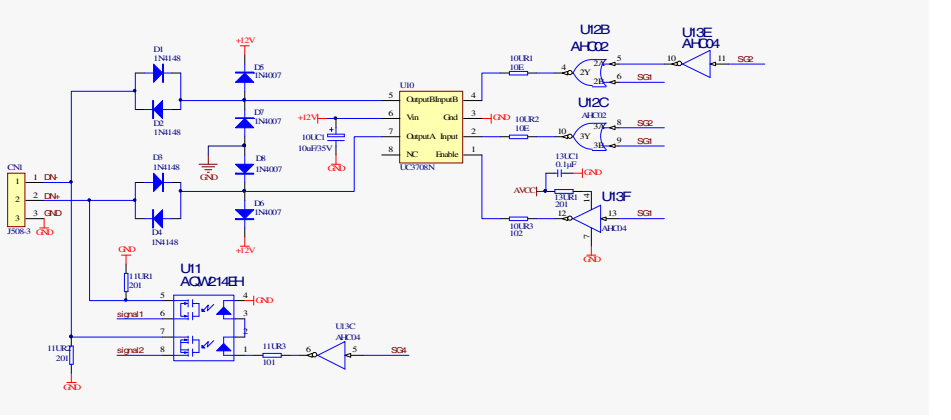


Fig. 1. One of the emission circuit

Emission circuit diagram is shown in Figure 1.

The circuit is consisted of two sets of emission circuit. The two sets of emission structure is basically the same circuit design, but SG4 input port of the downstream transmitter circuit input multi-access a NAND gate. The two sets of transmission circuits are connected to the open channel upstream and downstream transducers. We do the conversion transceiver by change the output level of microcontroller input control signal SG1, SG3 and SG4.

2.3 Receiver Circuit

Open channel water is dirty. Ultrasonic communication process will be influenced by impurities in the water to make the noise, the received ultrasonic contain other frequency noise is not conducive to extract the time difference information. Receiver circuit for signal filtering and amplification is to ensure that the signal to meet the input level requirements of the microcontroller. In addition, it needs to accept the circuit front-end impedance matching, in order to ensure that the transducer output signal has a larger power.[3]

In this design, the impedance matching and filtering module use TI company's integrated operational amplifier chip OPA603. The OPA603 is a high-speed current feedback op amp, the program delay, rise time and fall time is in nanoseconds. OPA603 may constitute a band-pass, low pass, high pass and band-stop filter by connecting capacitors and resistors in different ways, which characteristics of the filter center frequency stability. OPA603 is a low-power device, the maximum power dissipation is only 170mW.

The broadband amplifier to select the ADI company's gain-adjustable operational amplifier AD603. Disconnect the chip's 5 feet and 7 feet its bandwidth is 9MHz, and its gain is 10dB - 50dB. It can amplify 1MHz high-frequency signals, the amplitude reach

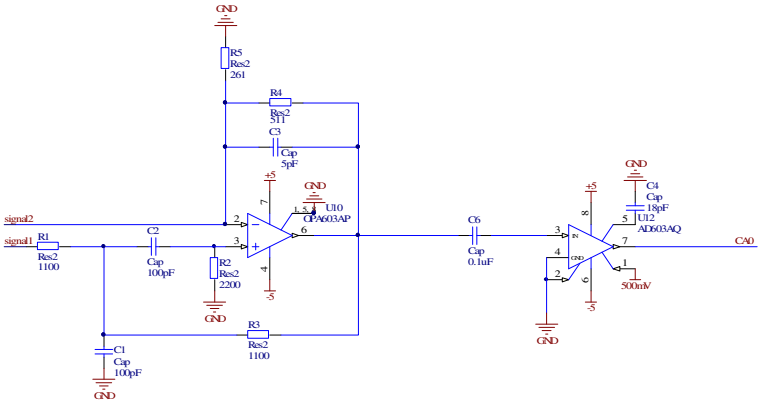


Fig. 2. Receiver circuit diagram

3V to meet the requirements of MCU. AD603 is a low-power device, 5V reference supply, consuming only 125mW. The receiver circuit diagram is shown in Figure 2. The receiver circuit output voltage signal input CA0 pin of microcontroller to generate the rising edge of the digital signal to complete the A / D converter.

2.4 MCU Control Circuit

The design selects TI company's low-power microcontroller MSP430F5438 as the control unit. In high-speed mode, the clock frequency of the MSP430 can achieve 8MHz, can emit high-frequency square wave of 1MHz.[4]

In addition, the ultrasonic transducer distance is about 0.5m, the speed of ultrasound in water is 1450/ s, the transceiver time is 10^{-3}—10^{-4}s, the timing unit of the micro-controller clock source can be up to 8MHz. It meets the system timing accuracy. Minimum measurement accuracy of flow rate of the system is up to 3cm.[5]

The MSP430 has a wealth of peripheral circuits. Its comparator and capture-comparator has the ability to capture the propagation time of ultrasonic. Its USART communication module can be connected to RS485 converter chip to communicate with the host PC.

2.5 The Communication Module

The MCU output level and the host PC's serial port module level are inconsistent. It needs converter chip to convert TTL level to RS485 level which the host PC is needed. In addition, due to the underground environment is complex and serious, it also has a high demand for anti-jamming capability of the converter chip and the optical isolation of the input signal.[6]

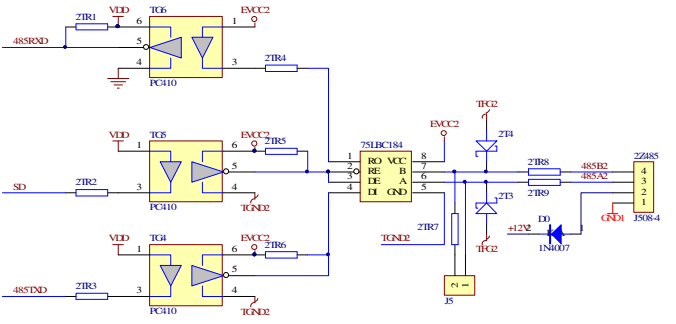


Fig. 3. Communications circuits

The design selected TI company's 75LBC184 as conversion chip. 75LBC184 has strong anti-jamming capability. It still maintains 250kps transmission rate even in a strong electronic interference. Optical isolation of the communication circuit chip selects SHARP company's PC410. The PC410 effect of isolation is good, instantaneous common-mode rejection ratio is high. It can achieve 500V/μs. Its time response is fast as nanoseconds. Communication circuit is shown in Figure 3.

3 Conclusion

In this paper, we measure ultrasonic propagation time to get the time difference information. Calculate the flow by principle of time-difference method. Make full use of the rich functional modules of MSP430, to time the ultrasonic propagation time. Transmitter circuit use the power amplifier chip UC3708 as the core, the output drive signal achieve voltage power amplifier. Receiving circuit use the OPA603 as a narrow bandpass filter, the filter works well to improve the signal to noise ratio. It is applicable to flow measurement in the mixed cropping water quality. Under the control of the microcontroller, the various functional modules of the system orderly work.

In addition, in order to adapt to the strong electromagnetic noise environment, the system uses a large number of opto-isolation devices. Isolation of input and output prevent external electromagnetic noise from going into the system. This is that a lot of the design is not considered. Therefore, the design is more suitable for the underground mine environment. In this design, the using components are all low-power components. It save the energy and more to meet the requirements iof underground mine explosion-proof.

References

1. Qiu, L., Wang, R.: Realization of ultrasonic gas flow measurement system. Transducer and Microsystem Technology 25(1), 47–48 (2006)
2. Huang, Y., Tian, T.: The Design and Realization o New Measure Method on Time Difference Ultrasonic Flow meter. Harbin Engineering University (2007)
3. Li, J., Mo, D.: The design of ultrasonic flow meter based on Improved transit-time method. China Instrumentation 12, 12 (2004)
4. L. Qin: Commonly used module of the MSP430 microcontroller and integrated system instance. Publishing House of Electronics Industry (2007)
5. Shen, J., Yang, Y.: Principles and Practice of 16 ultra-low power MSP430 series microcontroller. Beijing University of Aeronautics and Astronautics Press (2008)
6. RS-422 and RS-485 Application Note. B&B Electronics Mfg. Co. Inc. (1997)

Novel Ideas for Evaluating a Detection Method for the Harmonic and Reactive Currents of a Single-Phase Active Power Filter

Zicheng Li* and Guohai Liu

School of Electrical and Information Engineering, Jiangsu University, Zhenjiang, 212013, China

Abstract. Comparisons and analyses of the simulation waveforms of three typical detection methods for the harmonic and reactive currents of a single-phase active power filter lead to the following question: Does a higher detection precision of a detection method for the harmonic and reactive currents of a single-phase active power filter correspond to a better operation at a dynamic adjusting detecting state? To answer this question, novel ideas for evaluating a detection method for the harmonic and reactive currents of a single-phase active power filter are pro*posed by considering the Fourier series expression of a nonlinear load current with the operation principle and essential function of an active power filter as the theoretical bases.

Keywords: active power filter, detection precision, dynamic response time, fundamental active current, harmonic and reactive currents.

1 Introduction

The widespread use of power electronic appliances in power systems has caused an alarming increase in harmonic pollution. The use of active power filters (APFs) has been considered as an effective solution to compensate for the harmonics generated. Furthermore, the performance of APFs has been greatly influenced by the detection methods for harmonic current which they use [1]–[4].

Akagi et al. proposed the instantaneous reactive power theory (IRPT) (i.e., *p*–*q* theory) in 1983. Since then, IRPT has been continually developed. At present, IRPT-based methods for three-phase APF [5]–[7] are regarded as relatively mature methods [1], [4]. However, these methods require a complex circuitry to realize the transformation. For example, four to six high-precision analog multipliers are used in each phase. This makes the circuit sensitive to component parameter variations [1], [8]. As for a single-phase APF, there is no mature method available. Currently, many detection methods for the harmonic current [1]–[4], [8]–[15] have already been proposed. Although each method has its own strong points, there are still some problems which are difficult to overcome [3], [4]. In addition, the validation of the presence of an applied value in using a detection method for harmonic current is only

* Corresponding author.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 35–40.
Springerlink.com

performed by practical applications. This is because there is a shortage of evaluation standards in theory.

Examples of such methods are the double linear construction algorithm (DLCA) [4], the direct computation method (DCM) [4], and the reference method (RM) [14]. These three are typical methods for the harmonic and reactive currents of a single-phase APF. DLCA can compute the theoretical amplitude of the fundamental active current of a nonlinear load current accurately; hence, it can provide the theoretical sum of the harmonic and reactive currents of the nonlinear load current correctly. The main advantages of DCM are its simplicity, high detection precision, and fast tracking speed [4]. RM is the shortened form of the method based on IRPT. It has an excellent detection performance and can be regarded as a reference method for evaluating whether *a detection method for the harmonic and reactive currents of a single-phase active power filter* has applied value or not [14]. This paper aims to address the lack of a theoretical evaluation standard for the detection methods for harmonic current. Novel ideas for evaluating *a method* are proposed based on comparisons and analyses of the simulation waveforms of DLCA, DCM, and RM.

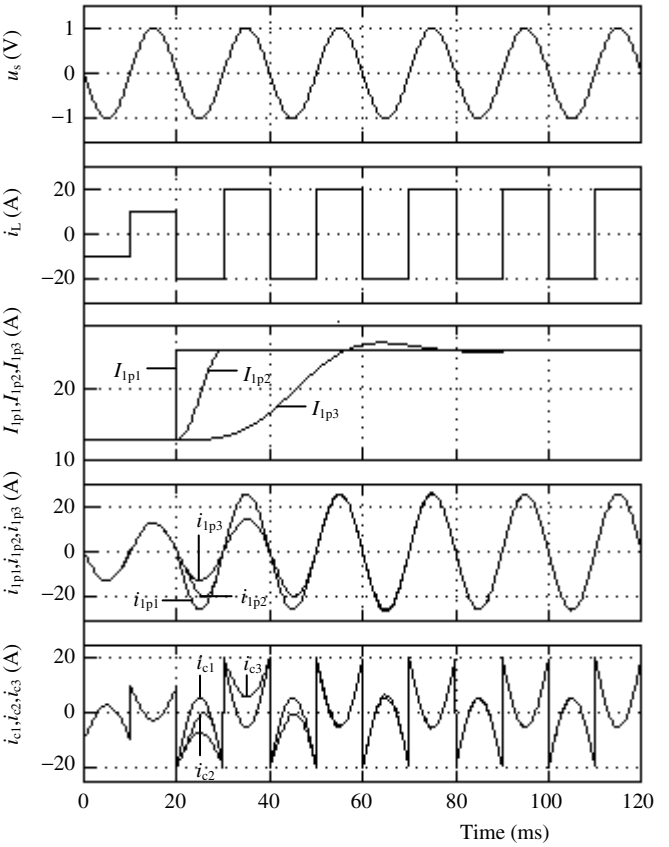


Fig. 1. Simulation waveforms of DLCA, DCM, and RM when a nonlinear load current linearly increases from 10 to 20 A in a cycle.

2 The Proposed Question

The simulation waveforms with a nonlinear load current linear increases from 10 to 20 A in a cycle are shown in Fig. 1, where u_s is the power supply voltage whose

amplitude is 1 V, and i_L is the nonlinear load current. Further, I_{p1}, i_{p1}, and i_{c1} detected by DLCA; I_{p2}, i_{p2}, and i_{c2} detected by DCM; and I_{p3}, i_{p3}, and i_{c3} detected by RM are the amplitude of the fundamental active current, the fundamental active current, and the sum of the harmonic and reactive currents, respectively. The low-pass filter of the RM adopts a fourth-order low-pass Butterworth analog filter whose cutoff frequency is 25 Hz [14], so its detection waveforms will not distort. When i_L is in steady state, the three methods will operate at a steady-detecting state. When they do so (see 0–20 ms and 90–120 ms in Fig. 1), I_{p1}, I_{p2}, and I_{p3}; i_{p1}, i_{p2}, and i_{p3}; and i_{c1}, i_{c2}, and i_{c3} are graphed superposing one another, respectively. DLCA can accurately compute the theoretical sum of the harmonic and reactive currents of i_L, so i_{c1} is the sum of the harmonic and reactive currents of i_L in theory. Thus, at this steady-detecting state, the three methods can accurately detect the sum of the harmonic and reactive currents of i_L. When i_L linearly increases from 10 to 20 A in a cycle (see 20–40 ms in Fig. 1), the three methods will simultaneously operate at a dynamic adjusting detecting state. When they do so (see 20–90 ms in Fig. 1), the detection precision (DP) of DLCA is higher than that of DCM, and the DP of DCM is higher than that of RM. The dynamic response time (DRT) of DLCA (0 ms) is shorter than that of DCM (10 ms), and the DRT of DCM (10 ms) is shorter than that of RM (about 50 ms). For the three methods, their DP and DRT are correlated with each other: the higher the DP is, the shorter the DRT is, and vice versa.

Based on the comparisons and analyses described above, the following question arises: Does a higher DP of *a method* correspond to a better operation of the method at a dynamic adjusting detecting state?

We assume that the answer to this query is yes. When the three methods operate at the dynamic adjusting detecting state, the DP of RM is low, and it is in fact lower than that of any one of DLCA and DCM as observed at 20–90 ms in Fig. 1. However, RM has an excellent detection performance. It can be regarded as a reference method which evaluates whether *a method* has an applied value or not. Very possibly, the assumed answer (yes) is incorrect because the DP of RM is much lower than that of DLCA or DCM.

3 Novel Ideas for Evaluating *a Method*

According to *Fourier* series, a nonlinear load current can be expressed as

$$\begin{aligned} i_L(t) &= I_p^* \sin(\omega t) + I_q^* \cos(\omega t) \\ &+ \sum_{n=2}^{\infty} \left[I_{np} \sin(n\omega t) + I_{nq} \cos(n\omega t) \right] = i_p^*(t) + i_c^*(t). \end{aligned} \tag{1}$$

Where

$$i_p^*(t) = I_p^* \sin(\omega t), \tag{2}$$

$$i_c^*(t) = I_q^* \cos(\omega t) + \sum_{n=2}^{\infty} \left[I_{np}^* \sin(n\omega t) + I_{nq}^* \cos(n\omega t) \right], \tag{3}$$

$$I_p^* = \frac{2}{T}\int_0^T i_L(t)\sin(\omega t)dt = \frac{2}{T}\int_{t-T}^{t} i_L(t)\sin(\omega t)dt. \quad (4)$$

In these equations, $i_p^*(t)$, $i_c^*(t)$, and I_p^* are the theoretical fundamental active current, the theoretical sum of the harmonic and reactive currents, and the theoretical amplitude of the fundamental active current of $i_L(t)$, respectively.

Note that the precondition of (1) is that $i_L(t)$ should be a periodic current.

The operation principle of a single-phase APF is shown in Fig. 2. It consists of a detection circuit for harmonic and reactive currents, a current-tracking control circuit, a driving circuit, and a main circuit. In Fig. 2, $i_s(t)$ is the power supply current, $i_L(t)$ is the load current, $i_c'(t)$ is the actual compensation current provided by APF, and $i_c(t)$ is the sum of the harmonic and reactive currents of $i_L(t)$. The variable $i_c(t)$ is detected by *a method* which realizes the function of the detection circuit for harmonic and reactive currents. Generally, the method should detect the I_p which is the amplitude of the fundamental active current of $i_L(t)$. Thus, the $i_p(t)$ which is the fundamental active current of $i_L(t)$ is $I_p\sin(\omega t)$, and the $i_c(t)$ which is the sum of the harmonic and reactive currents of $i_L(t)$ is $i_L(t)$–$I_p\sin(\omega t)$. The relation between $i_p(t)$ and $i_c(t)$ is given by

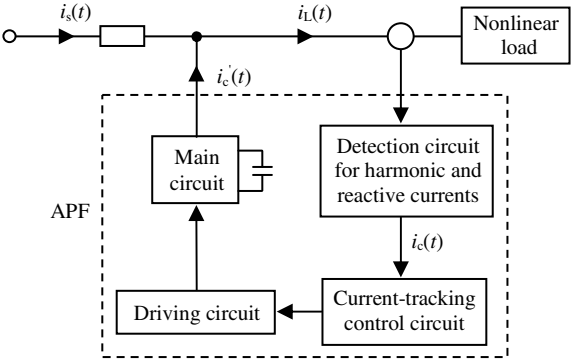


Fig. 2. Operation principle of a single-phase APF

$$i_p(t) = i_L(t) - i_c(t) = I_p\sin(\omega t). \quad (5)$$

Under ideal circumstances, $i_c'(t) = i_c(t)$, $i_c(t) = i_c^*(t)$. On the basis of *Kirchhoff*'s Current Law, and combining (5), (1), and (2), $i_s(t)$ can be expressed as

$$\begin{aligned} i_s(t) = i_L(t) - i_c'(t) = i_L(t) - i_c(t) = i_p(t) = I_p\sin(\omega t) \\ = i_L(t) - i_c^*(t) = i_p^*(t) = I_p^*\sin(\omega t). \end{aligned} \quad (6)$$

APF can be regarded as a current-controlled current source. In theory, APF should provide $i_c^*(t)$, that is, $i_c'(t) = i_c^*(t)$, and the power supply should provide $i_p^*(t)$, that is, $i_s(t) = i_p^*(t)$. Therefore, the function of *a method* is to distinguish $i_c^*(t)$ instantaneously and accurately; that is, $i_c(t)$ should equal $i_c^*(t)$. Based on this understanding, the higher the DP of *a method*, the better should such method ideally be.

This assumption is not completely correct. The function of APF is to compensate for $i_c^*(t)$, dynamically, instantaneously, and continuously. However, this is only its easily recognizable function. The essential function of APF is to make $i_s(t)$ a

sinusoidal current with the same frequency and phase as the power supply voltage $u_s(t)$ [$=U_m\sin(\omega t)$]. This is done to achieve the purpose of improving power quality.

When $i_L(t)$ is in steady state, *a method* will operate at a steady-detecting state. When it does so, the precondition of (1) is satisfied, so the I_p^* computed by (4) is precisely a constant. Therefore, $i_p^*(t)$ [$=I_p^*\sin(\omega t)$] is a sinusoidal current with the same frequency and phase as $u_s(t)$. Thus, this steady-detecting state can only accurately detect $i_c^*(t)$, and then make $i_c'(t) = i_c^*(t)$ and $i_s(t) = i_p^*(t)$. Through these, APF can realize its essential function to make $i_s(t)$ a sinusoidal current with the same frequency and phase as $u_s(t)$. Accordingly, at this steady-detecting state, the higher the DP of *a method*, the better is the method.

When $i_L(t)$ is in varying state, *a method* will simultaneously operate at a dynamic adjusting detecting state. When it does so, because the precondition of (1) is not satisfied, $i_L(t)$ cannot be expressed as (1) in theory, and I_p^* should not be computed by (4) but by DLCA. I_p^* is affirmatively not a constant as the I_{p1} of 20–90 ms in Fig. 1. Thus, $i_p^*(t)$ is not a sinusoidal current. Therefore, at this dynamic adjusting detecting state, making $i_c'(t) = i_c^*(t)$ and $i_s(t) = i_p^*(t)$ deters APF from realizing its essential function to make $i_s(t)$ a sinusoidal current with the same frequency and phase as $u_s(t)$. Therefore, there is a need to search for an approximate way based on logic which will make $i_s(t)$ an approximate sinusoidal current with the same frequency and phase as $u_s(t)$. On the basis of $i_s(t) = i_p(t)$, if $i_s(t)$ is an approximate sinusoidal current, then $i_p(t)$ should also be an approximate sinusoidal current. Therefore, $i_c(t)$ should assure that $i_p(t)$ [$=i_L(t)-i_c(t)$] would be an approximate sinusoidal current. On one hand, the closer the approximated sinusoidal current $i_p(t)$ is to a real sinusoidal current, the better the method in theory. On the other hand, the higher the DP of *a method*, the better is the method. However, both statements are ambivalent: the closer to a sinusoidal current $i_p(t)$ is, the lower the DP of the method, and vice versa. Accordingly, *a method* should make $i_p(t)$ an approximate sinusoidal current with a DP that is high enough. As much as possible, both parameters should be given attention. In addition, when $i_p(t)$ [$=I_p\sin(\omega t)$] is made to be an approximate sinusoidal current, the optimal scheme is that I_p should smoothly track I_p^*. Consequently, when *a method* operates at a dynamic adjusting detecting state, concluding that *a method* with a higher DP is better is incorrect.

4 Conclusions

When *a method* operates at a steady-detecting state, the higher its DP, the better is the method. However, when it operates at a dynamic adjusting detecting state, concluding that *a method* with a higher DP is better is incorrect. The amplitude of the fundamental active current of a nonlinear load current detected by the method should smoothly track the theoretical amplitude. The method should guarantee that the detected fundamental active current of the load current is an approximate sinusoidal current and that its DP should be high enough. As much as possible, both parameters should be given attention. These insights can be considered as novel ideas for evaluating *a method*.

Acknowledgments. This work was supported by a Project funded by the Priority Academic Program Development of Jiangsu Higher Education.

References

1. Zhou, L., Li, Z.: A novel active power filter based on the least compensation current control method. IEEE Trans. Power Electron. 15, 655–659 (2000)
2. Lin, H.C.: Intelligent neural network-based fast power system harmonic detection. IEEE Trans. Ind. Electron. 54, 43–52 (2007)
3. Wang, Q., Wu, N., Wang, Z.: A neuron adaptive detection approach of harmonic current for APF and its realization of analog circuit. IEEE Trans. Instrum. Meas. 50, 77–84 (2001)
4. Li, Z., Sun, Y.: A new compensation current real-time computing method for power active filter based on double linear construction algorithm. Sci. China Ser. E Technol. Sci. 49, 485–512 (2006)
5. Akagi, H., Kanazawa, Y., Nabae, A.: Instantaneous reactive power compensators comprising switching devices without energy storage components. IEEE Trans. Ind. Appl. 20, 625–630 (1984)
6. Soares, V., Verdelhq, P., Marques, G.D.: An instantaneous active and reactive current component method for active filters. IEEE Trans. Power Electron. 15, 660–669 (2000)
7. Peng, F.Z., Akagi, H., Nabae, A.: A study of active power filters using quad-series voltage source PWM converters for harmonic compensation. IEEE Trans. Power Electron. 5, 9–15 (1990)
8. Tepper, J.S., Dixon, J.W., Venegas, G., Morh, L.: A simple frequency-independent method for calculating the reactive and harmonic current in a nonlinear load. IEEE Trans. Ind. Electron. 43, 647–654 (1996)
9. Luo, S., Hou, Z.: An adaptive detection method for harmonic and reactive Currents. IEEE Trans. Ind. Electron. 42, 85–89 (1995)
10. Karimi-Ghartemani, M., Mokhtari, H., Iravani, M.R., Sedighy, M.: A signal processing system for extraction of harmonics and reactive current of single-phase systems. IEEE Trans. Power Delivery 19, 979–986 (2004)
11. Li, H., Zhuo, F., Wang, Z., Lei, W., Wu, L.: A novel time-domain current-detection algorithm for shunt active power filters. IEEE Trans. Power Syst. 20, 644–651 (2005)
12. Mazumdar, J., Harley, R.G., Lambert, F.C., Venayagamoorthy, G.K.: Neural network based method for predicting nonlinear load harmonics. IEEE Trans. Power Electron. 22, 1036–1045 (2007)
13. Mazumdar, J., Harley, R.G.: Recurrent neural networks trained with backpropagation through time algorithm to estimate nonlinear load harmonic currents. IEEE Trans. Ind. Electron. 55, 3484–3491 (2008)
14. Li, Z., Sun, Y., Liu, G., Li, F., Yang, J.: Simulation research on a referenced method for harmonic current detection of single-phase circuit. Low Voltage Apparatus 17, 30–33 (2008)
15. Mojiri, M., Karimi-Ghartemani, M., Bakhshai, A.: Processing of harmonics and interharmonics using an adaptive notch filter. IEEE Trans. Power Delivery 25, 534–542 (2010)

Mapping Cores to Network-on-Chip in Digital Circuit Design by Using Tabu Search Approach

Zhongliang Pan and Ling Chen

School of Physics and Telecommunications Engineering, South China Normal University, Guangzhou 510006, China
panz@scnu.edu.cn

Abstract. The network-on-chip(NoC) is one of important aspects in digital circuit design, it attracted more and more attentions in electronic engineering. The NoC has been proposed as a promising solution to act as communication architecture because the number of cores in the systems-on-chip(SoC) is increased greatly. A lot of work needs to be carried out in the design of NoC, such as the determinations of the NoC communication backbone and communication protocols, the allocation of cores to various available nodes in NoC topology, that is the core mapping. The task of core mapping is to determine the topological placement of cores onto the nodes in NoC topology structure, and choose a path for each pair of IP cores. In this paper, the NoC structure with mesh topology is considered, a new approach is presented for the core mapping of network-on-chip by using Tabu search algorithm. The evolution operations in genetic algorithms are added to the Tabu search process to enhance the computation efficiency. The experimental results show that the time being needed to obtain the core mapping scheme by the approach proposed in this paper is lower than conventional genetic algorithm.

Keywords: Digital circuits, network-on-chips, core mapping, Tabu search algorithm, genetic algorithms.

1 Introduction

With the advances in semiconductor technology, the integration of complex systems-on-chip(SoC) has become feasible, the SoC can offer a tremendous amount of computational functions. The network-on-chip(NoC) has been proposed as a promising solution to act as communication architecture owing to the increasing number of cores in the SoC. The NoC makes use of packet switching technique to replace the shared buses and dedicated wires in the SoC. The infrastructure of NoC is the combination of various elements (e.g., switches, links) and protocols (e.g., routing) that determine the communication architecture and modes [1,2]. A lot of work needs to be done in the design of NoC, such as the structure designs, interconnect, quality-of-service, core mapping, etc.

The structure designs of hierarchical NoC were studied. Bourduas [3] proposed a hybrid NoC structure that partitions a large 2D-mesh into a lot of smaller sub-meshes.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 41–47.
springerlink.com

Winter [4] discussed the realization methods of clustered and hierarchical 2D-meshes, and analyzed their performance by using the cycle accurate simulations.

For the interconnect of NoC, Stephan [5] discussed the ring network that has a slotted ring architecture to reduce hop counts and latencies of global traffic. Murali [6] studied the design issue of synthesizing the most power efficient NoC interconnect, and proposed a method to guarantee optimum throughput.

For the core mapping, Jingcao [7] studied the energy-aware and performance-aware mapping in NoC, and proposed a branch-and-bound algorithm to solve the problem of core mapping. Zhonghai [8] studied the method to tackle the mapping problem in 2D mesh structure by using simulated annealing. Ling [9] studied the run-time mapping on multiprocessor systems-on-chip (MPSoC), where the communication structure of the MPSoC is same as the network-on-chip.

In this paper, a new approach is presented for the core mapping of network-on-chip by using Tabu search algorithm, the evolution operations of genetic algorithms are added to the Tabu search process to enhance the computation efficiency.

2 Core Mapping by Tabu Search Approach

One of task in the design of NoC is the core mapping, which is to determine the topological placement of the IP cores onto different nodes in the NoC topology, and select one path for each pair of IP cores. The objective of the core mapping is to get the best energy/performance trade-off. In this paper, the mesh topology is used when we design the method for the core mapping.

2.1 Core Mapping Problem

The communication among the cores in a NoC can be represented by a communication graph (or called as core graph). The core graph G(V, E) is a directed graph. The V is the set of vertexes, the E is the set of edges. Each vertex, such as $v_i \in V$, represents a core. Each directed edge represents the communication between two cores. For example, (v_i, v_j) is the edge between the cores (vertexes) v_i and v_j, the (v_i, v_j) is denoted as e_{ij}, the $e_{ij} \in E$. The weight of an edge can be defined to represent the bandwidth requirement between the cores (vertexes) v_i and v_j. The w_{ij} is used to denote the weight of the edge e_{ij}.

For a NoC, its topology structure can be represented by the following NoC topology graph. The NoC topology graph T(U, M) is a directed graph. The U is the set of nodes, the M is the set of edges. Let u_i represent a node in the NoC topology structure, the $u_i \in U$. Every directed edge represents the communication between two nodes in the NoC topology structure. For example, (u_i, u_j) is the directed edge between the nodes u_i and u_j. The (u_i, u_j) can be denoted as h_{ij}, the $h_{ij} \in M$. The s_{ij} is used to denote the weight of the edge h_{ij}.

The mapping of the cores in core graph G(V, E) onto the nodes in NoC topology graph is defined by the following one-to-one function. $f(x)$: $V \rightarrow U$, such that for arbitrary $v_i \in V$, there is a u_j that is corresponding to the v_i, where $u_j \in U$, that is $f(v_i)=u_j$.

In general, some constraints can be added to mapping f(x). For two nodes u_s and u_t in NoC topology graph, suppose the p(s,t) represent the routing path between nodes u_s and u_t. For a link (i.e. an edge) which is denoted as L between the two nodes that are neighboring directly in NoC topology graph, define a function g(L, p(s,t)) as follows: the g(L, p(s,t))=1 when the edge L belong to path p(s,t), that is the edge L is one of the edges in path p(s,t). Let $B(L)$ denote the bandwidth of the link L in NoC topology graph. If the energy is considered, then the following model of Problem 1 is used. The energy metric E_b for the energy model of network routers is defined by each bit of data is transferred through the router. The E_b is expressed by equation $E_b=E_{sb}+E_{Lb}$.

The E_{sb} denotes the energy consumption of the switch, i.e., the energy consumption on both the interconnect wires and the logic gates inside of switch. The E_{Lb} denotes the energy consumption of the link between nodes in NoC. For two nodes u_i and u_j in NoC, suppose the u_i is source node, the u_j is destination node. When the nodes u_i and u_j are not directly connected, the source node u_i transmits data to destination node u_j by several other nodes on the path from nodes u_i to u_j. Let the number of these other nodes is k, then the average energy consumption for sending one bit of data from nodes u_i to u_j can be computed by the following equation.

$$E_b(u_i, u_j) = k \cdot E_{sb} + (k-1) \cdot E_{Lb}$$

Therefore, the total energy consumption can be given by

$$\sum_{v_i, v_j \in V} w_{ij} \cdot E_b(f(v_i), f(v_j))$$

The core mapping problem under energy consumption can be expressed as follows:

Problem 1

f(x): $V \rightarrow U$, satisfies $f(v_i)=u_j$ for $\forall v_i \in V$, $\exists u_j \in U$.
such that:

$$\text{Minimize} \quad \psi = \sum_{v_i, v_j \in V} w_{ij} \cdot E_b(f(v_i), f(v_j)) \tag{1}$$

and satisfies: (a) For arbitrary $v_i \in V$ and $v_j \in V$, $v_i \neq v_j$, the equation $f(v_i) \neq f(v_j)$ is true. (b) For arbitrary link L, $\sum_{v_i, v_j \in V} g(L, p(f(v_i), f(v_j))) \cdot w_{ij} \leq B(L)$.

In Problem 1, the following two aspects are taken into account: the communication bandwidth requirements, and the distance between the source node and destination node. The objective is to minimize the total communication energy consumption of cores mapping scheme. In the Section 2.2, we use the Tabu search approach to solve the problem of core mapping.

2.2 Core Mapping by Using Tabu Search

The Tabu search is an iterative algorithm for solving optimization problems, it can store the history information about the past search steps, and makes use of the history information to produce new feasible solutions in search space [10]. Because the history information about feasible solutions is recorded and is used, the Tabu search

algorithm can avoid handling the similar solutions repeatedly. The Tabu search algorithm consists of following steps for the minimization of a given function.

Algorithm 1

Step 1. Let the x_0 be an initial feasible solution. Set $x_c := x_0$, where x_c is the current feasible solution.

Step 2. For the current feasible solution x_c, a set S of feasible solutions is produced by a predefined criteria, which can be obtained by performing the simple modification to the current solution x_c. This modification is called as a move.

Step 3. A tabu list stores all the moves that cannot be applied for the current feasible solution x_c. These moves stored in the tabu list are those moves being performed recently according to the predefined criteria.

Step 4. Make use of the tabu list, the fittest feasible solution x' is selected from the set S. The x' is chosen as the next feasible solution.

Step 5. If the predefined termination conditions are satisfied, then stop the algorithm, else set $x_c := x'$, go to Step 2.

A merit of the Tabu search algorithm is that it can get high computation speed. The main disadvantage of the algorithm is that it has a relative dependence on the initial solution. The better is the initial solution, the more accurate the solution obtained by Tabu search is. Therefore, in the following, the evolution operations of genetic algorithms are employed to overcome its shortcomings.

The genetic algorithm (GA) starts with a randomly generated initial population being consisted of a lot of individuals, where every individual is evaluated by means of a fitness function [11]. Secondly, the new populations are produced by using genetic operations such as selection, crossover and mutation. The process is carried out repeatedly until the better individuals are obtained. In the following, the Tabu search with evolution is given, it is illustrated in the Algorithm 2.

Algorithm 2

Step 1. For the initial feasible solution x_0, produce N feasible solutions (i.e., individuals) by using the Tabu search algorithm given in the Algorithm 1. Create the initial population being made up of the N individuals.

Step 2. Evaluate the fitness of each individual in the current population.

Step 3. Create a new population by using evolution operations: selection, crossover and mutation.

Step 4. Choose m individuals randomly from the current population. Denote the m individuals as $I_1, I_2, \cdots, I_m$. Where the m is a positive constant.

Step 5. Make the m individuals chosen in the Step 4 as the initial feasible solutions, produce m new individuals by using the Tabu search algorithm given in Algorithm 1. Denote the m new individuals as $d_1, d_2, \cdots, d_m$.

Step 6. Use the individuals $d_1, d_2, \cdots, d_m$ to replace the individuals $I_1, I_2, \cdots, I_m$ in the current population. Form a new current population.

Step 7. If the termination condition is satisfied, then stop the algorithm and return the best individual in the current population, else go to Step 2.

For the core mapping of NoC, the coding of an individual is as follows. Let the nodes in NoC topology graph are u_1, u_2, ···, u_n, respectively. The coding of an individual is the vector that is made up of each value corresponding to every u_i (i=1, 2, ···, n). For example, the coding of the individual is (7 6 1 5 3 8 9 4 2). The fitness function of an individual is defined by the $1/\psi$, where the ψ is given in equation (1).

2.3 Experimental Results

The approach in this paper for mapping cores to network-on-chip by using Tabu search has been implemented in C++, and a lot of experiments have been carried out. These experiments are performed on a personal computer with 3.0GHz and 512MB memory under Windows operation system. Here, we give the experimental results about the minimization of energy consumption for core mapping. The approach in this paper is applied to a real Multi-Media System (MMS) [12]. The MMS system is an integrated video and audio system, which consists of an MP3 audio encoder, an MP3 audio decoder, an H.263 video encoder and an H.263 video decoder. The core graph of the MMS is given in the Fig.1. The weights of the edges in the graph are in MB/s. The cores in the communication graph of the MMS are mapped to a NoC topology with 25 nodes.

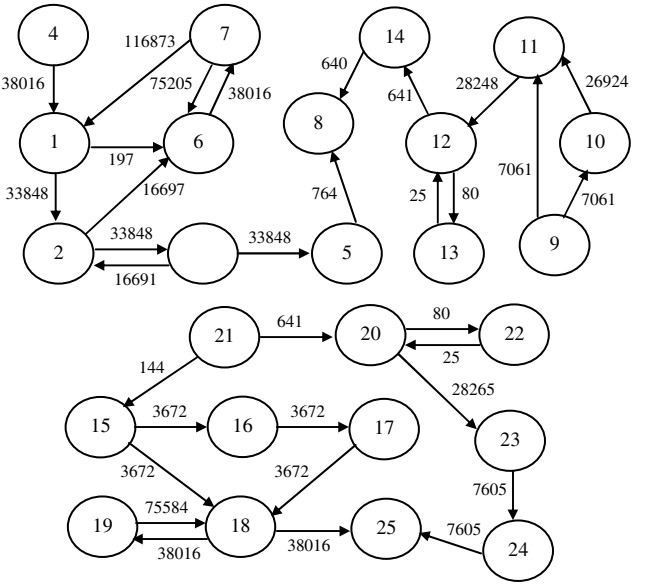


Fig. 1. The core graph of a multi-media system

In these experiments, the following parameters are used in the Algorithm 2: The maximal evolution generation is set as 1000; the population size (i.e., the number of individuals in populations) is set as 30, i.e., N=30; the crossover rate is set as 0.90; the mutation rate is set as 0.002; the value of parameter m is set to 8. We have compared the approach in this paper (i.e. Algorithm 2) with conventional genetic algorithm

(CGA). The main parameters that we have used for CGA are as follows: The population size is 30, the selection operation uses roulette wheel selection scheme; the two points crossover is used in crossover operation. The crossover probability of two individuals is chosen as 0.90, the mutation rate is set as 0.005.

For the core mapping of MMS system, the core mapping scheme is obtained in 102 generations by using the Algorithm 2, the core mapping scheme is obtained in 183 generations by using conventional genetic algorithm, where the numbers of generations are the 10 runs of each algorithm. The total time needed by Algorithm 2 is lower than conventional genetic algorithm.

3 Conclusions

The network-on-chip has been studied intensively in the area of electronic engineering, it makes use of packet switching to replace the shared buses and dedicated wires in SoC. One of the design aspects of NoC is the core mapping. In this paper, a new approach is presented for the core mapping of NoC by using Tabu search, where the mesh topology is considered. Some work needs to be done in the future for core mapping such as general topology structure and constraint conditions, the selection of the fitting values for the parameters in Tabu search algorithm, etc.

Acknowledgments. This work were supported by National Natural Science Foundation of China (No.61072028), Guangdong Province & Chinese Ministry of Education Cooperation Project of Industry, Education and Academy (No. 2009B090300339).

References

1. Hsiang, K.Y., Ming, Y., Sertac, A., Jonathan, C.H.: CNoC High-radix clos network-on-chip. IEEE Trans. on CAD 30(12), 1897–1910 (2011)
2. Elena, K., Vassos, S., Theocharis, T.: Intelligent hotspot prediction for network-on-chip-based multicore systems. IEEE Trans. on CAD 31(3), 418–431 (2012)
3. Bourduas, S., Zilic, Z.: A hybrid ring/mesh interconnect for network-on-chip using hierarchical rings for global routing. In: International Symposium on Networks-on-Chip, pp. 195–204 (2007)
4. Winter, M., Prusseit, S., Gerhard, P.: Hierarchical routing architectures in clustered 2D-mesh networks-on-Chip. In: International SoC Design Conference, pp. 388–391 (2010)
5. Stephan, B., Zeljko, Z.: Modeling and evaluation of ring-based interconnects for network-on-chip. Journal of Systems Architecture 57(1), 39–60 (2011)
6. Murali, S., Atienza, D., Meloni, P., Carta, S.: Synthesis of predictable networks-on-chip-based interconnect architectures for chip multiprocessors. IEEE Trans. on Very Large Scale Integration (VLSI) Systems 15(8), 869–880 (2007)
7. Jingcao, H., Radu, M.: Energy- and performance-aware mapping for regular NoC architectures. IEEE Trans. on CAD 24(4), 551–562 (2005)
8. Zhonghai, L., Lei, X., Jantsch, A.: Cluster-based simulated annealing for mapping cores onto 2D mesh networks on chip. In: IEEE Workshop on Design and Diagnostics of Electronic Circuits and Systems, pp. 1–6 (2008)

9. Ling, C., Ogras, U.: Energy- and performance-aware incremental mapping for networks on chip with multiple voltage levels. IEEE Trans. on CAD 27(10), 1866–1879 (2008)
10. James, T., Rego, C., Glover, F.: Multistart Tabu search and diversification strategies for the quadratic assignment problem. IEEE Trans. on Systems, Man and Cybernetics, Part A: Systems and Humans 39(3), 579–596 (2009)
11. Yin, Y.S., Kin, C.C.: A genetic algorithm that adaptively mutates and never revisits. IEEE Trans. on Evolutionary Computation 13(2), 454–475 (2009)
12. Jingcao, H.: Energy-aware mapping for tile-based NoC architectures under performance constraints. In: Asia South Pacific Design Automation Conference, pp. 233–239 (2003)

Voltage Transient Signal De-noising Based on Wavelet Decomposition Level

Yuanyuan Fan[1], Yingjun Sang[1,2], Qingxia Kong[1], Fei Huang[1], Qi Chen[1], and Bin Liu[1]

[1] Huaiyin Institute of Technology, Meichengstr.1,
223003, Huaian, China
[2] Electrical Engineering School, Southeast University, Sipailoustr.2
210096, nanjing, China
sangyingj@163.com

Abstract. In this article, the wavelet analysis method is used to detect and analyze voltage sag signal contaminated by white noise. A new threshold selection method based on the wavelet decomposition layer is proposed, in view of the characteristic of noise coefficients which decrease when the scale increases. This method can obtain the superior wavelet coefficients estimation through adjusting two adjustable parameters. This method has the better ability for reducing noise interference.

Keywords: Transient, wavelet transform, threshold de-noising.

1 Introduction

Electric power quality issues have captured increasing attention in electric engineering in recent years, for its degrading the performance and efficiency of customer loads, especially power electronics loads [1,2]. Power quality monitoring (PQM) will be an effective means to detect the power disturbances, such as voltage sags, oscillations, pulses. However, the signal from sensor is often contaminated by noise, so it's necessary to filter the signals with an efficient de-noising method.

Threshold method is common and important for wavelet de-noising, some researches have been carried out. Donoho and Johnstone proposed hard and soft threshold method, which can effectively reduce the noise interference with a certain threshold selection, but each has the drawbacks respectively, so a new threshold selection method is proposed in this article, which is appropriate for the voltage sag and other nonstationary situations of power quality monitoring system. The simulation results proved that the de-noising effect of proposed algorithm is better, compared with the other three algorithms.

2 Discrete Wavelet Transform (DWT)

Discrete wavelet transform was developed by Mallat from fast algorithm based on the conjugate quadratic filters (CQF) [3,4], J-level wavelet decomposition can be computed as follows:

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 49–54.
springerlink.com

$$c_{j,k} = \sum_n h_{n-2k} c_{j-1,n} \tag{1}$$

$$d_{j,k} = \sum_n g_{n-2k} c_{j-1,n} \tag{2}$$

Where $c_{j,k}$, $d_{j,k}$ are scale and wavelet coefficients derived from the projection of the signal on the space of scale and wavelet functions respectively. H_n and g_n are the low-pass and high-pass filters respectively, corresponding to the selected wavelet basis, the reconstruction algorithm is shown as following formula:

$$c_{j-1,k} = \sum_n \left[h_{k-2n} c_{j,n} + g_{k-2n} d_{j,n} \right], j = J, J-1, \cdots, 1 \tag{3}$$

3 Voltage Transient Signal De-noising Based on Wavelet Decomposition Level

Voltage sag is the most common type of voltage disturbances of power distribution system, and is commonly caused by the system short-circuit fault, large motor starting, transformer or capacitor switching, therefore monitoring and detection of voltage sag are necessary to improve the power quality of system.

In literature[5] wavelet singularity detection was used for multi-resolution analysis of transient power quality disturbance signals, and got the accurate breaking position information in time domain, literature [6] described the use of perturbation detection method with 3 Daubechies wavelet, and pointed that Daub4 wavelet is suitable for detecting short-term abrupt phenomena with wavelet multi-resolution analysis, we choose Daub4 wavelet as mother wavelet.

Voltage sag simulation circuit model shown in Fig. 1, the model voltage sag is caused by a ground fault, adjusting short-circuit resistance value of the "Three-Phase Fault" module can control the amplitude of the sag and regulating the short-circuit time can control the voltage sag duration.

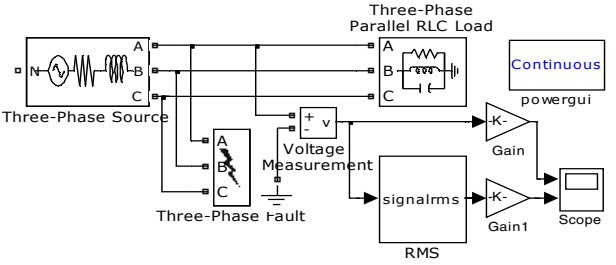


Fig. 1. Simulation model of voltage sag

3.1 Threshold-Based Wavelet De-noising

The noise in power quality monitoring system, is usually the Gaussian white noise, and the noisy signal is basically of the following form:

$$y(t) = f(t) + \alpha\varepsilon(t) \tag{4}$$

Where $\varepsilon(t)$ is the standard Gaussian white noise, α is the noise level, $f(t)$ is original pure signal. According to Donoho and Johnstone proposed threshold approach [5], there are two different threshold function, the hard and soft threshold.

The hard threshold function is not a continuous function, and thus the signal reconstruction process will appear oscillations. he soft threshold function is continuous, but when $|w| > T$, which is a constant deviation, and can impact the estimated degree of signal reconstruction, it will result in the loss of useful information, and bring a great distortion to reconstructed signal. Soft and hard threshold function can be expressed in Fig. 2

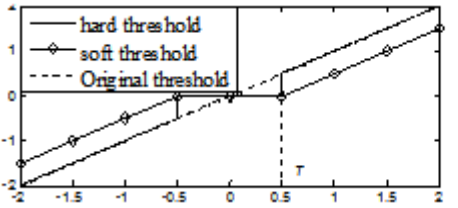


Fig. 2. Soft and hard threshold function

3.2 Subband Adaptive Threshold Algorithm

Other ways to overcome the problem of soft and hard threshold algorithm are estimating level-dependent thresholds. We know that the amplitudes of signal and noise have different characteristics with the increase of decomposition scale, the amplitude of noise displays a reduced trend when the decomposition scale is increased, whereas the amplitude of the useful signal are not related to the change of level.

So Zhongsheng Wang proposed the improved algorithm to solve the fault diagnosis problem of rotor engine [7], S. Poornachandra proposed the S-median algorithm to solve the problem of ECG signal noise reduction [8], however, The two algorithms have only one adjustable parameter to adjust the whole range of thresholds, have not the fine-tuning capability, and the algorithm does not adapt to power quality fault signal, so this paper presents a new level-dependent subband adaptive noise reduction algorithm, Subband adaptive algorithm, the threshold value function is shown as follows:

$$T = \frac{\sigma\sqrt{2\log(N)}}{S_{L,K} + b}, \quad S_{L,K} = 2^{nL/(nL-K)} \tag{5}$$

Where the original universal threshold function is improved, $S_{L,K}$ is associated with the amount of wavelet decomposition level, L is the amount of wavelet decomposition level, and K is the current level number where the threshold is calculated (for example, of 5 levels wavelet transform , L=5, K=1,2,3,4,5). On the one hand, n, as a fine-tuning factor, is used to adjust the threshold downward trend from low to high

decomposition level, with the range $n>1$; On the other hand, b, as a coarse adjusting factor, is used to adjust the whole range of thresholds.

The variation regulation is shown in Fig. 3, when the value of n becomes larger, the threshold curve becomes flatter, the difference between the n level and the first level becomes smaller, and the threshold at $n=\infty$ is almost a constant equal to the value of the first level, which just like the hard threshold algorithm. when $b=0$, the threshold value decreases from the value nearby 1.2, and when $b=4$, the threshold value decreases from the value nearby 0.4, so b can be used to change the threshold range greatly.

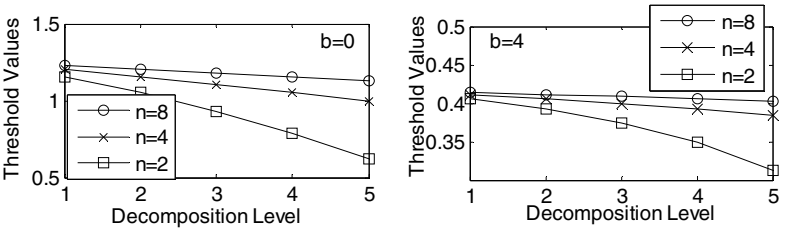


Fig. 3. Subband adaptive algorithm threshold variation regulation with parameter variation

In this paper, the noisy voltage sag signal is processed with four noise reduction algorithms, including Heuristics Sure algorithm, Minimax algorithm, Penalize Medium algorithm and Subband adaptive algorithm, the threshold under the different noise level is calculated, and the results are shown in Table 1.

Table 1. The threshold table of four algorithms under different noise level

α	Thresholding Algorithm	Decomposition Level				
		1	2	3	4	5
0.100	Heuristics Sure	0.3612	0.3410	0.3194	0.0653	0.2692
	Minimax	0.2271	0.2271	0.2271	0.2271	0.2271
	Penalize Medium	0.2832	0.2832	0.2832	0.2832	0.2832
	Subband adaptive	0.4452	0.4392	0.4324	0.4246	0.4157
0.125	Heuristics Sure	0.4382	0.4133	0.3862	0.0633	0.3270
	Minimax	0.2754	0.2754	0.2754	0.2754	0.2754
	Penalize Medium	0.3602	0.3602	0.3602	0.3602	0.3602
	Subband adaptive	0.5523	0.5429	0.5324	0.5205	0.5069

4 Signal De-noising Results and Discussion

The voltage sag fault signal is generated using simulation model, which is performed 5-layer wavelet decomposition in the case of white noise in-aliasing factor of 0.125, choosing daub4 as wavelet function, after the decomposition, the low frequency coefficients and high frequency coefficients are shown in Fig. 4, such as d1, d2, d3, d4, d5 and a5, the voltage sag signal is reconstruct to get the voltage sag estimated signal by threshold processing.

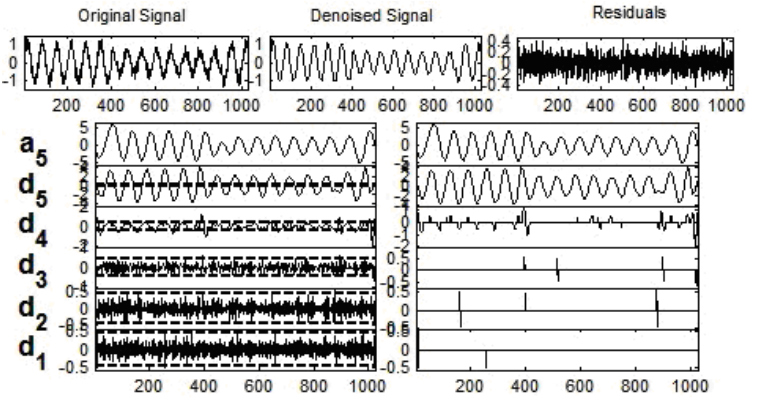


Fig. 4. Wavelet decomposition and de-noising effect diagram

From the noise reduction effect we can see, the noise cancellation signal is very close of the fault signal.

Power spectrum analysis is a powerful frequency-domain analysis tool. Via the power spectrum analysis of the residual signal, which is equal to the difference between original signal and de-noised signal, we will find the frequency components loss situation of the original signal and the noise signal distribution characteristic within a certain bandwidth, the residuals power spectrum using four de-noising algorithms are shown as Fig. 5, at the noise level 0.10:

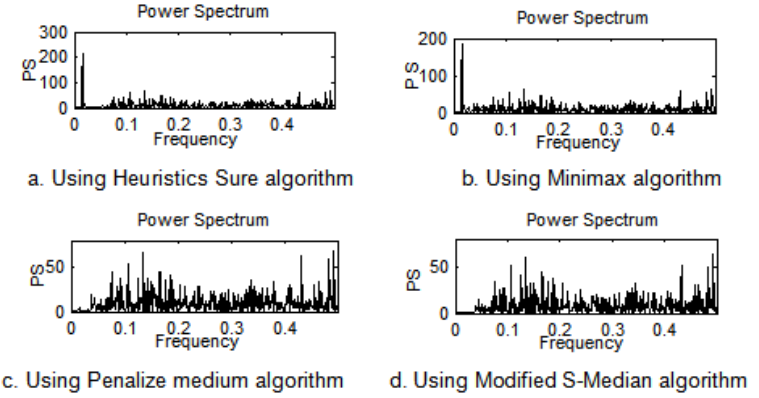


Fig. 5. Residuals power spectrum

We can see from Fig. 5(a, b), the low-frequency loss of signal is more serious using Heuristics Sure algorithm and Minimax algorithm than other two algorithms, and the low-frequency loss using Heuristics Sure algorithm is most serious, which low-frequency power spectrum amplitude exceeds the PS value 200. It illustrates that the low frequency components of signal are suppressed seriously using Heuristics Sure algorithm and Minimax algorithm, then too much useful information is removed, so

that the signal after noise reduction has a greater distortion. On the contrary, Fig. 5(c, d) show that there are very small low-frequency loss after the noise reduction using Penalize Medium algorithm and Subband adaptive algorithm, and the low-frequency loss is almost 0 using Subband adaptive Algorithm for de-noising while the corresponding high-frequency sequence is filtered, meanwhile, it is also found that the residual frequency components are evenly distributed in a large range, which is very close to the properties of the ideal Gaussian white noise distribution, so it can reflect the effect of the noise reduction method.

5 Conclusions

In this paper, a new level-dependent proposed noise reduction algorithm based on wavelet transform is proposed. It has improved the effect of power quality signal de-noising for PQM system. A voltage sag signal generated by simulation model is applied for simulation analysis, the simulation results show that the algorithm is more effective in de-noising voltage transient signal than other three de-noising algorithms.

Acknowledgements. The authors wish to acknowledge the support of Science Research Foundation (HGC1103) of Huaiyin Institute of Technology.

References

1. Wen, J., Liu, P.: A Method for Detection and Classification of Power Quality Disturbances. Autom. Elect. Power Syst. China 26(1), 42–44 (2002)
2. Huang, W.Q., Dai, Y.X.: Block-Thresholding Approach for Power Quality Disturbance De-noising. Transactions of China Electrotechnical Society 22(10), 161–165 (2007)
3. Ji, T.Y., Lu, Z., Wu, Q.H.: Detection of Power Disturbances Using Morphological Gradient Wavelet. Signal Processing 88, 255–267 (2008)
4. Mallat, S., Wen, L.H.: Singularity Detection and Processing With Wavelets. IEEE Transactions on Information Theory 38(2), 617–643 (1992)
5. Donoho, D.: De-noising by Soft-Thresholding. IEEE Transactions on Information Theory 41, 613–627 (1995)
6. Brito, N.S.D., Souza, B.A., Pires, E.A.C.: Daubechics Wavelets in Quality of Electrical Power. In: Proceedings of the 8th International Conference on Harmonics and Quality of Power, pp. 511–515. IEEE Press, Greece (1998)
7. Wang, Z.S., Jiang, H.K.: Robust Incipient Fault Identification of Aircraft Engine Rotor Based on Wavelet and Fraction. Aerospace Science and Technology 14, 221–224 (2010)
8. Poornachandra, S.: Wavelet-based De-Noising Using Sub-Band Dependent Threshold for ECG Signals. Digital Signal Processing 18, 49–55 (2008)

Excitonic Transitions Controlled by Voltage in a Single Self-assembled Quantum Dot

Peng Qiu, Guanglong Wang, Jianglei Lu, and Hongpei Wang

Institute of Nanotechnology and Microsystem, College of Mechanical Engineering, 050003 Shijiazhuang, China
qpfind@163.com

Abstract. Recently, tunnelling dynamics of electrons controlled by an electric field have been the focus of quantum dot electronic applications. This paper report electron and hole tunnelling phenomena in a single self-assembled quantum dot as a function of the applied electric field. The absorption spectroscopy is used to measure excitonic transitions under conditions where optical recombination can not be observed due to the high, ionizing, electric field. It shows that the tunnelling mechanisms are strongly dependent on the applied electric field. Hence the lifetime limited linewidth can be controlled directly by a gate voltage.

Keywords: Quantum dot, tunneling phenomena, excitonic transitions.

1 Introduction

Based on their promising applicability in quantum electrodynamics and quantum information, self-assembled semiconductor quantum dots (QDs) have been the focus of research for several years. Their atomic-like discrete energy states combined with their incorporation in a semiconductor heterostructure allows the study of the optical properties of single quantum dots, tuned by magnetic and electric fields, in a very controlled way. Recently, tunnelling dynamics of electrons controlled by an electric field have been demonstrated [1][2]. In this paper we report not only electron tunnelling to the continuum of the back contact but also resonant hole tunnelling phenomena to a two-dimensional hole system. We show that the tunnelling mechanisms are strongly dependent on the applied electric field. Hence the lifetime limited linewidth can be controlled directly by a gate voltage. In the experiment we present here, we investigated the excitonic transitions in self-assembled InAs/InGaAs quantum dots that are embedded in a field effect structure.

2 Excitonic transition in a Single Self-assembled Quantum Dot

In Fig. 1 the band structure of the sample is shown schematically. The quantum dots were grown with molecular beam epitaxy. They are separated by a 25 nm thick GaAs tunnel barrier from the back contact (highly n-doped GaAs). The QDs are overgrown

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 55–60.
springerlink.com

with 30 nm of GaAs. In order to prevent charge transfer from the semitransparent metallic top gate (5 nm of NiCr), a 120 nm thick blocking barrier consisting of a periodic AlAs/GaAs superlattice was deposited above the QDs. With this geometry, the electric field the QDs experience can be tuned in a controlled way, simply by applying a dc voltage between the back contact and the top gate. Furthermore the QDs can be charged with single electrons by shifting their electronic levels beneath the Fermi energy. The voltage V_g allows to control the electric field at the position of the QD. E_F is the Fermi energy, E_c and E_v are the conduction and valence ionization energies respectively. A labels the interface between the capping layer of GaAs above the dots and the AlAs/GaAs superlattice. E_0 is the energetic point of origin of the triangular potential well created at this interface.

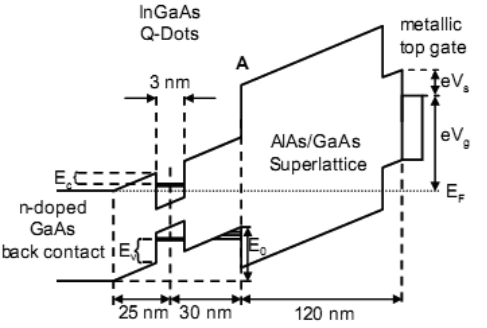


Fig. 1. Scheme of the energy band structure of the sample along the growth direction

In this letter we concentrate on the neutral exciton only. The measurements were carried out with Stark-shift modulation absorption spectroscopy[3]. This allows us not only to measure homogeneous linewidths of the excitonic transitions but also to measure at gate voltages where the QD is in the ionization regime. This is the regime where the tunnelling barrier to the back contact (which is voltage dependant) has become so small that the tunnelling rate of the electron in the QD is higher than the recombination rate. If the electron tunnels out of the QD before it can recombine with the hole under emission of photoluminescence (PL), the PL is quenched. However in absorption spectroscopy an exciton is created under absorption of a photon. This can still happen, even if the electron (and the hole) tunnel out of the QD before they can recombine optically. The shorter lifetime of the exciton in the QD will in this case lead to a broader linewidth than in the gate voltage region where PL is observable and the linewidth should represent the lifetime limited by spontaneous emission. The experiment was performed with a confocal microscope at cryogenic temperatures.

The energy dispersion of a neutral exciton (X^0) is shown in Fig. 2. Here the black dots represent the resonance energies of the neutral QD, measured in absorption. The fine structure of the neutral exciton is of the order of 20 μeV and cannot be resolved in this graph. The open dots represent the PL data of the very same dot. The voltage extent of the X^0 PL shifts to more negative voltages with increasing excitation power. In Fig. 2, we plot the PL positions extrapolated to very low powers. The origin of this

behavior is hole storage at the interface to the blockingbarrier (A in Fig. 1). The holes, a space charge, screen the applied electric field.

It can be clearly seen that, as expected, the excitonic transition can be observed to far more negative gate voltages in absorption, than in PL. Two absorption spectra are shown in Fig. 3. At gate voltages where PL can be observed we measure very narrow linewidths as small as 1.2 μeV (Fig. 3(a)). The reason why the measured linewidths are always larger than the expected values is due to fluctuations of the resonance energy of the excitonic transitions of the order of the linewidth. At large negative gate voltages, at which no PL can be observed, the linewidths are up to two orders of magnitude larger (Fig. 3(b)).

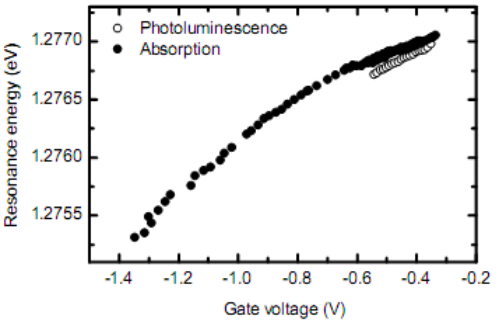


Fig. 2. Resonance energies of the neutral exciton (X^0) as a function of the applied gate voltage. The absorption data are represented by the black dots, the PL by the open circles. The small discrepancy in energy between the absorption and emission data is due to instrumental reasons.

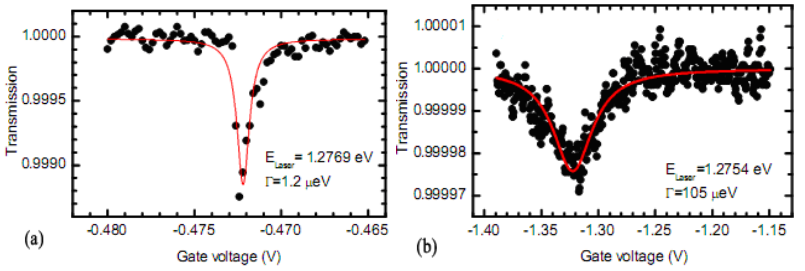


Fig. 3. Two absorption spectra measured at two different gate voltages

3 Electron Tunnelling

In Fig. 4(a) the measured linewidth is shown as a function of gate voltage. We observe an exponential increase of the linewidth with decreasing gate voltage. Further oscillations of the linewidth are noticeable this will be discussed later. In the regime (ionization regime) where no PL is observed ($V_g < -0.54$V) we conclude that the

lifetime of the excitonic state is limited by the tunnelling time of the electron. For gate voltages $V_g > -0.54$V, the lifetime is rather limited by the spontaneous emission of a photon (We neglect phonon assisted dephasing, as the experiment was performed under cryogenic temperatures T =4.2K, as well as power broadening due to the low excitation powers of $\approx 1\text{W/cm}^2$).

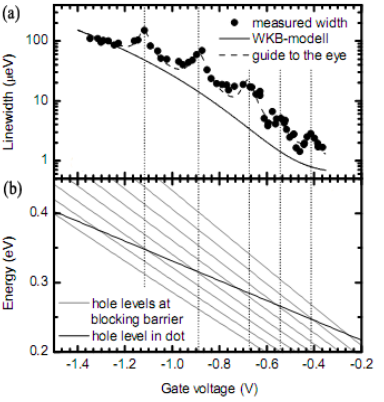


Fig. 4. Figure (a) shows the linewidth as a function of gate voltage. The solid line corresponds to the electron tunnelling that has been modelled with a WKB approximation. Figure (b) represents the hole energy levels E_h in the QD and E_n in the triangular potential well formed at the interface A (Fig. 1). The energies are defined relatively to the valence band edge at the back contact.

Considering a triangular tunnelling barrier between the quantum dot and the back contact as shown in Fig. 1 we can model the tunnelling time and hence the linewidth with a WKB approximation [3]:

$$\Gamma_{\text{WKB}} = \frac{\hbar^2 \pi}{2m^* L^2} \exp\left(\frac{4}{3\hbar} d(V_g)\sqrt{2m^* E_c}\right). \tag{1}$$

Here L is the height of the dot in z-direction (3 nm), E_c is the conduction band offset (Fig. 1) and $m^*=0.07m_e$ is the effective mass of an electron in GaAs. Also we use the gate voltage dependant tunnelling distance:

$$d(V_g) = \frac{E_c}{e(V_g - V_s)} d \tag{2}$$

with the built in voltage V_s =0.62V. Fitting Eq. (1) with an offset $\Gamma_0 = 0.62\mu$eV that corresponds to the lifetime limited linewidth we get a value for the conduction band offset of E_c = 134meV. This corresponds very well with the value received from analyzing the charging behavior of the dot.

4 Hole Tunnelling

The oscillations of the linewidth as a function of the gate voltage represent a resonant hole tunnelling effect. The AlAs/GaAs superlattice forms a triangular potential well at the interface to the GaAs capping layer (A in Fig. 1). This leads to quantized two-dimensional states in the valence band as schematically indicated in Fig. 1.

The steepness of the triangular potential well and therefore the energies of the quantized states are dependent on the gate voltage. In Fig. 4 (b) the energies of the quantized states in the triangular well and the hole state in the QD are plotted as a function of gate voltage. Comparing the gate voltages of the maxima of the linewidth and the crossing points of the two energy levels (the dashed lines are a guide to the eye) reveals a striking agreement. To calculate the energies in the triangular quantum well we used the result [4]:

$$E_n = E_0(V_g) - c_n \left(\frac{(eF(V_g)\hbar)^2}{2m^*} \right)^{\frac{1}{3}} \tag{3}$$

with m^*=0.5m_e the effective mass of the heavy hole in GaAs, $c_n \approx (3\pi(n-1/4)/2)^{2/3}$ and the electric field $F(V_g)$= (V_g+V_s)/175 nm. The parameter $E_0(V_g)$ = e(V_g+V_s)/η' (with the lever arm η'= 175/55) describes the point of origin of the potential well as shown in Fig. 1. The hole level in the dot is described by[3] [5]:

$$E_h = e\frac{V_g + V_s}{\eta} + E_v - E_b \,. \tag{4}$$

here the lever arm is η = 175/25 and the valence band offset E_v = E_c(1/Q-1)=97meV with E_c = 134meV from before and Q =0.58. The shift of the hole level in the QD by the binding energy of the exciton (E_b =6.2meV) can be calculated in an analog way and is determined to be E_b =2.7meV. The hole tunneling is fast whenever a 2D valence continuum state with in-plane wave vector less than $k_{\parallel}^{\max} \approx 1/r_{\parallel}$ is available where $r_{\parallel}$ is the in-plane extent of the quantum dot wave function. However, when the only available continuum state has $k > k_{\parallel}^{\max}$, tunneling is suppressed. In these experiments, $r_{\parallel} \approx$ 5 nm, implying $k_{\parallel}^{\max} \approx 2\times10^8$ mm^{-1}, yet the energetic separation between the $k_{\parallel}$= 0 2D hole states is about 20meV, such that a hole wave vector of up to 5 × 10^8 m^{-1} is required away from resonance with a $k_{\parallel}$= 0 state. This consideration explains the strongly oscillatory nature of the hole tunnelling in Fig. 4. We note that electron tunnelling is not suppressed by the same argument as there is no strong quantization in the z-direction in the back contact.

Another argument why hole tunnelling is expected in our sample structure is that in the ionization regime we still observe absorption [6] [7]. If only the electron tunneled and the hole remained in the dot, further absorption would be blocked. Our measurements show that this is not the case. The exponential increase in the linewidth is a compelling argument for both electron tunnelling and hole tunnelling.

5 Conclusions

This paper has shown an electric field dependency of the linewidth of transitions to the excitonic ground state of single self-assembled quantum dots. The linewidth can be tuned over two orders of magnitude, from linewidths limited by spontaneous emission to linewidths limited by the tunnelling times of the electron or the hole stored in the quantum dot. We modelled the electron tunnelling with a straightforward WKB approximation. Oscillations of the linewidth are deduced to be caused by resonant hole tunnelling to discrete states in a triangular potential well that is formed in our sample by an AlAs/GaAs superlattice.

References

1. Karrai, K.: Discovery of quantum dot-Landau level hybridized states. Nature 427, 135 (2004)
2. Smith, J.M.: Observation of spin relaxation through voltage-dependent cotuneling. Phys. Rev. Lett. 94, 197–402 (2005)
3. Kroner, M., Hogele, A., Seidl, S., et al.: Voltage-controlled linewidth of excitonic transitions in a single self-assembled quantum dot. Physica E 32, 61–64 (2006)
4. Srinivasan, K., Painter, O.: Momentum space design of high-Q photonic crystal optical cavities. Opt. Express 10, 670–684 (2002)
5. Högele, A., Seidl, S., Kroner, M., et al.: Voltage-Controlled Electron-Hole Interaction in a Single Quantum Dot. Journal of Superconductivity: Incorporating Novel Magnetism 18, 245–249 (2005)
6. Masia, F., Moreels, I., Hens, Z., Langbein, W., Borri, P.: Four-wave-mixing imaging and carrier dynamics of PbS colloidal quantum dots. Physical Review B 82, 115–302 (2010)
7. Yoshie, T., Scherer, A., Hendrickson, J., Khitrova, G., Gibbs, H.M., Rupper, G., Ell, C., Shchekin, O.B., Deppe, D.G.: Vacuum Rabi splitting with a single quantum dot in a photonic crystal nanocavity. Nature 432, 200–203 (2004)

Atmospheric Refraction Correction of Optical Measuring Single Station Using Reference GPS Data

Yuhua Dong and Jifeng Ding

College of Information & Communication Engineering, Dalian Nationalities University, Dalian, China
dongyuhua@dlnu.edu.cn

Abstract. In the flying experiments, the single optical theodolite can not form crossing location, so the atmospheric refraction can not be corrected, and the precision of equipments also can not be analyzed. Using reference GPS location data to position the height of the target, calculating the correction value of the target elevation angle, the equipment precision of the optical theodolite can be analyzed. The simulation results show that the method of refraction correction of optical measuring elevation angle by reference GPS data is feasible, and the correction value is credible. So the method is available, and it has value of popularization and application.

Keywords: the optical theodolite, single station, atmospheric refraction, elevation angle.

1 Introduction

The optical theodolite is the main equipment to complete the initial stage measuring task of flying target. Because there are only angle measurement elements, and no location measurement elements, so two or more crossing location of measurement station should be formed to calculate the track of the target, then the atmospheric refraction correction can be carried out. If only an optical theodolite detect the target during some time period, the measurement element data of the period can not correct the refraction. In the paper, the reference GPS positioning data is used as the target track of the period, then determine the target height, calculate the correction value of elevation angle refraction, then the atmospheric refraction can be obtained. The flying experiment data is used in the paper, and the refraction correction value and data error characteristics are analyzed, which verify the method is feasible.

When the relative height of optical measuring equipment is less than 200m or higher than 78000m, the approximate correction is adopted. The problem of atmospheric refraction correction with relative height from 200m to 78000m is mainly discussed in the paper.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 61–66.
springerlink.com

2 Atmospheric Refraction Correction of Optical Measuring Single Station Using Reference GPS Data

The known conditions are space geodetic rectangular coordinates of measuring station A is (X, Y, Z) and space geodetic rectangular coordinates of GPS positioning data of target M is (X_M, Y_M, Z_M), then the distance between measuring station and the target is

$$R = \sqrt{(X - X_M)^2 + (Y - Y_M)^2 + (Z - Z_M)^2} \tag{1}$$

The atmospheric refraction error of optical theodolite is determined by the light wave winding transmission path[1], as shown in Fig.1. If the measuring equipment has the function of ranging, the measuring distance is R_e in Fig.1, and in fact, the actual value of distance is R. Because reference GPS data is the true value of the target, so the calculated distance between measuring station and target is the true slant range R.

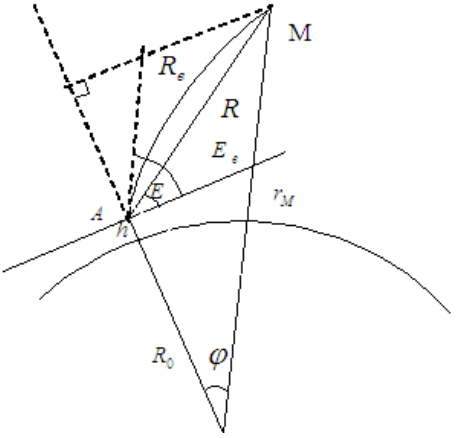


Fig. 1. The atmospheric refraction correction of light wave

The slant range R is used as the measuring distance data R_e of measuring distance, then the equipment can be used as a device with range data.

Firstly, calculate the initial value $\bar{r}_M$ of integral height

$$\bar{r}_M = \sqrt{R^2 + r_0^2 + 2Rr_0 \sin E_e} \tag{2}$$

Where $r_0 = R_0 + h$, R_0 is the mean radius of earth curvature, h is the geodetic height of measuring station, E_e is the observed data of elevation angle.

Secondly, calculate the earth center distance r_M according to the apparent range $\overline{R}_e$, and the formula of $\overline{R}_e$ is

$$\overline{R}_e = \int_{r_0}^{\bar{r}_M} \frac{rn^2(r)dr}{\sqrt{n^2(r)r^2 - n_0^{\,2} r_0^{\,2} \cos^2 E_e}} \tag{3}$$

Where $n(r)$ is the atmospheric refractive index varying with earth center distance r, which can be obtained by temperature pressure and humidity of r. n_0 is the atmospheric refractive index at the atmospheric stratification of measuring station.

Thirdly, the Eq.(3) uses numerical iterative of spherical integral to obtain R_e, compare R_e with the measuring value R, satisfying $|R_e - R| \le \delta$, (δ is a specified value), then $r_M = \bar{r}_M$.

Fourthly, after obtaining earth center distance r_M, the layered numerical integral of the geocentric angle φ is calculated by

$$\varphi = \int_{r_0}^{r_M} \frac{n_0 r_0 \cos E_e dr}{r\sqrt{n^2 r^2 - n_0^2 r_0^2 \cos^2 E_e}} \tag{4}$$

Where n is the atmospheric refractive index at the atmospheric stratification of the target, r_0 is the earth center distance of the measuring station, n_0 is the atmospheric refractive index of measuring station[2].

As shown in Fig.1, according to the earth center distance r_M of the target and the geocentric angle φ, the true elevation angle E of the measuring station is

$$E = arctan\frac{r_M \cos\varphi - r_0}{r_M \sin\varphi} \tag{5}$$

The refraction correction value of measuring station is

$$\Delta E = E_e - E \tag{6}$$

It should be noted that, only when the distance between measuring station and target is more nearer, the difference between the true slant range R and the apparent range $\overline{R}_e$ is small, then the above method can be established.

3 Examples and Analysis

In the process of simulations , the flying experimental data is produced, in which an optical theodolite equipment record the data continuously, and at first only the equipment detect the target, the data segment is called as a single segment. The following measuring data combined with other optical theodolite equipment can form

the common merger segment, and the data segment is called as merger. The two data segments satisfy the condition that the distance between equipment and the target is relatively near.

3.1 The Different Refractive Correction Methods Analysis of the Same Data

The merger can not only use the atmospheric refractive correction of single station, but also use that of multi-station merger, because there are other equipments tracking the target. The correction value using the single station method to correct the atmospheric refraction is shown in Fig.2. Then for the same data, using the multi-stations to calculate the correction value of the atmospheric refraction, the difference of two correction value is shown in Fig.3.

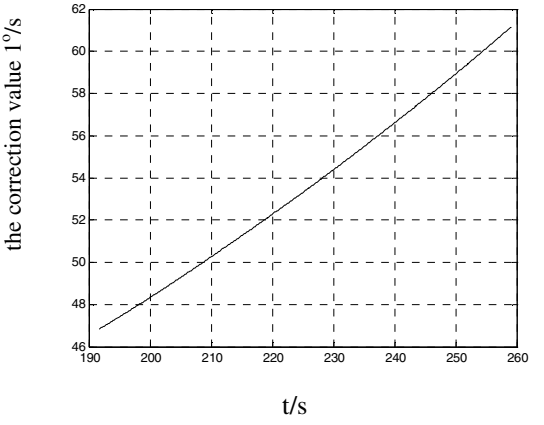


Fig. 2. The refractive correction value of merger using single station

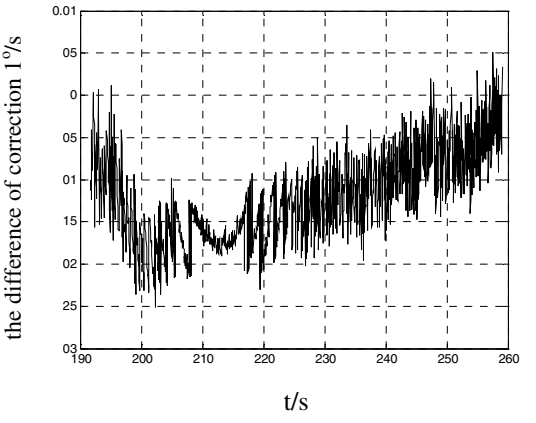


Fig. 3. Difference between correction values by two methods

From Fig.3, we can see that the correction value difference is within 0.03° /s, and different methods can not affect the correction effect.

3.2 The Influence of Data Error Characteristic by Different Refractive Correction Methods

The data segment is corrected using single station, and the data of merger segment is corrected by the means of multi-station merger. The difference between correction value and inverse calculating measuring element by reference GPS is calculated respectively, then draw plot and calculating statistical residual error. Because the time of single segment data and merger data is continuous, so the two residual error curves can also be connected together, as shown in Fig.4. The dotted line is single station correction, and the solid line is correction of multi-station merger. From the figure, it can be seen that the transition of the curves is smooth, and the change of refraction methods does not influence the equipment status.

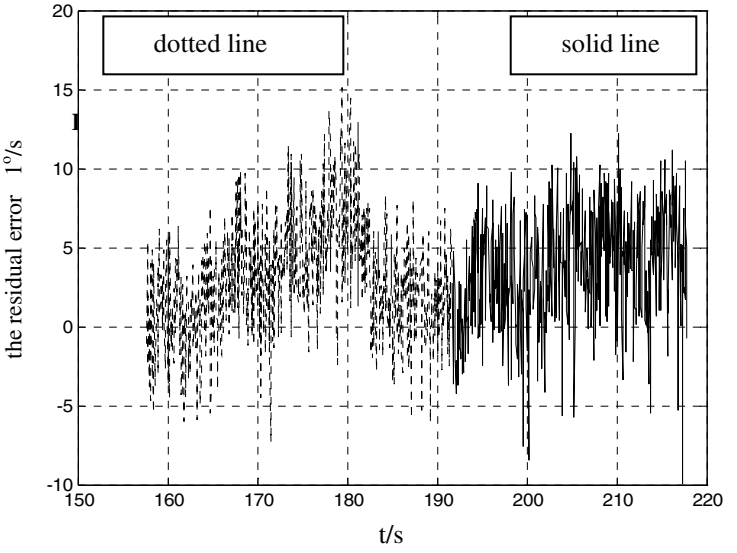


Fig. 4. The joint residual plot of two refraction method

Table 1 is the statistical results of residual error. From Table 1, it can be seen that the different refraction methods have little effect on the statistical results of equipment precision.

Table 1. The statistical residual of elevation angle by two refraction methods

	single segment (single station)	merger segment (single station)	merger segment (multi-station)
The mean of residual	3.05	3.71	3.76
The mean square derivation of residual	3.64	3.60	3.61

4 Conclusion

The optical measuring equipment is high precision equipment, and the obtained every bit of data is very valuable. In the flying experiments, the refraction correction is done by reference GPS data, while the measuring element data can not be merged by optical measuring equipment, the method can take advantages of limited measuring data to the maximum degree, and master the actual status of equipment more accurately, which can provide the basis of equipment alignment calibration. Although the method has its applicable constraints of short distance, but the method proposed in the paper can satisfy the precision requirement of flying experiment data processing, the method uses the reference GPS data to correct the atmospheric refractive correction of single station of optical measuring equipment. The task of optical measuring equipment is to track the flying target in initial stage, when the difference between the target flying tracking and theoretical trajectory is small, using theoretical trajectory to correct the atmospheric refraction of elevation angle can meet the requirements, then the optical measuring equipment can provide more additional data resources for the target location.

References

1. Wang, M., Hu, S., An, Z.: The measurement data error analysis technology of exterior ballistics and its application. National Defence Industry Press, Beijing (2008)
2. Liu, L.: Post processing of the exterior tracking & measurement data. National Defence Industry Press, Beijing (2000)

A Circuit Diagram Encryption Technology Based on Position Disordering and Gray-Scale Transformation

Li Tu[1] and Jianjun Wu[2]

[1] School of Computer Science and Engineering, Hunan City University, Yiyang, Hunan, 413000, China
[2] Center of Network, Hunan City University, Yiyang, Hunan, 413000, China
tulip1907@163.com, 40545092@QQ.com

Abstract. In this paper we have analyzed the various existing circuit diagram encryption techniques, and proposed a new algorithm, this algorithm combined with position disordering and gray-scale transformation, in the processing of position disordering, we have transformed the gray value of an circuit diagram reversibly, then we have used sensitivity analysis and relevance analysis to evaluate the new algorithm. The results showed that this method can get better circuit diagram scrambling effect, and increase the difficulty of decryption.

Keywords: Position disordering, XOR processing, Gray-scale value.

1 Introduction

There are n individuals in a circle, their number are from 1 to n, now they began to report their number from number s, let the people number to m out of the line, then they began to report their number from the next people, let the people number to m out of the line again……To do so until all the people out of the line. For any given parameters n, s and m, We can get the order of n people out of the line. This is the classic Josephus Problem. If we look on the order out of the line as a traversal sequence, it is called Josephus Traversing.

The algorithm is used in circuit diagram encryption, in order to obtain good results.

In the electronics industry, the circuit designing is a trade secret, intellectual property rights protected by law, so the circuit diagram of confidentiality is seriously.

2 Encryption Method

Suppose the picture need to be encrypted is P. When initializating an circuit diagram, we set up a simple matrix sequence, the size of the sequence is as big as the circuit diagram that need to be encrypted. In Joseph traversing algorithm, matrix C is changed into a one-dimensional matrix, and then we got a new one-dimensional matrix by changing of parameter s and parameter m, and through Matlab the new one-dimensional matrix was converted a two-dimensional matrix, it was matrix C1.Transfer each pixel of matrix P to the corresponding figures position of matrix C1.Circuit diagram P1 is an encrypted circuit diagram, when decrypting circuit

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 67–71.
springerlink.com

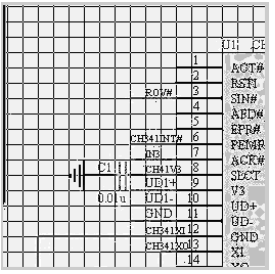
diagram p1,we need a decryptiing matrix C2,matrix C2 is obtained through C1 (Joseph encryption matrix). Finally we transferred each pixel of matrix P1 to the corresponding numbers position of matrix C2. This is the encryption of the pixels

The grayscale encryption is similar to the pixel encryption, the gray value of each pixel is carried out in the XOR processing, calculate each pixel and Joseph matrix with XOR processing, we get a new gray value, thus the circuit diagram is encrypted. In the decryption process, we can restore the original circuit diagram through XOR processing once again.

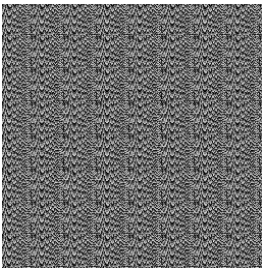
The encryption method in this paper is the combination of the two kinds of encryption.

3 Encrypted Results

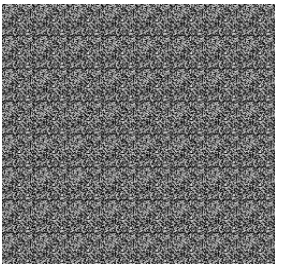
Changed the gray value of the circuit diagram, then disordered the position of the circuit diagram pixels, we obtained the following encryption circuit diagram, compared the gray value of the unencrypted circuit diagram and the encrypted circuit diagram, the result is shown in figure 1:



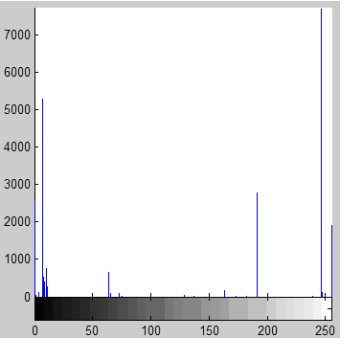
The original image



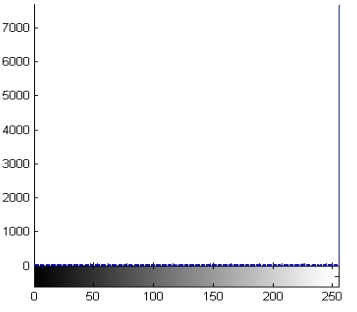
Grayscale encrypted image



Gray-scale encrypted in first, then position disordering encrypted image



The histogram of the image before encryption



The histogram of the image after encryption

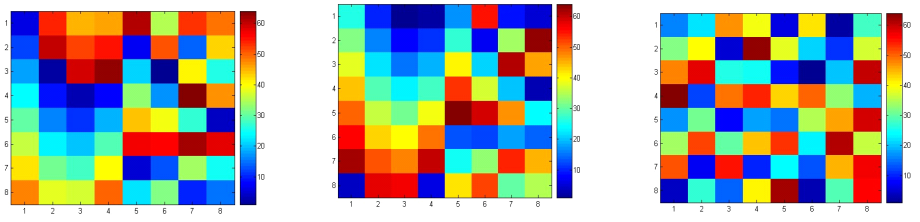
Fig. 1. Effect of encryption

4 Sensitivity Analysis and Relevance Analysis of the Algorithm

4.1 Sensitivity Analysis of the Algorithm

It can be seen from Figure 2, different values of parameter S and parameter m have a great impact on the matrix, and by changing the values of parameter S and parameter m constantly, we can get different encryption algorithm.

By adjusting values of parameter s and parameter m, the Joseph traversing sequence of the length of an n- sequence are n! kinds, so the disordering results of an M×N circuit diagram are (M!) N×(N!) M kinds.

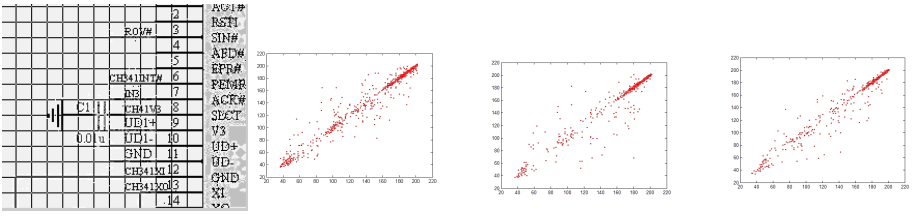


s=1,m=6 s=21,m=6 s=1,m=16

Fig. 2. Sensitivity analysis

4.2 Relevance Analysis of the Algorithm

The correlation between neighboring pixels of an circuit diagram can reflect the effect of disordering. It can be seen from Figure 3,in an original circuit diagram, the gray values of adjacent pixels are close, the maps of level correlation, diagonal or vertical degrees are basically in a straight line of equation y=x. In an encrypted circuit diagram, adjacent pixels in each circuit diagram are scattered, the correlation maps are more decentralized.



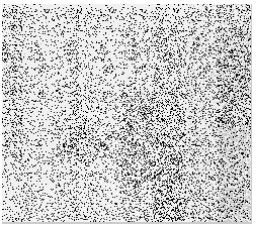
The original image (256*256)

level correlation of the pixels

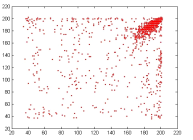
diagonal correlation of the pixels

vertical correlation of the pixels

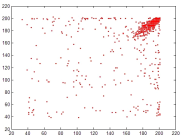
Fig. 3. Relevance analysis



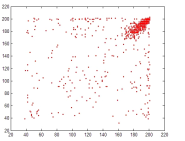
Position Disordering
（r=c=32）



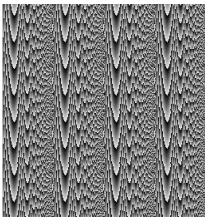
level correlation of the pixels



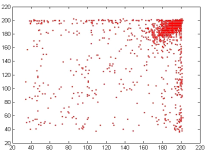
diagonal correlation of the pixels



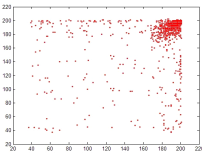
vertical correlation of the pixels



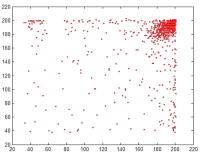
Position Disordering
（r=c=64）



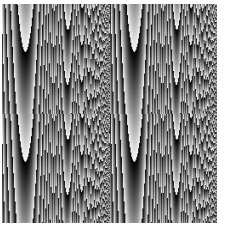
level correlation of the pixels



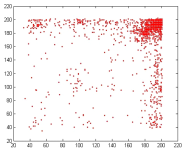
diagonal correlation of the pixels



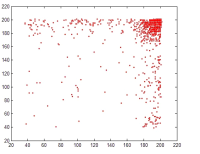
vertical correlation of the pixels



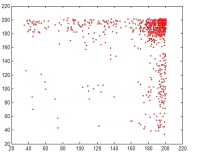
Position Disordering
（r=c=128）



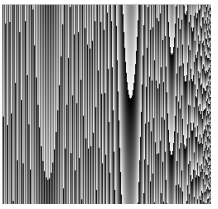
level correlation of the pixels



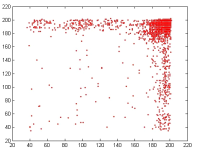
diagonal correlation of the pixels



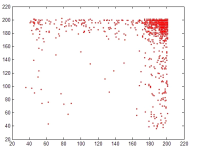
vertical correlation of the pixels



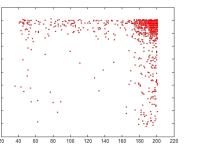
Position Disordering
（r=c=256）



level correlation of the pixels



diagonal correlation of the pixels



vertical correlation of the pixels

Fig. 3. (*continued*)

5 Conclusion

Position disordering changes the pixel position of an circuit diagram by mathematical transformation to achieve the purpose of scrambling, but there is a defect, it is, no matter how complex transform is, the histogram of the circuit diagram does not change. Gray-scale transformation only changes the histogram of the circuit diagram,

but dose not change the position of the circuit diagram pixels. This algorithm is the perfect combination of the two encryption methods.

The sensitivity analysis shows that, different encryption matrix can be gotten from changing the values of parameter r and parameter c, and the encrypted e matrix affects the encryption of circuit diagrams. It shows that the more obvious of disordering of circuit diagrams, the more obvious the effect of encrypting.

The algorithm not only can be applied to the circuit diagrams encrypted transmission, but also can be used for the encryption of all grayscale images.

Acknowledgment. This work was supported by Scientific Research Fund of Hunan Provincial Science Department under Grant (No: 2011FJ3022, 2011FJ3016).

References

1. Lu, W., Chen, T.: New conditions on global stability of Cohen-Grossberg neural networks. J. Neural Comput. 15(5), 1173–1189 (2003)
2. Valle, M.E.: A class of sparsely connected auto associative morphological memories for large color images. IEEE Trans. Neural Networks 20(6), 1045–1050 (2009)
3. Huang, W.Z., Huang, Y.: Chaos of a new class of Hopfield neural networks. Appl. Math. C. 206, 1–11 (2008)
4. Huang, Y., Yang, X.S.: Hyperchaos and bifurcation in a new class of four-dimensional Hopfield neural networks. Neurocomputing 69, 1787–1795 (2006)

Research on Routing Algorithm Based on Genetic Simulated Annealing Algorithm in Electronic Engineering

Yang Liu

School of Information Science and Engineering
Hunan City University
Yiyang 413000, China
lykk006@tom.com

Abstract. To improve the efficiency and effectiveness classical routing algorithm that still has some shortcomings. A genetic simulated annealing routing algorithm based on genetic algorithm, is studied for routing computing in electronic engineering. The simulation results show that the proposed algorithm can effectively speed up searching process, boost the chance of achieving the best path, and improve the quality of network service.

Keywords: wireless sensor network, routing algorithm, link state algorithm, distance vector algorithm, simulated annealing algorithm.

1 Introduction

With the rapid development of wireless communication technology, electronics technology, sensor technology and micro systems, wireless sensor networks more and more attention has been paid. The sensor network is a wireless network deployed in the observation environment within a large number of micro sensor nodes through wireless communication [1]. Composed of sensor network nodes, including sensors and aggregation nodes (Sink). Limited energy of sensor nodes is very difficult to replenish their energy again, and after deployment, sensor networks severe energy constraints.

The routing algorithm is efficient delivery of its functions, routing protocols must minimize the overhead of software and applications. The routing algorithm contains the global routing algorithms and decentralized routing algorithm. Distributed routing algorithm, each router only and it is directly connected to the router, each router in the network. These algorithms are also known as DV (Distance Vector, DV) algorithm [2]. Distance vector algorithms are based on the exchange of only a small amount of information. Distance vector algorithm, each router to send all or part of the routing table, but only distributed to its neighbors. Distance vector algorithms require each router sends its routing table all or part of the information, but only sent to neighboring nodes. The overall routing algorithm, each router has all the information of other routers in the network and network traffic status. These algorithms are also known as the LS (Link State, LS) algorithms. Link state algorithms, sometimes called

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 73–78.
springerlink.com

shortest path first or distributed database protocols, are built around a well-known algorithm from graph theory. The link-state algorithm to distribute routing information to each node of the network, but each router sends only describe part of its own link state routing table. Link-state algorithms to send routing information to all nodes on the Internet, however, each router sends only describes that part of its link state routing table. The link state algorithm everywhere send fewer update information, and the distance vector algorithm updates only to neighboring routers to send more information.Fig.1. is the Determining the route. Routing is the task of finding a path from a sender to a desired destination. For the above routing algorithm and its shortcomings, this paper presents a wireless sensor network routing algorithm based on genetic simulated annealing algorithm, called genetic simulated annealing routing algorithm (GSARA). The proposed algorithm combines genetic algorithm and simulated annealing to establish a link between the source and destination nodes and multiple iterations of the strategy to improve the genetic algorithm in wireless sensor networks for routing, effective in reducing frequency congestion, and reduce the routing overhead, balances node energy overhead, extending the network life cycle.

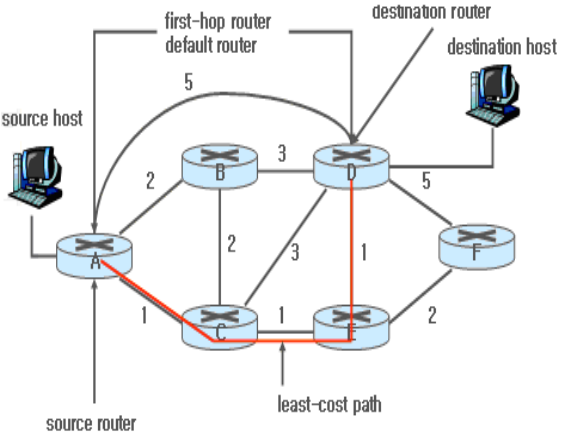


Fig. 1. Determining the route

This paper is organized as follows: a brief introduction to the routing algorithm design goals is given in section 2. Section 3 describes the routing algorithm metrics .Section 4 describes the proposed algorithms and presents the experimental results. In Section 5 summarizes the results and draws a general conclusion.

2 Routing Algorithm Design Goals

Routing algorithms typically have the following design objectives of one or more of: optimization, simple, low power, robust, stable and rapid convergence, flexibility [3,4].

(1) Optimization: refers to the ability to choose the best path routing algorithm. To calculate the metric values and weights.

(2) Simplicity: the algorithm is simple in design; routing protocols must be efficient delivery of its functions, to minimize the overhead of software and applications. The high efficiency is particularly important when the routing algorithm, the software must be running on a computer with limited physical resources.

(3) Robustness: routing algorithms in irregular or unpredictable environment, such as hardware failure, the load is too high or operational errors, can be run correctly. The routers in the network connection point will have serious consequences when they fail. The best router algorithm is usually able to withstand the test of time, and proved to be reliable in a variety of network environments.

(4) Fast convergence: the convergence of all routers in the judgment of the best path to achieve the same process. When a network event caused by routing available or unavailable, the router will issue updates. Routing updates throughout the network, causing re-calculate the best path, and eventually reach all routers unanimously recognized as the best path. The slow convergence of the routing algorithm will cause the path loops or network outages.

(5) Flexibility: The routing algorithm can quickly and accurately adapt to a variety of network environments. For example, a segment fails, the routing algorithm to be able to quickly find fault, and select the best path to use all routes of the segment.

3 Routing Algorithm Metrics

Use many different metrics to determine the best path routing algorithm. Complex routing algorithm may use a variety of metrics to choose the route, through weighted computation, and merge them as a single composite measure, and then fill in the routing table as a standard routing. Commonly used metrics: path length, reliability, delay, bandwidth, load, communication cost [5].

The path length is the most common routing metric. Some routing protocols allow network to each network link artificial Fu to the generation of value, this case, the route length through each link the sum of the consideration [6]. Other routing protocols define hop count; the packet must go through in the journey from the source to the destination network products, such as the number of routers. Reliability, more network links in the routing algorithm middle finger dependent (usually described in bit error rate), some network links than other failure, network failure, a number of network links than the other it easier or faster repair . Any reliability factors can give reliable rate assignment counted, usually conferred by the network to the network link metric value.

Routing delay refers to the packet from the source to take the time to reach the destination through the network. Many factors affect the delay, including the bandwidth of intermediate network links, each router port queue congestion, as well as physical distance of all intermediate network links. Because delay is a mixture of a number of important variables, it is relatively common and effective metric.

Bandwidth refers to the circulation of the links available capacity. In all other conditions are equal, 10Mbps Ethernet link 64kbps leased line is preferable. Although the bandwidth is the maximum throughput of the link available, but through links

with greater bandwidth do routing is not necessarily better than the route through slower links. For example, if a quick link busy, the packets arrive at the destination to spend the time may be longer. Load refers to the network resources, such as the peak of the router. Load can use a lot of calculations, including CPU usage and deal with the number of packets per second. Continue to monitor these parameters itself is very resource-intensive. The communication cost is another important metric; especially to some companies may be operating costs even in the performance. Even if the line delay may be longer, they also prefer to send data through their own line instead of the expensive public line.

4 Genetic Simulated Annealing Routing Algorithm

Simulated annealing algorithm derived from the principle of solid annealing, solid heated to sufficiently high, let it slowly cooling, heating, solid internal particles with the temperature rise becomes disordered like, can increase, while slowly cooling the particles gradually more orderly, and to reach equilibrium at each temperature, and finally reach the ground state at room temperature, can be reduced to minimum [7]. It is based on Mente-Carlo iterative solution strategy of a random optimization algorithm, the starting point is to optimize the similarity between the problem and the general combination of solid matter in the physical annealing process. Simulated annealing algorithm from the beginning of a higher temperature departure, accompanied by the continuous decline in the temperature parameters, combined with the probability characteristics of the sudden jump randomly in the solution space to find the global optimum of the objective function, namely the probability to jump out of local optimal solution and tends to the global optimum. The simulated annealing algorithm is a general purpose optimization algorithm, in theory; the algorithm has a probability of global optimization. The simulated annealing algorithm is changed by giving a kind of search process and ultimately tends to zero the probability of the sudden jump, which can effectively avoid falling into local minima and ultimately tends to the serial structure of the global optimum of the optimization algorithm.

The GA adopted the idea of biological evolution through natural selection and survival of the fittest competition strategy to solve the optimization problem. GA has strong global search performance, but in practical applications prone to premature convergence problem and the evolution of post-search less efficient. SA algorithm originated in statistical physics methods, and for the first time was Kirkpatric such as the introduction of optimization problems. The SA algorithm has good local search ability, but the strong dependence of the parameter. Genetic simulated annealing routing algorithm combining GA and SA algorithm consisting of an optimized algorithm. Poor local search capability of GA, but the ability to grasp the search process in general; SA algorithm to search the overall poor, but to take advantage of local search ability. GSARA algorithm to combine the advantages of the GA algorithm and the SA algorithm fully, greatly improving the operating efficiency of the algorithm and solution quality.

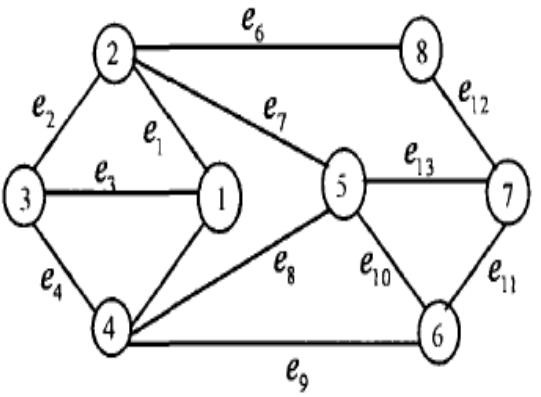


Fig. 2. Network topology diagram

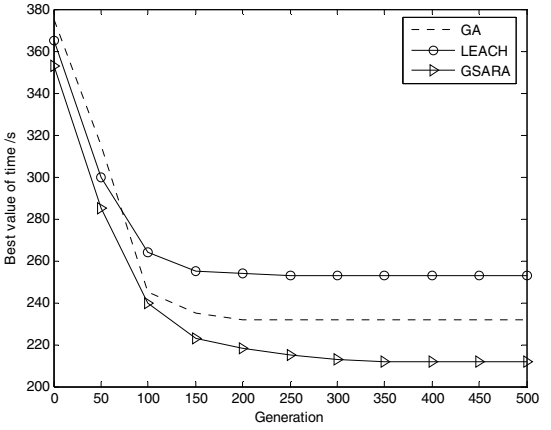


Fig. 3. Best value of time for three routing algorithms

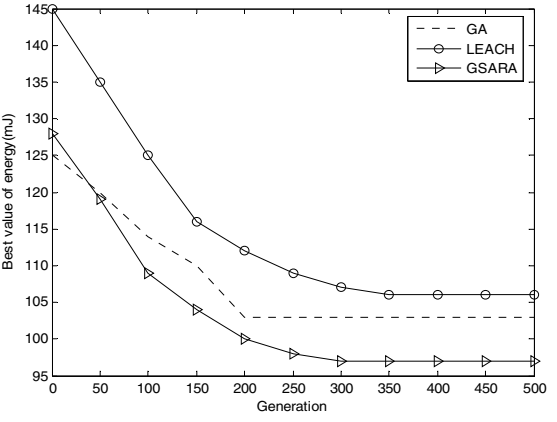


Fig. 4. Best value of energy for three routing algorithms

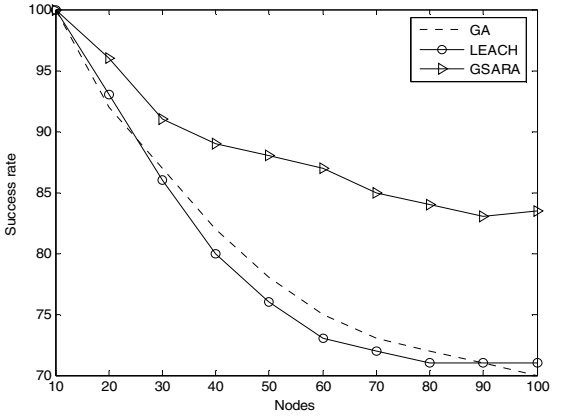


Fig. 5. Routing request success rate for three routing algorithms

5 Conclusions

In this paper, the improvement of existing electronic engineering, ant colony routing algorithm based on improved path searching existing routing algorithm, and weigh the mutual restraint between routing convergence speed of the network life cycle. Its application in electronic engineering, routing has a high value for improving the efficiency of the network of wireless sensor networks, to extend the network lifetime. Simulation results illustrate that genetic annealing algorithm not only overcomes the low convergence speed in simulated annealing algorithm but also solves the group degeneration problem in genetic algorithm.

Acknowledgments. The authors gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

References

1. Han, J., Park, K., Kim, W., et al.: A fast and highly adaptive Peer-to-Peer lookup system for medium-scale network. In: Proc of the Eighth IEEE International Symposium on Computers and Communication, ISCC 2003 (2003)
2. Liu, W., Rui, M., Xu, H.: Balanced wireless sensor network energy consumption of AODV improvement program. Computer Engineering 34(22), 143–147 (2008)
3. Liang, H., Chen, W., Shuai: A wireless sensor networks, ant colony optimization routing algorithm. Sensor Technology 20(11), 2450–2455 (2007)
4. Bing, L., Zheng, W.: Wireless sensor networks Multigroup ant colony optimization routing algorithm. Computer Applications 7(26), 2686–2690 (2009)
5. Gunes, M., Sorges, U., Bouaziz, I.: IARA-the-ant-colony based routing algorithm for MANETS. In: International Conference on Parallel Processing Workshops, ICPPW 2002, pp. 79–85 (2002)
6. Kassabalidis, I., El-Sharkawi, M.A., Marks, R.J.: Swarm intelligence for routing in communication networks. Global Telecommunications 6(6), 3613–3617 (2001)
7. Stutzle, T., Hoosh, H.: Max-Min ant systems. Future Generation Computer Systems 16(19), 889–914 (2000)

A Limited Feedback Scheme for User Selection in Electronic Communication Systems Based on Threshold

Bo Yu[1], FanZhi Kong[2], Hua Zhang[1], and XinHui Quan[1]

[1] Northeastern Petroleum University, Daqing, China
yblj9000@163.com
[2] Tianjin University of Science and Technology, Tianjin, China
kongfanzhi@tust.edu.cn

Abstract. In electronic communication engineering, feedback is an important way to improve the system performance. In multi-user MIMO electronic systems, when the total amount of feedback is limited on uplink, a small number of users with high-rate feedback have prior sum-rate to a large number of users with low-rate feedback. But a small amount of users' feedback is not conducive to multi-user diversity. To this point, an algorithm for user selection based on threshold is proposed and to determine the users for feedback among the users of elected according to optimal feedback rate, on the basis of user feedback channel direction information and channel quality information, the zero-forcing beam-forming algorithm is applied to precoding. Simulation results show that the proposed method has a higher sum-rate than non-threshold selection algorithm with optimal feedback rate.

Keywords: electronic engineering, multiple input multiple output (MIMO), user selection, precoding.

1 Introduction

In MIMO electronic communication system, Dirty Paper Coding (DPC) which is hard to realize in actual device can achieve broadcast channel capacity[1-2]. To avoid non-linear processing of DPC, linear processing methods proposed by researchers are liable to be implemented, such as Zero-Forcing Beam-Forming (ZFBF) in transmitter or channel inverse and so on[3-4].

It is necessary for transmitter to be aware of the information of some known users in order to get multi-user diversity and effective precoding in multi-user electronic system, which, however, is got by users' limited feedback in FDD electronic system. The total feedback limited uplink multi-user MIMO system proposed in [5] analyzes the optimal user feedback rate, but less user feedback is not conducive to multi-user diversity. [6] proposed a scheme for user selection and precoding based on threshold. Although utilize threshold has reduced the number of candidate users, the number of candidate users we got is much larger than the number of users in reference [5] where optimal feedback rate was required. In other words, due to the feedback rate of per user is low, its performance still has room for improvement.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 79–83.
springerlink.com

In multi-user MIMO electronic system with total feedback rate is fixed, a novel scheme for user selection and precoding based on threshold optimization is proposed to the issue of mutual restraint between user feedback rate and multi-user diversity. Simulation reveals the scheme proposed has higher sum-rate than that in reference [6].

2 System Model

Consider a MIMO broadcast channel, assume base station has M_t antennas and system owns K users equipped with single antenna. Channel output y_k is:

$$y_k = \boldsymbol{h}_k \boldsymbol{x} + n_k, \quad k = 1, \cdots, K \quad , \tag{1}$$

where n_k is additive white Gaussian noise with zero mean and variance one, $\boldsymbol{h}_k$ is the vector of channel coefficients from the *K*th user antenna to the transmitter antenna array and $\boldsymbol{x}$ is the vector of channel input symbols transmitted by the base station .The channel input is subject to the average power constraint, and fits $\mathrm{E}[\| \boldsymbol{x} \|^2] \le P$. Channel accords with block-fading model, is an ergodic stationary spatially white jointly Gaussian process.

Each user is assumed to know its own channel. Each user quantizes its channel to α bits and feeds back the bits perfectly and instantaneously to the BS. The vector quantization is performed using a codebook that consists of 2^{α} M_t-dimensional unit norm vectors $C_k = \{\boldsymbol{c}_{k,1}, \cdots, \boldsymbol{c}_{k,N}\}$, $N = 2^{\alpha}$. Each user quantizes its channel vector to the quantization vector that forms the minimum angle to it. Thus, the *K*th user quantizes its channel to $\boldsymbol{q}_k$, chosen according to:

$$\boldsymbol{q}_k = \arg\max_{\boldsymbol{c}_{k,j}^{\dagger}} | \tilde{\boldsymbol{h}}_k \boldsymbol{c}_{k,j}^{*} | \quad , \tag{2}$$

where $\tilde{\boldsymbol{h}}_k = \frac{\boldsymbol{h}_k}{\|\boldsymbol{h}_k\|}$ is the *K*th user's channel direction information (CDI).

3 Optimized Threshold Based on Probability Distribution

To further decrease the number of users above the threshold in [6] and satisfy the necessary users for optimal feedback rate[5]. The following is threshold optimization based on probability distribution of the channel gain.

Let us define $\boldsymbol{h}_k$ as the *K*th user's $1 \times M_t$ channel vector, whose elements meet the complex Gaussian distribution with zero mean and variance one. Channel gain $X = \|\boldsymbol{h}_k\|^2$ meets the Gamma distribution with parameter is M_t and 1, that is $X \sim \Gamma(M_t, 1)$.

The cumulative distribution function of X is as follows:

$$F(X) = \frac{\gamma(M_t, X)}{\Gamma(M_t)} \quad , \tag{3}$$

where $\gamma(M_t, X) = \int_0^X t^{M_t-1} e^{-t} dt$, $\Gamma(M_t)$ is Gamma function, When M_t is positive integer, there exists $\Gamma(M_t) = (M_t - 1)!$. When M_t is integer, (3) can be expressed as:

$$F(X) = 1 - \sum_{i=0}^{M_t-1} \frac{X^i}{i!} e^{-X} \quad . \tag{4}$$

If there are K users, β is threshold of users' channel gain, there are at least N users' channel gains meet $X = \|\boldsymbol{h}_k\|^2 \geq \beta$, its cumulative distribution is expressed as:

$$F_{K,N}(\beta) = \sum_{m=0}^{K-N} C_K^m F^m(\beta)(1 - F(\beta))^{K-m} \quad . \tag{5}$$

When the total uplink feedback bits Z are constant, optimal feedback bit α_{opt} and users N which needs to feedback CDI can be got from [5], then BS will select threshold β_{opt} according to (5), satisfy $F_{K,N}(\beta_{\text{opt}}) \approx 1$, according to users K.

4 User Selection and Precoding Strategies

The specific procedures for user selection and precoding based on threshold optimization are as follows:

(1) First determine the total feedback rate Z, and then estimate optimal feedback rate α_{opt} and number of optimal users N by (3). According to number of system users obtain threshold β_{opt} from (5).

(2) When user channel gain is $\|\boldsymbol{h}_k\|^2 \geq \beta_{\text{opt}}$, one bit feedback conducted, N users with channel information are randomly selected from candidate users by base station.

(3) According to the user feedback, base station using semi-orthogonal user selection algorithm to complete the user selection. Algorithm procedures as follows:

A. Define user set is $S_0 = \{1, \cdots, N\}$, the user feedback unit norm CDI quantitative information is expressed as $\boldsymbol{q}_s$, $s \in S_0$;

B. According to the channel gain of user feedback, choose users with the largest gain, the expression of user's serial number s_1 is: $s_1 = \arg\max_{s \in S_0} \|\boldsymbol{h}_s\|^2$.

C. Define $S_1 = S_0 - \{s_1\}$, $\kappa_j(1) = z(\boldsymbol{q}_{s_1}, \boldsymbol{q}_j)$, $\forall j \in S_1$, where, $z(\boldsymbol{q}_{s_1}, \boldsymbol{q}_j) = \|\boldsymbol{q}_{s1} \boldsymbol{q}_j^{\text{H}}\|^2$.

D. For $2 \leq m \leq M_t$, repeat the following steps: a. $s_m = \arg\min_{j \in S_{m-1}} \kappa_j(m-1)$; b. $S_m = S_{m-1} - \{s_m\}$; c. $\kappa_j(m) = z(\boldsymbol{q}_{s_m}, \hat{\boldsymbol{h}}_j) + \kappa_j(m-1)$, $\forall j \in S_m$.

In the above algorithm, $\kappa_j(m-1) = \sum_{i=1}^{m-1} z(\boldsymbol{q}_{s_i}, \boldsymbol{q}_j)$ is used to measure the orthogonality between the alternative users CDI quantization vector $\boldsymbol{q}_j$ and the selected users CDI vector $\{\boldsymbol{q}_{s_i}\}_{i=1}^{m-1}$.

(4) Conduct Beam-Forming [7] by utilizing ZFBF.

5 Simulation Results

Reference [6] analyzed the threshold value to the users selecting joint ZFBF precoding scheme based on threshold feedback, and the threshold t ranges between $\ln K - \ln\ln K$ and $\ln K$, this paper defines $t = \ln K - \ln\ln K$ as the threshold 1 lower bound, while $t = \ln K$ as threshold 1 upper bound.

The performance of optimal feedback algorithm and its rate are simulated and analyzed in figure1 and figure2 towards non-threshold selected randomly, threshold 1 lower bound, threshold 1 upper bound and the threshold in this paper and so on. The simulations are realized in the condition of the base station equipped with 4 antennas and K users each equipped with one antenna, the channel meets the static Rayleigh fading, SNR=10dB(P=10dB).

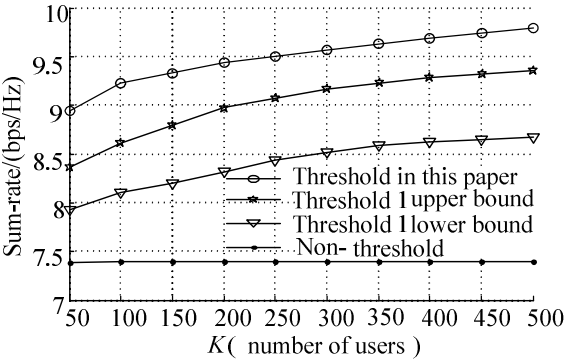


Fig. 1. Sum-rate versus the number of users for Z=50 bits

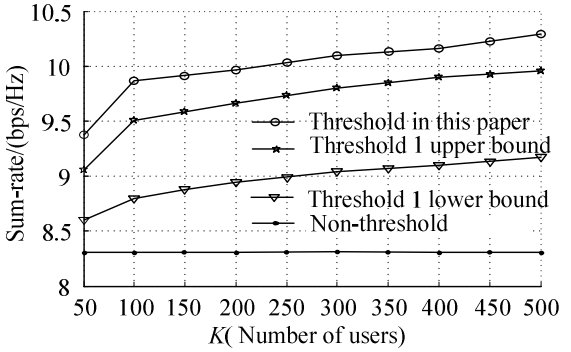


Fig. 2. Sum-rate versus the number of users for Z=100 bits

In figure 1, simulate the performance of several programs when the total amount of uplink feedback $Z = 50$bits . When the number of system users is $K = 300$, the scheme in this paper is superior respectively 0.3, 1.0 and 2.3bps/Hz over threshold value 1 upper bound, threshold lower bound and non-threshold scheme in reference [5].

The performance of several schemes can be seen in figure2 when the system total amount of uplink feedback $Z = 100\text{bits}$. Obviously, similar conclusions can be drawn as figure 1. By this time, when the number of system users $K = 300$, the scheme in this paper is superior respectively 0.3, 1.1 and 1.8 bps/Hz over threshold value 1 upper bound , threshold lower bound and non-threshold scheme.

6 Conclusion

Multi-users diversity and single user feedback rate restrict mutually in electronic engineering with uplink total feedback rate restricted. A solution based on threshold for this problem is proposed. Analyze the threshold of the channel gain for multi-user diversity, and propose an optimal feedback rate scheme on the basis of threshold optimization. With the single user feedback rate unchanged, utilize threshold to control channel gain, improve the multi-user diversity gain, thus improve the electronic system sum-rate, the simulation results demonstrate the excellent performance of the scheme.

Acknowledgement. The work is supported by Key Sciences and Technologies Program of Heilongjiang Educational Committee (No. 11551025).

References

1. Weingarten, H., Steinberg, Y., Shamai, S.: The capacity region of the Gaussian multiple-input multiple-output broadcast channel. IEEE Trans. Inform. Theory 52, 3936–3964 (2006)
2. Gesbert, D., Kountouris, M., Heath, R.W.: Shifting the MIMO paradigm: From single user to multiuser communications. IEEE Sig. Proc. Mag. 24, 36–46 (2007)
3. Peel, C.B., Hochwald, B.M., Swindlehurst, A.L.: A vector- perturbation technique for near capacity multi-antenna multiuser communication-part I: Channel inversion and regularization. IEEE Trans. Comm. 53, 195–202 (2005)
4. Yoo, T., Gold Smith, A.: On the optimality of multi-antenna broadcast scheduling using zero-forcing beamforming. IEEE Jour. Select. Areas in Comm. 24, 528–541 (2006)
5. Ravindran, N., Jindal, N.: Multi-user diversity vs. accurate channel feedback for MIMO broadcast channels. In: IEEE, ICC 2008, pp. 3684–3688 (2008)
6. Bayesten, A., Khandani, A.K.: On the user selection for MIMO broadcast channels. IEEE Trans. Inf. Theory 54, 1086–1107 (2008)
7. Yu., B., Yang, X., Li, J.: Joint Multiuser Random Beamforming and Zero-forcing Precoding for SDMA Systems with Limited Feedback. In: IEEE, WiCOM 2011, vol. 1, pp. 1–4 (September 2011)

The Simulation and Algorithm of Chaotic Polynomial Family for Secure Electronic Communication

Chunling Sun

Heilongjiang University Harbin, 150080, China
chunling6666@163.com

Abstract. The principle of constructing chaotic polynomial and the constructing process of the chaotic polynomial families were discussed. Several important simulation results were listed. The new chaotic time sequence created by iteration of the chaotic polynomial family is differed from the traditional Logistic chaotic time sequence. A simple algorithm for electronic secure communications ions were proposed.

Keywords: Chaotic polynomial family, Simulation, Simple algorithm, electronic secure communications.

1 Introduction

The most relevant literature in recent years that we have found has focused on its application research. Currently several famous chaotic systems that are researched widely (such as Logistic map Lorenz system, etc.) can produce a clear and simple chaotic time sequence, which is currently the main aspect of application, but they are several only the several "seeds" for chaotic systems to produce chaotic time sequence, which cannot satisfy various demands in real application, especially for electronic communication.

Nowadays, how to open up the source of chaotic sequence has become one of the most important problems in the application research of chaotic time sequence. We analyze the mechanism of the chaos and construct chaotic polynomial family. The iteration of function will produce a new chaotic time sequence differing from traditional Logistic chaotic sequence.

2 The Constructing Process of the Chaotic Polynomial Family

Now let's study the restriction of the cubic polynomial self-map. Firstly we will deduce the formula of the maximum of $f(x)$ and make it satisfy the demands of the self-map in the domain (0, 1).

$f'(x)$ is defined as the derived function of $f(x)$,

$$f'(x) = a(3x^2 - 2(1 + x_1)x + x_1) = 0 \tag{1}$$

We arrange it and have

$$3x^2 - (2x_1 + 2)x + x_1 = 0\,. \tag{2}$$

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 85–90.
springerlink.com

Then the extracting root formula is

$$x = (-b \pm \sqrt{b^2 - 4ac})/(2a) \tag{3}$$

Thus we can get two roots which satisfy $x^1 < 1 < x^2$ by calculation. Due to the foregoing supposition: $x_1 > 1$, we can abandon one root and get the max point at the interval (0, 1):

$$\mathrm{x}^1 = (x_1 + 1 - \sqrt{x_1^2 - x_1 + 1})/3 \tag{4}$$

$f(\mathrm{x}^1)$ is the maximum value at the interval.

$$f(\mathrm{x}^1) = a\mathrm{x}^1(\mathrm{x}^1 - 1)(\mathrm{x}^1 - x_1) \ (x_1 > 0) \tag{5}$$

And $f(\mathrm{x}^1)$ satisfies the following request:

$$\begin{cases} a\mathrm{x}^1(\mathrm{x}^1 - 1)(\mathrm{x}^1 - x_1) > 0 \\ a\mathrm{x}^1(\mathrm{x}^1 - 1)(\mathrm{x}^1 - x_1) < 1 \end{cases} \tag{6}$$

$f(x)$ is the cubic self-map. The final restriction of the constructed cubic self-map function is arranged as:

$$0 < a < (\mathrm{x}^1(\mathrm{x}^1 - 1)(\mathrm{x}^1 - x_1))^{-1} \tag{7}$$

So when we confirm the value of x_1 (the third zero point), we can apply the formula (4) to confirm the maximal value of $f(x)$. Put it into the restriction (7), we can confirm the span of the controllable parameter, and get one cubic polynomial self-map function. When different x_1 into the cubic polynomial self-map were put, we can get one domain of a, and find a right a which leads the cubic function into chaos, accordingly renew a set of chaos.

3 The Simulation Analysis on the Chaotic Polynomial Family

Fig. 1 shows the max of changing cubic polynomial function $f(x)$ when the fixed $x_1 = 4.0$, we protract the image of function. In the figure the numbers marked by each corresponding image of function are the values of a of each cubic polynomial function. In the definition domain (0, 1), the maximum of the self-map function is the peak of the image, we can get the maximum by substituting the foregoing formula (4) and see the area of the max is (0, 1) from the image. The perturbation area of a is (0, 1.1523), the upper bound of the perturbation area can be obtained by putting the foregoing fixed value into the formula (7).

We will investigate the cubic chaotic polynomial functions for different x_1 (the third zero point). The function curve of a set of different cubic polynomial function is shown as Fig. 3. The curve has the same maximum of the interval. In the figure, the numbers marked by each corresponding the image of function are one zero point of the cubic function outside interval. We choose $x_1 = 1.0, 1.5, 2.0, 4.0 and 1024.0$ as the five representative curves.

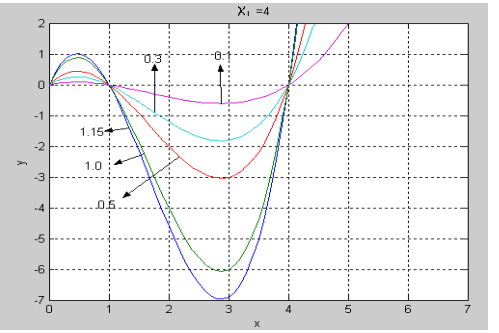


Fig. 1. The simulation of f as $x_1 = 4.0$ and different a

When 1.0 is chosen as the value of a, the cubic polynomial function will have the multiple roots 1.0. Here the cubic polynomial function is the initial point of the constructing functions and has significantly representative meanings. We can know that the span of a is (0, 6.75) by calculating. The value of x_1 can be chosen 2.0 which are ones of the mid-course amounts of the cubic polynomial function family. Because the numerical value of the computer simulation can never come to the infinite point, here $x_1 = 1024.0$ is used to substitute for x_1 at the infinite of the real axis.

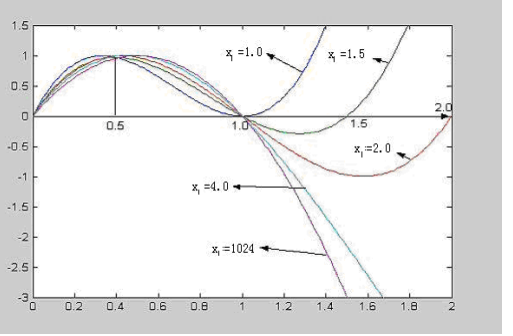


Fig. 2. The simulation of f below different x_1

In the definition domain (0, 1), the maximum of the polynomial self-map function is the peak of the image. Here we keep on making the maximum close to 1. By simulation research we know that the bigger x_1 is, the closer the maximum to 0.5, and the closer to the image of the quadratic parabola the cubic polynomial image at the segment (0,1).

If we fix x_1, different a has the periodic frequency of the corresponding time sequence in the perturbation area of a. When $x_1 = 2.0$, we put the value of x_1 into formula (7), the upper bound of the perturbation area can be obtained within the span of (0, 2.5901) by calculating.

Different a in the area (0,1) can reflect the change of the function. When $a = 1.7999, x_1 = 2.0$, the time sequence of the corresponding cubic polynomial self-map function can be produced by iteration is 2 periodic. That is to say, the dynamic behaviors of time sequence can be predicted. As an augments, the state of time sequence directly enters chaotic state from 2 periodic to 2n periodic.

Table 1. The table of the doubling bifurcation as $x_1 = 1.0, 2.0.4.0$

Periodicity	x1=1.0 Operator-value of a. at bifurcated point,	x1=2.0 Operator-value of a.	x1=4.0 Operator-value of a.
1->2	4.000005	1.799999	0.810010
2->4	4.970005	2.190001	0.97255
4->8	5.224000	2.270010	1.003549
Entering chaos	5.30045	2.288901	1.01256

When a operates the value near to 2.5901, the time sequence which is produced by the iterative system is in the chaotic state. In Table 1, when noting $x_1 = 1.0, 2.0 and 4.0$, correspondingly different a, the time sequence enters chaotic several typical value of a by the doubling bifurcation.

Fig. 3 shows the cubic self-map's bifurcation as $x_1 = 2.0$ whether the series of value of a satisfy the Feigenbaum constant or not needs further investigation.

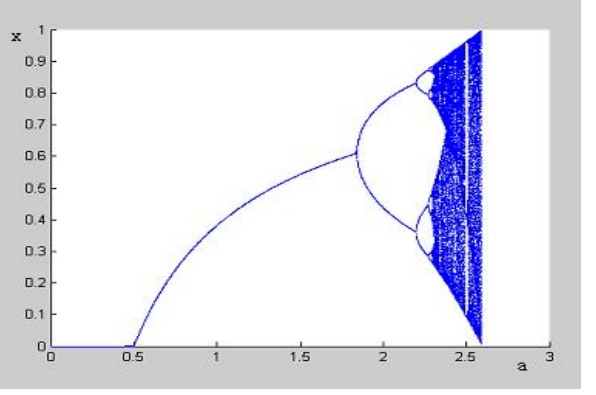


Fig. 3. The simulation of the cubic self-map's bifurcation as $x_1 = 2.0$

4 The Proposed Simple Algorithm

A class of Sequences (8) is generated by iterating the cubic polynomial family $\{f^m(x)\}(m = 0,1,2,...)$.

4.1 The Details of Scheme Are Described as Follows

$$\{x_n^m\}, (n = 0,1,2..., m = 0,1,2,...2^t) \tag{8}$$

here $2^t = 8$.

Employs the chaotic cubic polynomial family map

$$x_{n+1}^m = \alpha^m x_n^m (x_n^m - 1)(x_n^m - \overline{x_1}^m), 0 < x_0^m < 1 \tag{9}$$

4.2 Convert the Whole Phase [0,1] into [0,2l -1], Here l=8

$$\begin{aligned} c_n^m &= \text{int}(x_n^m * 255) + 1, \\ &(n = 0,1,2..., m = 0,1,2...8)) \end{aligned} \tag{10}$$

4.3 For the Convenience of Description, We Reset the Time Sequence $\{c_n^m\}$

$$\{c\}_j = \{c_j^1 c_j^2 c_j^3 c_j^4 c_j^5 c_j^6 c_j^7 c_j^8\}_j, (j = 0,1,2...) \tag{11}$$

4.4 Divides the Plaintext Message S into Blocks Sj of Length k Bytes, Here k=8

$$s_j = \{s_1 s_2 s_3 s_4 s_5 s_6 s_7 s_8\}_j (j = 0,1,2,....) \tag{12}$$

4.5 If All the Plaintexts Have Already Been Encrypted, the Encryption Process Is Finished. Otherwise, Calculates

$$d_j = c_j \oplus s_j = \{c_j^i \oplus s_i, i = 1,2...,8\} (j = 0,1,2,....) \tag{13}$$

Each block of plaintext can be encrypted with the time sequence created by the different cubic chaotic polynomials. Meanwhile, in order to improve the security of the cryptosystem, some special operation can be employed such as enlarging m (increasing the number of the cubic polynomials) and adjusting the status value of the cubic map in the iterating process.

The decryption process is almost the same as the encryption one. We only need to repeat step 4.1,4.2 and 4.3 and do the inverse operation in substitution.

5 Conclusion

The general quadratic polynomial map really expends a function family which the function family stretches towards two directions at both horizontal and vertical axes. Because the function families are defined in the real number space (in practical calculation in the rational number of the finite precision), the constructed function families are (close to) density.

By the simulation for the construction of the cubic polynomial self-map family, it approves the validity and correctness to the constructing process from the angle of

numerical value experiment. It provides support for further analyzing on the dynamic character of time sequence. The constructing process also needs to be further perfected.

The cubic chaotic polynomial families bring huge advance for the application demand. Theoretical and simulation results indicate that the cipher distribution of this scheme is even. In practice, of course, the proposed algorithm can be further improved.

References

1. Gao, J.-S., Sun, B.-Y., Han, W.: The function structure of the control orbit based on chaotic theory. Journal Electric Machines and Control 2, 150 (2002)
2. Qiu, Y., He, C., Zhu, H.: A kind of feedback-chaotic stream cipher of feedforward structure. Journal Transaction of Xi'an Jiaotong University 36(3), 309–312 (2002)
3. Wu, X., Chen, Z., et al.: Introduction of chaos. Shanghai science and technology literature book concern, Shanghai (1996)
4. Gao, J., Sun, B., et al.: The study of encryption process which is based on chaotic theory. Journal Techniques of Automation and Applications 6, 13 (2001)
5. Grank, J.: Chaos — inaugurate new science (Zhang, S. (trans.)). Shanghai translation book concern (January 1991)
6. Lorenz, E.N.: The essence of chaos. Liu Shida interprets. Weather book concern, Beijing (1997)
7. Hao, B.: Discussion from parabola — introduction to chaotic dynamics. Shanghai science and technology education book concern, Shanghai (1993)

Chaotic Communication System Prediction Based on Dimension-Exponent Mapping

ZhenChao Wang, Pei Du, and Chen Zhao

College of Electronic and Information Engineering, HeBei University, BaoDing, China
wangzhenchaohbdx@163.com, ziyoumudi@yahoo.com.cn,
416441383@qq.com

Abstract. Basing on analysis of the limitations of the traditional high embedding dimension chaotic time series prediction algorithm, this paper introduces a novel chaotic communication system model in which dimension-exponent mapping is applied to transforming original phase points. The improved prediction model is established by fitting new one-step evolutionary relationship between the transformed phase points and the next original phase point. The Euclidean distance of transformed phase points is regarded as criterion of relevant points. Compared to the original algorithm, the changes of new one-step evolutionary relationship tend to be steady over time. So the new model is easier for predicting. The Euclidean distance of the transformed phase points not only reflects relevancy between neighboring points and the focal point, but also expresses the relativity between different components of relevant points and the predicted value. Therefore, the Euclidean distance can reflect dynamic characteristics of predicted series correctly. Finally, chaotic time series of high embedding dimension is predicted by using this improved algorithm, the results of one-step prediction and multi-step prediction demonstrate the effectiveness of the algorithm.

Keywords: Chaotic communication system, High embedding, The Euclidean distance, Prediction.

1 Introduction

Recently chaotic theory is applied to many domains[1]. And the main research field is chaotic time series prediction algorithm based on phase space reconstruction[2]. However, the existing chaotic time series modeling methods and prediction technology are more suitable in low embedding dimension time series than in high embedding dimension. Sometimes the results of high embedding dimension time series are hard to be satisfied[3]. This paper introduces a novel model in which dimension-exponent mapping is applied to transforming original phase points. The improved prediction model is established by fitting new one-step evolutionary relationship between the transformed phase points and the next original phase point. And the Euclidean distance of transformed phase points is regarded as criterion of relevant points.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 91–95.
springerlink.com

2 Transforming Dimension-Exponent Mapping

$x_{n+1} = F(x_n)$ is a chaotic system, then a mapping $x_n' = g(x_n)$ and $x_{n+1} = G(x_n')$ is introduced to transform time series generated by formula $x_{n+1} = F(x_n)$, and $x_{n+1} = F(x_n) = G(g(x_n)) = G(x_n')$. In this way, the relation between x_n and x_{n+1} is replaced by the relation between x_n' and x_{n+1}. Then the absolute value of derivate changes correspondingly. Generally $g(x_n) = kx_n$, so k is the transforming coefficient. Take logistic chaotic mapping $x_{n+1} = h(x_n) = \mu x_n(1 - x_n)$ for example, $x_n' = g(x_n) = kx_n$ is used to transform. And $dH(x_n')/dx' = dh(x_n)/kdx$ is deduced, so $1/k$ is the transformed derivate. Before predicting chaotic time series, appropriate transforming mapping is applied to changing coordinate to smooth change of original one-step evolutionary relationship and reduce the complexity of prediction algorithm.

When k is greater than 1, separability of two close tracks will reduce. Relativity of series value goes down exponentially with delay time to chaotic time series. Moreover, relativity of each dimension component and prediction value in reconstruction phase space also goes down exponentially with delay time[4]. This means that the more relative dimension component is, the more important it is for predicting. Therefore, coefficient k of former components should be less than later components to counteract the adverse effect caused by increment of delay time. And coefficient k raises proportionately to Lyapunov index, and varies inversely as time delay.

Fig. 1 is principle of dimension-exponent mapping. x_n is the central point, and x_k, x_j, x_L are reference points of x_n. x_n', x_k', x_j', x_L' are new phase points after transforming. x_{k+1}, x_{j+1}, x_{L+1} are original one-step evolutionary points, and x_{n+1} is the new phase point. The model will be established after fitting one-step evolutionary relationship $G(x_n')$. Correlation dimension and Lyapunov index of attractor will not change after linear transforming[5]. Therefore, the Euclidean distance with new transforming points can be criterion of distinguishing phase points. The new distance can be expressed as Eq. 1.

$$d_i = [\sum_{n=1}^{m}(X_{k_l}^n - X_k^n)^2]^{\frac{1}{2}} = [\sum_{n=1}^{m} e^{\lambda(m-n+1)\tau}(X_{k_l}^n - X_k^n)^2]^{\frac{1}{2}} \quad (1)$$

The new distance formula describes not only correlation degree of different points but also relativity of different components and prediction value. Therefore, the new distance counteracts adverse effect caused by increment of delay time.

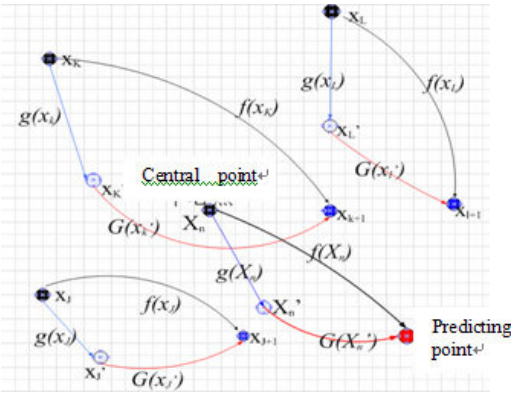


Fig. 1. Principle of dimension-exponent mapping

3 Process of Chaotic Time Series Prediction Based on Dimension-Exponent Mapping

The delay time τ is calculated by mutual information method. And correlation dimension d is calculated by G-P algorithm. Then embedding dimension m is determined by Takens theorem. The reconstruction phase space can be expressed as Eq. 2.

$$Y_t = (X_t, X_{t+\tau}, \cdots, X_{t+(m-1)\tau}) \subset R^m \tag{2}$$

The expression of linear coordinate transformation is as follow.

$$\begin{bmatrix} Y_1' \\ Y_2' \\ Y_3' \\ \vdots \\ Y_M' \end{bmatrix} = \begin{bmatrix} diag(Y_1) & 0 & 0 & 0 & 0 \\ 0 & diag(Y_2) & 0 & 0 & 0 \\ 0 & 0 & diag(Y_3) & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & diag(Y_M) \end{bmatrix}_{M*M} * \begin{bmatrix} A & 0 & 0 & 0 & 0 \\ 0 & A & 0 & 0 & 0 \\ 0 & 0 & A & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & A \end{bmatrix}_{M*M}$$

$diag(Y_i)$ is diagonal matrix composed of vector Y_i. Y_k' is transformed reconstruction vector, which can be expressed by Eq. 3.

$$\begin{aligned} Y_t' &= ((e^{m\lambda\tau})X_t, (e^{(m-1)\lambda\tau})X_{t+\tau}, \cdots, (e^{\lambda\tau})X_{t+(m-1)\tau}) \\ &= (X_t', X_{t+\tau}', \cdots, X_{t+(m-1)\tau}') \subset R^m \end{aligned} \tag{3}$$

Eq. 1 is used as criterion of distinguish reference points. $\hat{Y}_k' = \{Y_{k_1}', Y_{k_2}', \cdots Y_{k_q}'\}$ is reference vector set of Y_k', and d_i is distance between Y_{k_i}' and Y_k'. The weight of Y_{k_i}' is as Eq. 4, where $d_m = \min_{i=1,\cdots q}\{d_i\}$

$$P_i = \frac{e^{-\alpha(d_i - d_m)}}{\sum_{i=1}^{q} e^{-\alpha(d_i - d_m)}} \tag{4}$$

One-step weighting local linear fitting is as Eq. 5, where $e = (1,1,\cdots 1)^T$ and a, b are fitting real coefficients. And a, b can be calculated by Eq. 6. Prediction vector will be calculated by Eq. 6. The first vector of prediction vectors is the next prediction value of time series.

$$\begin{bmatrix} Y_{k_1+1} \\ Y_{k_2+1} \\ \vdots \\ Y_{k_q+1} \end{bmatrix} = \begin{bmatrix} e & Y_{k_1}' \\ e & Y_{k_2}' \\ \vdots & \vdots \\ e & Y_{k_q}' \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \tag{5}$$

$$\begin{cases} a\sum_{i=1}^{q} P_i X_{k_i}' + b\sum_{i=1}^{q} P_i X_{k_i}'^2 = \sum_{i=1}^{q} P_i X_{k_i}' X_{k_i+1} \\ a + b\sum_{i=1}^{q} P_i X_{k_i}' = \sum_{i=1}^{q} P_i X_{k_i+1} \end{cases} \tag{6}$$

4 Simulation

Logistic system generates two sets of data with different largest Lyapunov index, and Henon system generates a set of data. Each set has 3000 data, where the previous 2950 data is training samples and other 50 data is prediction samples. The chaotic parameters of simulation data are shown in Table 1. Table 2 shows inaccuracy of dimension-exponent method and original weighted methods.

Table 1. Chaotic parameters of simulation data

	Delay time	Embedding dimension	Largest Lyapunove index
Logistic 1	16	9	0.1443
Logistic 2	12	9	0.5645
Henon	15	9	0.3680

Table 2. Inaccuracy of improved method and original methods

	Improved method	Original method 1	Original method 2
Logistic 1	0.0143	0.2839	0.1964
Logistic 2	0.0030	0.4929	0.5564
Henon	0.2190	0.5319	0.7105

Table 2 indicates that dimension-exponent mapping algorithm has higher accuracy than original weighted methods. And the improved method can effectively predict chaotic time series.

5 Conclusion

This paper proposes a novel algorithm, in which dimension-exponent mapping is implied to transform actual phase points to phase changed points. The improved prediction model is established by fitting new one-step evolutionary relationship between the changed phase points and the actual phase points of the next moment. The Euclidean distance of phase changed points is regarded as criterion of relevant points. Compared to the original algorithm, the new one-step evolutionary relationship changes over time more slowly and is predicted easier. Euclidean distance of phase changed points not only distinguishes between different relevance of neighboring points to the focal point, but also expresses the prediction contribution of each dimension of neighboring points to predict. Chaotic time series of high embedding dimension is forecasted using this improved algorithm, the result of one-step prediction and multi-step prediction demonstrates the effectiveness of the algorithm.

References

1. Li, X.-H., Wang, Z.-X., Lv, T.-Y., Che, X.-J.: A Novel Incremental Clustering Algorithm Based on Chaos and Immune Response. Acta Automatica Sinica 36, 208–214 (2010)
2. Xie, J.-X., Chen, C.-T., Zhou, G.-H., Sun, Y.-M.: Noncoherent detection method for chaos-shift-keying communication system based on unscented kalman filt. Journal of Electronics and Information Technology 36, 684–689 (2008)
3. Zhou, Y.-D., Ma, H., et al.: Prediction of the chaotic time series using multivariate local polynomial regression. Acta Physica Sinica 56, 6809–6814 (2007)
4. Wang, H., He, J.-G., Hou, Z.-G.: Chaotic local adding-weight linear forecasting algorithm based on included angle cosine 35, 1483–1487 (2009)
5. Wang, Z., Zhao, C., Zhang, S., Zhao, Y.: Dimension-expont adding-weight one-rank local-region method for prediction of chaotic time series. Electrical Measurement & Instrumentation 47, 12–15 (2010)

Implementation of Automatic Electronic Control System for Temperature and Humidity in the Operating Room

Chunxiang Xu*, Wei Shi, and Haijun Song

College of Engineering Technology, Zhongzhou University, 6 Yingcai Street, Zhengzhou, Henan, 450044, China

Abstract. According to the system requirements for operating rooms and structure of the air conditioning system, the automatic temperature and humidity control is achieved based on EBI and electronic technology, which includes system network structure, air conditioning system composition and functions, CARE software programming, and EBI monitoring platform configuration design. It proves that the designed system with the friendly interface is able to meet the control requirements and has gone into service with the performance of high reliability in real mode.

Keywords: EBI, CARE, Temperature, Humidity, Air Conditioning System, Project Configuration, electronic technology.

1 Introduction

As the development of electronic techniques of information, controlling and communication, the management of information and automation for hospitals improves rapidly and many advanced and mature techniques are widely used for the environmental construction of the medical system, in which the air conditioning control system for the operating room is being adopted gradually as a significant component of the automatic building system. The project presented in the paper is based on some hospital of Hebei province of China, and adopts the industrial automation digital control system with the configuration platform of EBI by Honeywell Co. to achieve the control for temperature and humidity, and the management for field devices to meet demands for the environment of the operating rooms.

The EBI (Enterprise Building Integrator) is an integrated software to provide the automatic solution from the physical devices to the central monitor platform, and contains subsystems of the automatic building control, the safety and protection, the fire alarm, the digital video management, and possession and staff management. As the word leading system for the open building monitor, EBI meets customer's demands to the maximum extent and support the integration of field devices of BACnet, LonWorks, OPC, Modbus and C-Bus etc[1].

* Chunxiang Xu (1968-), female, associate professor, master's degree, engaged in research on the automatic control.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 97–102.
springerlink.com

2 The Function of Air Conditioning Electronic Control System for Operation Rooms

The environmental parameters controlled in operation rooms are temperature, relative humidity, press, clearness and freshness. The basic task of air conditioning system is to provide the function of leading airflow, filtering, cooling, drying, heating, and humidifying.

2.1 The Automatic Electronic Control Principle for Temperature and Humidity

Operation rooms demand the temperature and humidity to keep a constant value, and the control for the two significant parameters is implemented by DDC (Directly Digital Controller) and explained in Fig. 1.

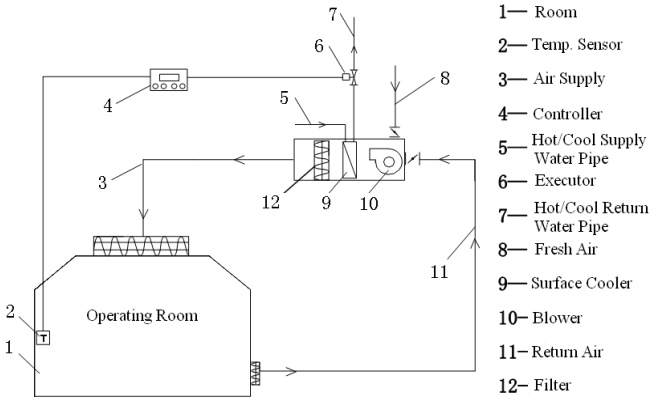


Fig. 1. Schematic of Temperature and Humidity Control for Operating Room

The combination of the fresh air and the return air flow correspondingly outside and from the room is sent onto the surface cooler by the blower and then into the room to keep the work state of cooling in summer. Depending on the preset rules and the parameter of temperature acquired on sensor, the DDC controller handle and adjust the water valve to keep the balance between the cold and the heat quantity correspondingly from the cooling water and the combinational air in real time to realize the constant indoor temperature.

2.2 System Requirements[2]

Two working modes are designed, summer mode and winter mode. In winter mode, it requires heating and humidifying generally by control the hot water valve and humidifier because parameters of temperature and humidity of the flow are lower than the set values, while it requires cooling and drying in the summer mode. The system requires higher priority for the humidity adjustment, which means the controller adjusts the traffic of cool water with the temperature no more than 8℃ to make the air cool and dry, and start the electrical heating device to make the compensation for temperature

error when it is very wet in summer. While the air humidity meets requirements, the system begins to work in the temperature controlling state automatically.

For the air conditioning system, it requires that the cycle unit works stably at a state of keeping the constant temperature and humidity during a whole year; and for the fresh air handling unit, it require to provide the humidity of 45% with the temperature of above 12℃ in summer and 8℃ in winter. The precision of temperature and relative humidity should be controlled less than ±1℃ and ±5%RH.

Field process requirements are indicated below. (a) Available control panel in the operating room for controlling the parameters of temperature and humidity, and the working state of units. (b)Available EBI server for the remote management to display all the monitored data dynamically, control parameters and working state, start the air conditioning system on time, make the self diagnosis and the repair alarm with timely printing, send maintenance indication on the accumulated working time and preset plan, provide the data management and analysis, inquire historical records, alert and archive, record and print running states, and provide interfaces for OS and maintenance etc. (c) Available outdoor control cabinet for manual handling, state displaying, parameter setting on the DDC panel.

2.3 Composition of the Air Conditioning System[3]

The air conditioning system consists of the blast pipe, the fresh air unit dealing the fresh flow with initial filtering and treatments of temperature and humidity, and the cycle unit operating on the flow from the fresh unit and the return flow from the operation room with the same deal above to control the parameters of temperature, humidity and pressure of the room and corridor within a allowable range. Four cycle units and one fresh unit is adopted in the project.

3 Electronic Structure of Air Conditioning System

The electronic control system based on EBI includes two levels, the field control point on physical layer and the central management server[4,5] showed in Fig. 2.

The field control point consists of all the meters and the DDC control node to realize the operations of comparison and calculation on the acquired data from sensors and detectors according to the set parameters and program within the DDC controller, send

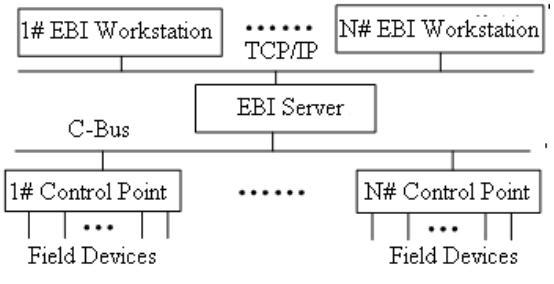


Fig. 2. Structure of the Control System

control signals and receive operations from the EBI server, panels of field cabinets and the operating room.

The central EBI server takes the task of monitoring and managing all the air conditioning system. All the controllers on the physical layer are connected by the Honeywell C-Bus and then to the EBI server by the RS485/232 converter. With the server, the optimized operations and document archive can be achieved.

The Ethernet interface of 100Mb/s based on TCP/IP is reserved for the internet connection to realize the remote management. And the control system can be embedded into the hospital automatic building control system when ready.

4 The Software Design

Referring to the system requirements and hardware schemes, the software of project configuration for the host computer is designed on the EBI platform while the slaver machine of DDC controller is programmed with Excel CARE.

4.1 Program and Function of Excel CARE

Excel CARE, a tool for Honeywell system integration and debug, provides the program and timing strategy for DDC in graphic mode. The program design includes creation of DDC, schematic design, data definition, switching logic and control strategy etc. with the design flow showed in Fig. 3.

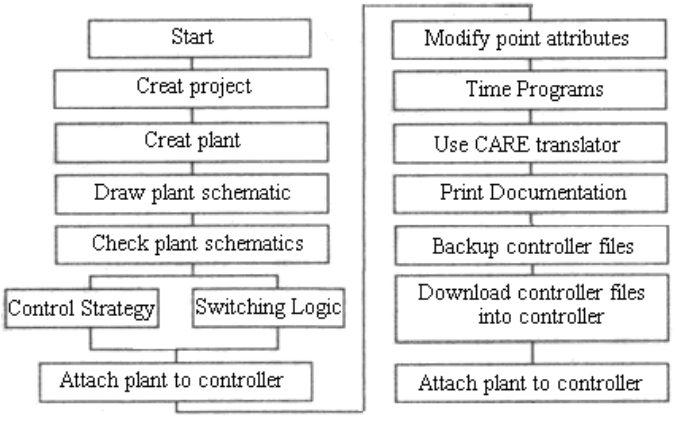


Fig. 3. Design Flow of Excel CARE

All the functions of CARE are based on device. Controllers with different volume can control one or more devices, however, not the same device. One device such as the air processor, boiler or cooling device is identified to be a separated controlled system.

A plant schematic consists of some segments in sequence to indicate different components and the relationship.

Data owns two types, the hardware data point from sensors and executor, and the software data point from calculation results of software. Also the analog point and the digital point, the input point and the output point are available.

The switching logic builds the relation of data points with DDC. As shown in Table 1 of the switching logic for the cool water valve, the valve signal for switching on is only available when the blower is running and has no faults in summer mode.

Table 1. Controlling Signals for Cool Water Valve

Function	Data Point	Value
State of Cool Water Valve	K3-CoolVlv	1
State of Blower	K3-RunStatus	1
State for Winter/Summer Switch	K3-WinterSummer	1
State of System Alarm	K3-SystemAlm	0

For the software control strategy, the regulation for temperature and humidity is controlled by PID which is widely used in the industrial control system because of the simple algorithm, good robust performance and high reliability. The control variable is determined by the difference between the setup and output value. The incremental PID expression is showed as the equations below.

$$\Delta u(k) = K_p \Delta e(k) + K_i e(k) + K_d [\Delta e(k) - \Delta e(k-1)] \tag{1}$$

$$K_i = K_p T / T_i \tag{2}$$

$$K_d = K_p T_d / T \tag{3}$$

The parameters of K_p, K_i, and K_d are the corresponding coefficient for proportion, integration and differentiation in the equation (1). In equation (2) and (3), T, T_i and T_d indicate the sample period, integral time and differential time correspondingly.

The incremental PID control is achieved by the internal PID regulator embedded in the Honeywell DDC. Fig. 4 shows the parameter setup of the standard PID regulator on CARE for the valve control.

The common PID controller can not achieve good performance for the nonlinear and time-varying system and the fuzzy model system. However, CARE also provides the enhanced controller of EPID which can keep the controlled item at a setup value to

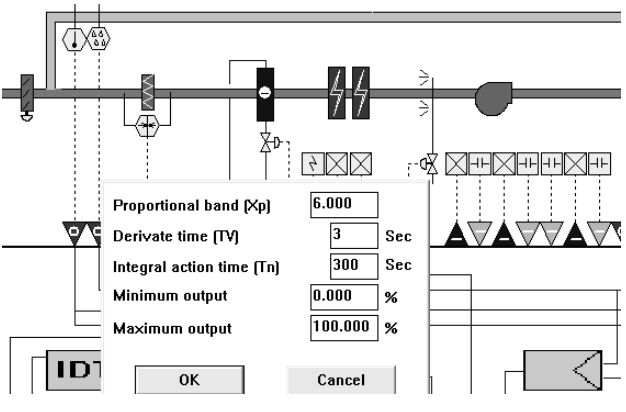


Fig. 4. Parameter Setup of Standard PID Regulator for the Cool Water Valve

avoid the floating change along the setup value at the PID beginning period. More about EPID can be gotten on document[7].

4.2 Software Design on EBI Server[5]

The EBI softer platform includes Station, Quick Builder and Display Builder. The design flow on EBI for the device monitor is showed below.

Create a new project in Quick Builder to define channels, controllers, points, stations and servers, setup configuration and then download data into EBI real database. The project configuration in the paper is showed in Fig. 5, which contains five DDC controllers and one channel.

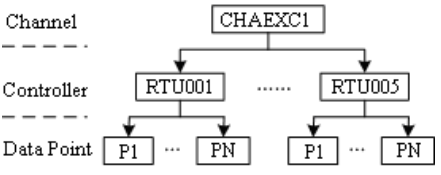


Fig. 5. Structure of Project Configuration

Design the human machine interface for each DDC with EBI Display Builder, where all the control points are connected in the graphic interface.

Achieve access and operation to the controlled points in EBI database to monitor devices and sub-systems by executing the human machine interface in the environment of Station.

5 Conclusion

The control system based on EBI and CARE for temperature and humidity in the operating room has advantages of high precision and stability in real mode. The system designed in the paper has gone into service and been working stably, which brings good project and application value.

References

1. Yang, S., Xing, L.: Discussion about Application of LonWorks Fieldbus in Intelligent Building. Low Voltage Apparatus 18, 11–13 (2007)
2. Zhang, G., Zhu, J.: The Design and Realization of Air Condition Control System for Clear Operating Room. Modern Electronic Technique 1, 97–100 (2009)
3. Wang, Y.: Intelligent Building Technique, pp. 30–35. Post and Telecom Press (2008)
4. Ge, S., Ma, F.: Establishment of the Open Control System of Air Conditioners Based on LonWorks. Journal of Hefei University of Technology 28(5), 558–561 (2005)
5. Honeywell Limited Australia: EBI Overview (2003), http://www.honeywell.com.au
6. Honeywell Inc.: en1b0101-ge51r0909g (2009), http://ecc.emea.honeywell.com
7. Honeywell Inc.: Excel live CARE User Guide (2004), http://honeywell.com
8. Jia, X.: Application Set Condition Software Coming True the System Control of New Wind. Journal of Jilin Institute of Architectural & Civil Engineering 25(4), 6–65 (2008)

A Resistor-Less CMOS Voltage Reference

Qianneng Zhou[1,2], Hongjuan Li[3], Li Wang[1], Qi Li[1], and Qiulin Zhang[1]

[1] College of Electronic Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
[2] National Labs of Analog Integrated Circuit, Chongqing 400060, China
[3] College of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
zhouqn@cqupt.edu.cn

Abstract. A CMOS voltage reference without resistor is designed and simulated in SMIC 0.18μm CMOS process in this paper. Simulation results show that the power supply rejection (PSRR) of the voltage reference at 10Hz, 100kHz and 10MHz achieves, respectively, 44dB, 44dB and 37dB. When temperature is in range from -50 to 99°C, the voltage reference achieves an output voltage of about 1.21V with a variation of ±25mV.

Keywords: CMOS voltage reference, current source circuit, resistor-less.

1 Introduction

Voltage references are one of important blocks in many analog, digital, and mixed-signal circuits and systems. It should be free from variations of fabrication process, power supply voltage, and temperature. Actually, in order to achieve a good temperature perform, bandgap voltage reference is one of the most popular voltage reference, which adopts a weighted sum of the forward bias voltage across a p-n diode and the thermal voltage V_t [1,2]. The relative weighting is commonly adjusted by trimming the ratio of the resistors. The presence of resistors is a drawback for some applications. In a standard digital CMOS process, the model of resistors is not accurate. Moreover, using resistors increases not only the chip area but also the noise coupling from substrate. As an alternative, CMOS resistor-less voltage references have been reported [3,4,5].

In this paper, a resistor-less CMOS voltage reference, which only employs MOS transistors and parasitic vertical pnp transistors in standard CMOS process, is presented. Furthermore, the resistor-less CMOS voltage reference has a simple architecture and achieves a well performance.

2 Analysis and Design of Voltage Reference

The circuits of the designed resistor-less CMOS voltage reference is shown in Fig.1, and all MOS transistors adopt the long channel transistor so that the channel-length modulation effect is negligibly small. The resistor-less CMOS voltage reference is

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 103–107.
springerlink.com

made up of the core circuit of voltage reference, current source circuit and a start-up circuit. The current source circuit will provide a current proportional to T^{α}, where T is the absolute temperature and α is a constant. The current source circuit has two possible equilibrium points, so a start-up circuit is necessary and is made up of transistor M_{s1}~M_{s5}. The core circuit of voltage reference will provide an output reference voltage, which has a good temperature. For convenience, in this paper, it is assumed that the drain current of transistor M_i is I_i, W_i and L_i are, respectively, the channel width and length of transistor M_i.

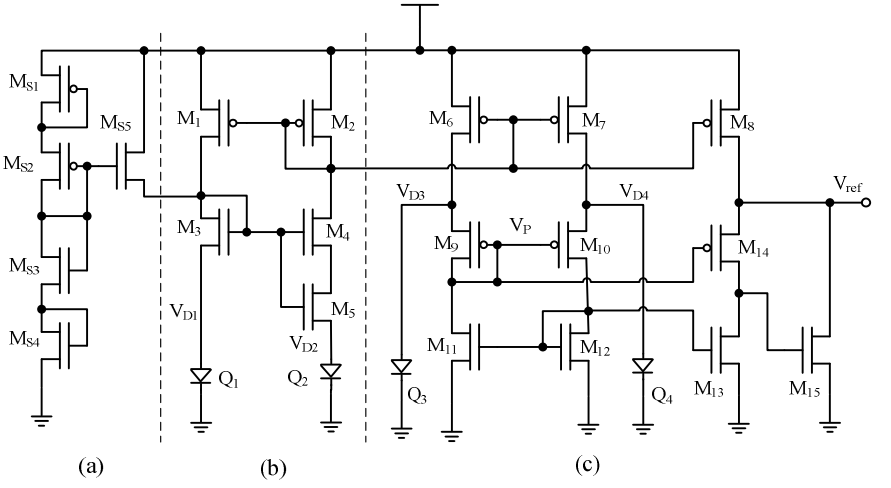


Fig. 1. Circuit of the resistor-less voltage reference (a) Start-up circuit; (b) Circuit of current source; (c) Core circuit of voltage reference

2.1 Current Source Circuit

The current source circuit is shown in Fig.1, which is made up of transistors M_1~M_5 and vertical pnp transistors Q_1~Q_2. Q_1 and Q_2 are entirely same, whose base and collector are tied together to form a diode respectively. M_1~M_4 operate in the saturation region, M_5 is in the triode region. The aspect ratios of M_1 and M_3 are larger than those of M_2 and M_4, respectively, by a factor N. So, the drain currents I_1 and I_2 have that I_1=N·I_2, and the source voltage of M_3 is equal to that of M_4. In fact, given a bias current I_c in the pnp transistor, the forward voltage V_D of diode can be given as

$$V_D = \frac{kT}{q} \ln \frac{I_c}{I_s} \tag{1}$$

where, k is Boltzmann's constant, q is electronic charge, T is absolute temperature, and I_s is the saturation current. The mobility of an electron μ_n is a temperature function with a reference temperature T_r, and which can be modeled as[6]

$$\mu_n(T) = \mu_n(T_r) \times (T/T_r)^{-\beta_{\mu n}} \tag{2}$$

where $\mu_n(T_r)$ and $\beta_{\mu n}$ are, respectively, the mobility of an electron at temperature T_r and the mobility exponent of an electron. According to the circuit of Fig.1, the drain current of M_2, i.e. I_2, can be expressed as

$$I_2 = \frac{C_{ox}}{2}\left(\frac{W_5}{L_5}\right)^2\left(\sqrt{\frac{L_4}{W_4}}+\sqrt{\frac{L_4}{W_4}+\frac{W_5}{L_5}}\right)\times\left(\frac{k}{q}\ln N\right)\times\frac{\mu_n(T_r)}{T_r^{-\beta_{\mu n}}}\times T^{2-\beta_{\mu n}} = G\times T^{\alpha} \tag{3}$$

where C_{ox} is the gate oxide capacitance per unit area, $\alpha=2-\beta_{\mu n}$ and G is a constant. This equation indicates that the current I_2 is proportional to T^{α}.

2.2 Core Circuit of Voltage Reference

As shown in Fig.1, transistors M_6~M_{15} and vertical pnp transistors Q_3~Q_4 form the core circuit of voltage reference, and transistor M_6~M_{15} are all in saturation region. It is similar to that reported in Ref. [4] but the bias circuit is replaced by a current source circuit. To form a diode, the base and collector of Q_3 and Q_4 are tied together respectively. M_{11} and M_{12} are entirely same, so they have the same drain current, i.e. $I_{11}=I_{12}=I_9=I_{10}$. M_2-M_6, M_2-M_7 and M_2-M_8 form, respectively, current-mirror pairs. The channel lengths of M_9 and M_{10} are the same, but the channel width of M_9 (i.e. W_9) is the greater than that of M_{10} (i.e. W_{10}). Here A is the ratio W_9/W_{10}. V_{D3} and V_{D4} are, respectively, the forward bias voltage of Q_3 and Q_4. According to the circuits of Fig.1, the gate voltage of M_9, i.e. V_P, can be expressed as

$$V_P = \frac{\sqrt{A}V_{D3} - V_{D4}}{\sqrt{A}-1} - |V_{TP}| \tag{4}$$

where V_{TP} is the threshold voltage of PMOS transistor. M_{12} and M_{13} form current mirror pair, which have that $(W/L)_{13}=B\cdot(W/L)_{12}$, so their drain currents have $I_{13}=B\cdot I_{12}$. Transistor M_{14} and M_9 form current mirror pair, whose width-length ratios have that $(W/L)_{14}=P\cdot(W/L)_{10}= P\cdot B^{-1}\cdot(W/L)_9$. Here, P>1 and P<B. According to the circuit of Fig.1, the output voltage V_{ref} can be expressed as

$$V_{ref} = V_{D4} + (\frac{\sqrt{B}}{\sqrt{P}}-1)\frac{\sqrt{A}}{\sqrt{A}-1}(V_{D4}-V_{D3}) \tag{5}$$

where, the parameters have B>P and A>1. For the circuit in Fig.1, Q_3 has an emitter area that is M times that of Q_3. The forward bias of Q_4 (i.e. I_{Q4}) is larger than that of Q_3. In equation 5, the second term is proportional to the absolute temperature T, and V_{D4} has a negative temperature coefficient. So, the $(\partial V_{ref}/\partial T)_{T=Tr}=0$, where T_r is room temperature, can be obtained by optimizing the circuit parameters. However, the temperature dependence of V_{D4} and V_{D3} is not perfectly linear and an absolutely temperature-independent V_{ref} is not possible in the temperature range, so a nonlinear temperature- dependent error voltage appears at V_{ref}.

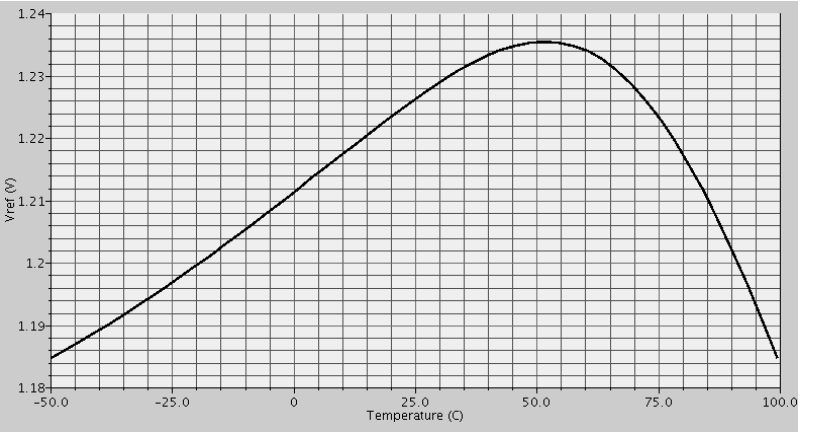


Fig. 2. Simulated Output voltage versus temperature

3 Simulation Result

The CMOS voltage reference is designed and simulated in SMIC 0.18μm CMOS process. Fig.2 shows the output voltage V_{ref} of the designed voltage reference circuit as a function of temperature with a 3V power supply. Simulation result indicates that output voltage V_{ref} has only a variation of 50mV with temperature ranging from -50 to 99°C. The simulated power supply rejection (PSRR) is shown in Fig.3. The PSRR at 10Hz, 100 kHz and 10 MHz achieves, respectively, 44dB, 44dB and 37dB. Fig.4 shows the output voltage V_{ref} as a function of the power supply. The output voltage V_{ref} has very small variation when power supply voltage V_{dd} changed from 3V to 5V.

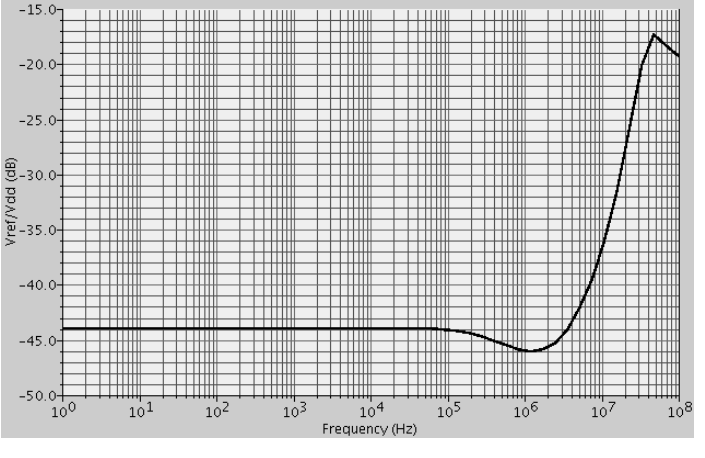


Fig. 3. Simulated PSRR versus frequency

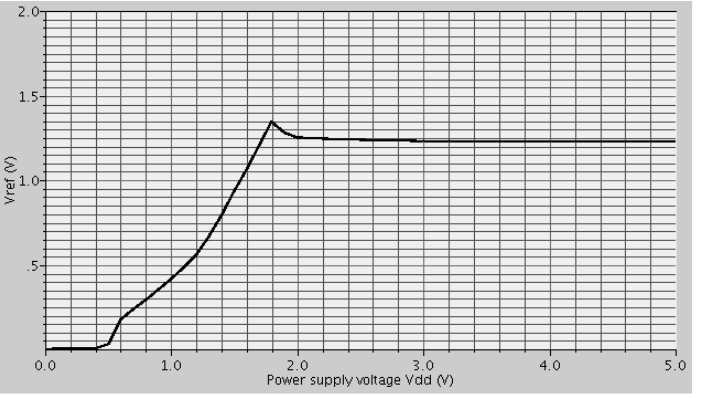


Fig. 4. Simulated output voltage versus power supply voltage V_{dd}

4 Conclusion

A CMOS resistor-less voltage reference, which only employs MOS transistors and parasitic vertical pnp transistors, is designed and analyzed in SMIC 0.18μm CMOS process. Simulation results show that the voltage reference achieves excellent stability, relatively small temperature dependency and relatively well PSRR. The voltage reference is well suited for system-on-chop applications.

Acknowledgments. Supported by Natural Science Foundation Project of CQ CSTC (cstcjjA40011) and Supported by Dr. Start Fund of Chongqing University of Posts and Telecommunications.

References

1. Song, B.-S., Gray, P.R.: A Precision Curvature-Compensated CMOS Bandgap Reference. IEEE J. Solid-State Circuits SC-18(6), 634–643 (1983)
2. Andreou, C.M., Koundounas, S., Georgiou, J.: A Novel Wide-Temperature-Range, 3.9ppm/°C CMOS Bandgap Reference Circuit. IEEE J. Solid-State Circuits 47(2), 574–581 (2012)
3. Buck, A.E., McDonald, C.L., Lewis, S.H., Viswanathan, T.R.: A CMOS Bandgap Reference Without Resistors. IEEE J. Solid-State Circuits 37(1), 81–83 (2002)
4. Adriana, B.G., Viswanathan, T.L., Viswanathan, T.R.: A Low-supply-Voltage CMOS Sub-Bandgap Reference. IEEE Trans. Circuits Syst. II, Exp. Briefs 55(7), 609–613 (2008)
5. Ming, X., Ma, Y.Q., Zhou, Z.K., Zhang, B.: A High-Precision Compensated CMOS Bandgap Voltage Reference without resistors. IEEE Trans. Circuits Syst. II, Exp. Briefs 57(10), 767–771 (2010)
6. Tsividis, Y.: Operation and Modeling of MOS Transitor, 2nd edn. McGraw-Hill, New York (1998)

Design and Application of Distance Measure Ultrasonic Sensor

Shuhai Wang[1], Qiuzhen Liu[2], Shuwang Chen[1], and Yuxi Xue[1]

[1] Institute of Information Science & Engineerning, Hebei University of Science and Technology, Shijiazhuang, 050018, China
[2] Tonghe Thermoelectric Co., Ltd. Qinhuangdao, 066012, China
xueyuxixyx@163.com

Abstract. An ultrasonic sensor to detect the distance measurement system is introduced in the paper. The paper presents the working principle of ultrasonic distance measurement, the structure properties of ultrasonic sensors, the ultrasonic sensors in the transmitter and receiver probe and their inter-action problem in details. Ultrasonic sensor is a component of detecting the distance. STC89C52 SCM (single chip machine) is a control component. The system mainly consists of five parts, such as the smallest single chip system, the ultrasonic transmission circuit, the ultrasonic receiving circuit, the digital display circuit and the alarm circuit. Application software of ranging is written. The accuracy and error in the other related issues of the distance ultrasonic sensor is studied by experiment. The system is designed for portable, low power consumption and high precision. It can be applied to water level measurement, robot obstacle avoidance, parking sensor and other areas.

Keywords: Ultrasonic sensor, Distance, Measurement.

1 Introduction

As ultrasonic technology continues to be researched deeply, it has been more progress and development than before. It enhances performance in many areas, such as in precision, non-contact, non-destructive and other aspects. With the technological breakthroughs in these areas, ultrasonic sensor is applied more and more widely. Ultrasonic sensor has its unique advantages in some respects about the distance measurement, such as level detection, robot avoid obstacle, parking sensors, and so on. It has the unique advantages in the toxic, harmful and corrosive environments.

2 System Design

According to the design requirements and tasks, the choice in the program must be paid attention to the operation and price. There are many tools and methods in measuring distance in present, such as laser ranging, infrared ranging, ultrasonic ranging and

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 109–114.
springerlink.com

moreadvanced satellite ranging, and so on. The advantage of laser ranging is the good color, strong direction, very long distances measurement. Its disadvantage is the blind spot in 15 meters. The advantage of infrared ranging is long distance measurement. Measurement distance can reach to 1-5 kilometers. Its disadvantage is higher price than ultrasonic ranging. The advantage of the satellite location is the farthest distance measurement. But the price of satellite ranging is the most expensive. It is applied in national defense and the military often. Cheap is the biggest advantage of ultrasonic ranging. Its circuit is designed simply and is operated easily. It can achieve centimeter-level accuracy [1]. But the measurement distance is relatively close, usually between a few meters and dozens of meters.

2.1 Ultrasonic Principle

Ultrasonic wave means the sound wave whose frequency is more than 20 kHz. Ultrasonic principle is to use two ultrasonic transducer (also called ultrasonic probe which includes sending probe and receiving probe) to complete the distance measurement under the control of SCM. Ultrasonic sending probe is the application of converse piezoelectric effect of the piezoelectric crystal materials, so the electrical energy is transformed into mechanical vibration and sends the ultrasonic wave out. Ultrasonic receiving probe is the application of the piezoelectric effect of the piezoelectric crystal materials, so the mechanical vibration is turned into electrical signals. Conversion chart of ultrasonic sensor is shown in Fig.1. The method of the distance calculation is described as followed. At a certain moment, the probe in ultrasonic sensor sends an ultrasonic signal. When this ultrasonic signal meets obstacles, it will return. The ultrasonic receiving probe receives the signal. The spread time of the ultrasonic wave from the sending probe to the obstacle and its return time can be gained. The distance between the ultrasonic probe and the obstacle can be calculated [2]. The calculation of distance is shown in Eq.1.

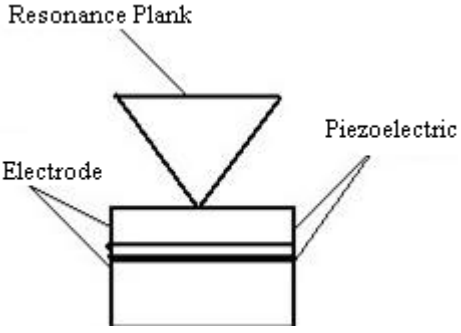


Fig. 1. Conversion Chart of Ultrasonic Sensor

$$d=s/2= (vt/2). \quad (1)$$

Where d is the distance between the ultrasonic probe and the obstacle, and s is the whole trip distance of ultrasonic wave spread, and v is speed of the ultrasonic wave in the medium, and t is the ultrasonic wave spared time in the round trip. Ultrasonic speed in the air is shown in Table 1.

Table 1. Some Temperature of the Sound Velocity

Temperature T[℃]	-30	-20	-10	0	10	20	30
Sound velocity v[m·s^{-1}]	313	319	322	331	337	344	350

3 Hardware Design

This system design of the hardware circuit includes the single chip minimize system, the ultrasonic sending circuit, the ultrasonic receiving circuit, the display circuit and the alarm circuit. The five parts contain a whole system under the control of SCM. It completes the design requirement and function [3]. The diagram of the hardware circuit is shown in Fig.2.

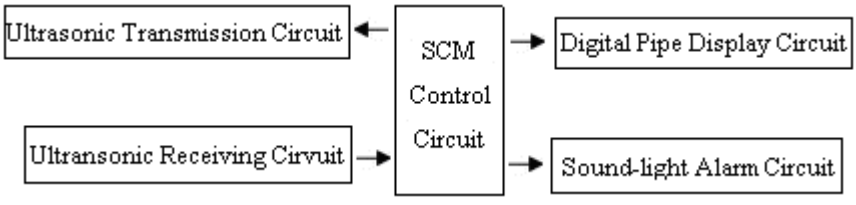


Fig. 2. Hardware Design Diagram

3.1 SCM Minimum System

SCM system contains the minimum micro-controller (STC89C52), the crystals (12 MHz) and the reset circuit. The STC89C52 chip controls the whole circuit. The main function of the reset circuit provides the SCM protection. The design of the reset circuit with a resistor can effectively solve the battery voltage problems. The drops of the battery voltage can discharge quickly by the capacitor. The state of the voltage can not make the circuit to stop working. The design of resetting with button makes SCM resetting. When the button is pressed, the system is reset. Crystals of this circuit is the clock source [4].

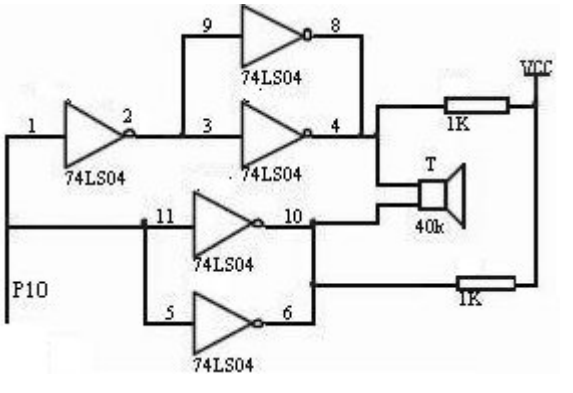


Fig. 3. Ultrasonic Transmission Circuit

3.2 Ultrasonic Transmission Circuit

Ultrasonic sending circuit mainly contains the chip of 74LS04 and ultrasonic sending probe [5]. Ultrasonic sending circuit is shown in Fig.3. The chip of 74LS04 contains a 6-chip inverter. Inverter role is to promote the ability of the circuit with load. Micro-controller port P1.0 outputs a square wave signal. The ultrasonic sending probe sends ultrasonic wave. Outputs of the two parallels inverter improve greatly their on loading abilities. Two pull-up resistors contain two effects. They can improve the loading capability of the 74LS04 chip. On the other hand, they can improve the damping effect of ultrasonic probes. So the sending probe reduces its free oscillation time.

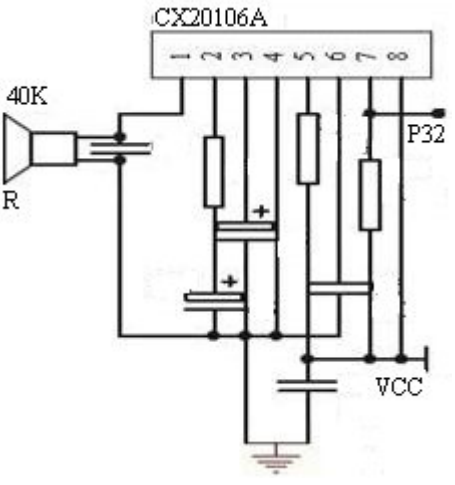


Fig. 4. Ultrasonic Receiving Circuit

3.3 Ultrasonic Receiving Circuit

Ultrasonic receiving circuit contains an ultrasonic receiving probe and a CX20106A chip [6]. The CX20106A chip is a special infrared receiver chip. Ultrasonic receiving probe is the application of the piezoelectric effect of the piezoelectric crystal materials, so the mechanical vibration is turned into electrical signals. The ultrasonic receiving circuit is shown in Fig.4.

3.4 Display Circuit

With the development of the society, all kinds of display device emerge on market now, such as the LED display, the LCD display, the LED nixie tube, and so on.

Nixie tube display contains static display and dynamic display. Static display is the each pin of nixie tube controlling by I/O port of SCM. The biggest advantage of static display is easy to program. But this kind of display mode uses the I/O port too much. Dynamic display is that all of the nixie tube of the corresponding pins connected together with controlled by the I/O port. The display mode can greatly save the SCM pins. The

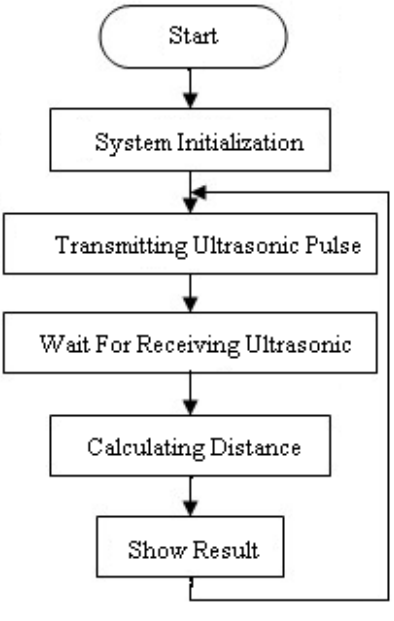


Fig. 5. The Software Design Flow Chart

design of the display circuit is a total of four anodes LED nixie tube and dynamic display. [7].

3.5 Alarm Circuits

The alarm circuit contains two parts. They are sound alarm and light alarm. It contains a buzzer and a light emitting diode (LED). This kind of alarm circuit compares strong light environment and large noise environment. We know that the alarm circuit is working. The alarm circuit protects the system greatly.

4 Software Design

The software design of the system is very important. The circuit modules of the system work together by the designing software. The designing software is used by a simple modular design approach [8]. The software design flow chart is shown in Fig.5. The software programs of the system contains main program of the design, the interrupt service routine design, the display design, distance calculation design and alarm design. The ultrasonic sending circuit is connected to the single chip P1.0 port. The ultrasonic receiving circuit is connected to the P3.2 port. Each pin of the display circuit is connected between the P0.0 ~ P0.7 port and P2.4 ~ P2.7 port. The alarming circuit chip is connected to the P2.0 port.

5 Data Analysis

Experimental data are recorded in Table 2.

Table 2. Experimental Data

Times	1	2	3	4	5	6	7	8	9
Actual range [cm]	99	100	100	104	105	110	120	148	155
Measure Distance [cm]	101	101	102	106	107	109	121	149	157

The smallest distance of measurement is about 30 cm. It is the mainly the influence between the ultrasonic sending probe and the ultrasonic receiving probe. The ultrasonic sending probe sends the ultrasonic wave conical but not the cylindrical. When the ultrasonic receiving probe receives the ultrasonic wave, we must consider the influence of the wave of sending. This is the reason that we do not measure the distance from zero centimeter.

6 Conclusions

The accuracy of the ultrasonic ranging design can reach to cm level. The measurement error is ±2 cm. The measure distance is about 5m. The size of the system is small to carry easily. Its supply is 5V D.C. As its low power consumption, the system can be used in low carbon and environmental protection application.

References

1. Yang, X., Liu, X.: Design of Ultrasonic Wave Distance Measuring System Based on Microprocessor. Machine Tool & Hydraulics 39(8), 106–108 (2011)
2. Guo, Q.: Ultrasonic Ranging and Anti-collision System Based 0n STC89C52. Instrument Technique & Sensor (6), 74–77 (2011)
3. Li, S.: Design of Ultrasonic Wave Distance Measuring System Based on AT89S52. Automation & Instrumentation (6), 44–47 (2009)
4. Wu, Q., Feng, W., Ma, W.: Design and Implementation of the Automobile Reversing Radar System. Modern Electronic Technique 32(9), 191–194 (2009)
5. Li, X., Tian, W.: The Design of Car Reversing Anti-Collision Warning System. Journal of Taiyuan University of Science and Technology 32(3), 172–176 (2011)
6. Zhang, H., Liu, Y.: Ultrasonic Ranging Back-dram Anti-collision Alarm System Based on AT89C52. Modern Electronic Technique (3), 205–206 (2010)
7. Sun, Y.: Design of Ultrasonic Distance Measuring System with High Precision. Mechanical Management and Development 23(6), 102–103 (2008)
8. Chen, X., Zhou, X.: Design and Realization of Ultrasonic Range Finder. China Measurement & Testing Technology 34(6), 120–122 (2008)

The Application of Quartz Crystal Micro Balance in Electronic Engineering

Xuepeng Liu and Dongmei Zhao

Zhongshan Polytechnic, P.R. China
liuxuepeng1026@yahoo.com.cn, zdmeihn@163.com

Abstract. The basic theory of quartz crystal microbalance (QCM) is simply introduced and its applications are discussed in areas of biology, chemistry and other aspects including immunosensor, DNA sensor, and electrochemical quartz crystal microbalance. The paper introduces its application in electronic engineering. The research tendency is also presented in the future

Keywords: Quartz crystal microbalance, immunosensor, DNA sensor, electrochemical quartz crystal microbalance.

1 Introduction

Quartz crystal micro (QCM) balance is a kind of quality sensors newly developed in recent years. Its advantage is high precision (minimum 1 ng/cm2), good stability in low temperature [1-4], and with the development of further research, the application of QCM has been greatly expanded. It not only can be used to measure quality, also can measure other physical properties, so in the physical, chemical and biological aspects it has been widely used

2 The Theoretical Basis of QCM

A QCM is combined by a micro electron wafer and two electrodes added the above together.

Quartz crystal micro balance sauerbrey equation in the gas phase $\Delta f = \frac{-2 f_0^2 \Delta m}{A\sqrt{u_q \rho_q}} = -c\Delta m$. Quality effect theory $\Delta f = -[f_0{}^2/(\rho_q N)][\Delta m/(1+\Delta m/m)]$. Coupling theory $\Delta f = -f_0^{3/2}(\frac{\rho_L \eta}{\pi u_q \rho_q})^{1/2}$

The relationship of Quartz crystal micro balance impedance and the frequency in the liquid medium

$$R_1 = (2\pi f \rho_L \eta)^{1/2} A / k^2$$

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 115–118.
springerlink.com

3 Quartz Crystal Micro (QCM) Balance in the Biological Applications

Piezoelectric biological sensors are a new type of biological sensors fixing the biological sensitive components on quartz crystal. It has the characteristic of high selectivity of biological materials and high sensitivity of piezoelectric sensors. Compared with the traditional biological analysis method, the sensors is a simpler device, with fast analysis speed, high sensitivity, no marker, and automation.

Biomedical quartz crystal sensor is composed by quartz crystal probe, signal detection circuit and data processing system, etc. Quartz crystal structure and combination of probe according to the needs of the study have different structure and combinations. For example QCM for liquid system can use single touch liquid or double touch liquid style, also a tandem (SPQC), liquid format (ESPS), array type form.

Frequency stability depends mainly on quartz crystal resonant frequency. And in liquid, liquid damping effect make crystal resonator phase characteristic of slow down, quality factor, which should be paid attention to the resonance frequency of the crystal, circuit performance, amplification link to phase characteristic.

When measuring the reference crystal probe into will not contain remains to be measured in the solution of the target samples, will examine crystals in measuring target to probe the sample containing the solution, because the quality of different reaction, produce change such as different also, the two oscillating circuit of the frequency of the signal output is different also. Through on time-multiplier by composed of frequency mixing circuit can detect both difference frequency signal Δ f. Δ f difference frequency signal sent by the plastic surgery to enlarge, dedicated controller or computer data processing system composed of data processing, get measurement analysis results.

With basic crystals produce of the oscillation signal as the benchmark signals, detection crystals produce of the oscillation signal as detection signal, both for the correlation processing, get a reaction tiny changes output signal. For example, the benchmark crystals for 9 MHZ, detection crystals is also 9 MHZ, the benchmark oscillation signal for 9 MHZ, oscillating signal detection for 9 MHZ-Δ F (Δ < 9 F), will signal MHZ correlation processing, get two road oscillating signal frequency difference of Δ F.

Frequency difference Δ F can correct response of different detection markers, the frequency, suitable for testing smaller values, but voltage value to the triangle wave, we must carry on the plastic surgery to enlarge, make its output for 5 v square wave signal to meet the standard, computer test requirements

Use computer parallel detection signal, parallel signal through series and parallel to register, and with the parallel port parallel port to the constant change of the signal, parallel of the register with potential changes, and increasingly refreshed. Through the CPU out instructions, every once in a while will parallel of the register to potential dynamic memory a units. This is achieved by computer to signal sampling work.

4 The Application of Quartz Crystal DNA Sensors

Quartz crystal DNA biosensors are combined by the single nucleotide sequences of DNA (also SsDNA probe) electrode and transducer two parts. Quartz crystal DNA biosensors for quartz crystal sensor matrix and its mechanism is in with a metal membrane surface, fixed a connection, will be known that the single nucleotide sequences of DNA (also called ssDNA probe) fixed forming on the surface sensor electrode, sensor electrode ssDNA probe and stay on the test sample target DNA hybrid, form double chain DNA (dsDNA), hybrid reaction in the sensor electrode done directly, transducer will hybrid produced during the change such as the quality of the converted into frequency signal, oscillation frequency with electrodes on the surface of the increase and decrease, the quality of the sensitivity, g level, through the test frequency changes to achieve the purpose of measurement. Because of the electrode ssDNA probe structure can be different, that form a different quartz crystal DNA biosensor type.

Quartz crystal biological sensors construction and testing process normally steps:

(1) Crystal surface functionalization. Through the chemical reaction to crystal surface suitable for sensitive membrane materials; connection

(2) Carrier film and the DNA probe fixed;

(3) The DNA probe activation and balance;

(4) Hybrid and testing;

(5) Sensitive membrane regeneration, namely using chemical reagent method makes crystal surface hybrid of double-stranded DNA has restored as an molecular degeneration, chain-free, can be reused.

5 Electrochemical Quartz Crystal Micro Balance

Electrochemical quartz crystal micro balances (electrochemical quartz Crystal microbalance, EQCM), mainly include crystals, oscillating circuit and frequency measurement of three parts. In EQCM, piezoelectric quartz crystal single solution and contact, touch the electrode is liquid quartz crystal incentive electrode electrochemical system work, is also the electrode, crystal surface is electrochemical places. Piezoelectric quartz crystal on the other side is in air, so on one hand is to reduce the vibration energy dielectric loss, reduce the conductivity of the influence of the solution changes, more important is to ensure that EQCM in high concentration of electrolyte can support work, because of quartz crystal can provide an additional process quality information of the gain and loss of electrode, can help to elucidate the mechanism of electrode process.

EQCM research focuses on metal deposition and dissolution process, chemically modified membrane electrochemical properties

The application of Piezoelectric sensors in chemical solution include that measurement Based on the quality of change, analysis applications on Viscosity response pattern, frequency shift analysis of conductivity of change.

6 The Temperature Self-detection and Compensation of Quartz Resonant Force Sensor

For many cut type resonators, AT-cut resonators due to produce simple, high sensitivity and the use of the most widely used. AT-cut resonators sensitivity is a function of temperature AT room temperature (less affected by temperature), so when AT-cut resonators as force sensitive components must take into account the effect of temperature. Use with stress and temperature compensation ability of the SC double cut resonators c model experiments of measurement temperature new methods, and good results have been achieved. The method and mode frequency of incentive c and three times, using a frequency sound 3 times frequency and three times the sound difference frequency measurement temperature, realize the crystal oscillator microprocessor temperature compensation. This method was applied successfully to temperature compensation chip design and quartz crystal micro balance temperature compensation.

7 Quartz Crystal Micro Balance Development Tendency

The trend of Quartz crystal micro balance concentrated in tandem piezoelectric sensors, tandem surface acoustic conductance, liquid every sensor electrode type piezoelectric sensors, double channel density of the piezoelectric sensor. In these areas attention should be paid to development theory, in order to play a greater role.

References

1. Sullivan, C.K.O., Guilbault, G.G.: Commercial quartz crystal microbalances-theory and applications. Biosensors & Bioelectronics 14, 663–670 (1999)
2. Wang, X., Ding, B.: A highly sensitive humidity sensor based on a nanofibrous membrane coated. Nanotechnology 21, 663–670 (2010)
3. Dessler, A.E., Sherwood, S.C.: 323, 1020–1032 (2000)
4. Zhou, X.F., Zhang, J.: Sensors Actuators 135, 209–217 (2007)

Super Large Optical Mirror Surface Form Nano-positioning Technology Based on Displacement Sensor Detection

Xuepeng Liu and Dongmei Zhao

Zhongshan Polytechnic, P.R. China,
69121 Heidelberg, Germany
liuxuepeng1026@yahoo.com.cn, zdmeihn@163.com

Abstract. Nano-positioning technology is one of the key technologies in frontier science, engineering technology. The control method and principle of super large optical mirror surface form nano-positioning technology are analyzed. For measurement and control system, the sensor design and the design for the parameters of the electric displacement detector are discussed. Positioning precision is: surface form control error is just 50nm; drift error is lower than 5nm.

Keywords: Nano positioning technology, super large optical mirror surface form nano-positioning technology, flexure hinge, sensor design, displacement sensor detection.

1 Introduction

With the development of science and technology, space technology, microelectronics engineering, measure science and technology, optical and photoelectron engineering, precision engineering, biological engineering, nano science and technology urgently need nano positioning technology. Nano positioning technology has become a frontier science, engineering technology, one of the key technologies. In the present the prosperous development of the nano science and technology field, nano positioning technology and Scanning Probe Microscopy techniques are inseparable, has become a Scanning Probe Microscopy (SPM), is one of the core technology to realize the processing and nano measurement of nano key technology, nano positioning technology has become a nano measurement and atomic operation engineering research and to the premise condition of industrialization and the work foundation. Nano science and technology development depend on nano location technology development, the urgent need to nano location technology development. The existing micro displacement positioning technology, often choose sliding guide or rolling guide a micro displacement positioning, and because it has a bigger gap and mechanical friction, it is difficult to reach the positioning accuracy nanometer level; Floating guide though can achieve nanoscale positioning accuracy, but the high cost and the large size, use the occasion is very limited. The mechanical transmission declined, and because it has the displacement drive big mechanical clearance, friction

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 119–122.
springerlink.com

and creep phenomenon, its motion accuracy can reach the nanometer level. Flexible hinge is a new type of elastic guide rail, with no form mechanic friction, no clearance, sports high sensitivity, the advantages of simple processing, especially for positioning technologies. Nano Piezoelectric, electricity to stretch in the structure is compact, micro displacement displacement resolution is high, the simple control, no fever, is the ideal micro displacement devices. Flexible hinge for elastic guide rail, piezoelectric, Electricity to drive the displacement in stretch for is to realize the effective method of nano resolution positioning.

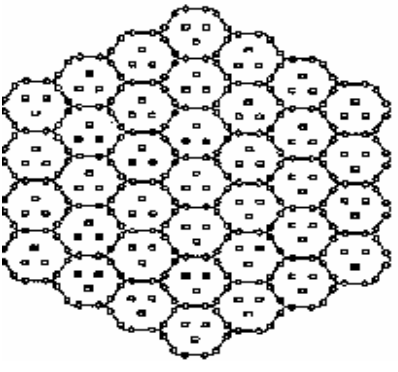


Fig. 1. The sensor and prop device

2 Measurement and Control System

Mirror control computer from displacement of the detectors receiving output signal and actuators, the position of each mirror piece of signal transmitter two-way many paths by computer and Lord provide drivers oscillator control the can in six two-way displacement detector and three actuators transfer between information.

By displacement sensor detection of computer data, according to face form control equation and using least square method to optimize the modifier, and its transmission give actuators to mirror piece of imposed appropriate amount to correct. Mirror piece of mobile, to get a new sensor detection data, again by computer repeat the above circulation, make each lens optimization of optical lens to focus overlap to achieve the best mirror distinguishing ability.

The system used in the displacement detector for capacitor sensor, it has wide measuring range, high sensitivity, easy to use and may to the small displacement of non-contact measurement, and many other advantages. The sensor for differential, detection data ($S = S_1 - S_2$) through computer processing, then lost to actuators to control the relative position between the son mirror. The parallel plate capacitor sensor principle equation:

$$v_o = \frac{C_s V_s}{C_g} \approx \frac{C_s V_s}{C_T} = -\frac{C_s V_s}{\varepsilon A} S_i$$

v_o , The voltage output signal detection; C_s , Standard capacitor; V_s , Standard square wave signal; A sensors to head end face effective area; C_g Total equivalent capacitance sensor; C_T Sensors to head the detection effective capacitance; S_i Sensors to be measured between head and that the distance parameters (i= 1, 2).

From the analysis, capacitor sensor detection error will stray capacitance ($\Delta C = C_g - C_T$), test environment ($\varepsilon = \varepsilon_0 \varepsilon_r$) and the stability of the signal source (V_s) from several aspects such as factors. According to using the environment characteristics and requirements, capacitive displacement detector design from the following two aspects:

- Sensor design

Sensors designed to capacitive differential operation, its characteristic is that sensor: sensor high sensitivity, can lead to the capacitance sensor basic eliminated; Use "drive technology" of cable operations type circuit excellent linearity, high accuracy, range and resolution of larger is better; The change of dielectric constant between the plate will change the displacement signal, but the sensor differential design is effectively overcome the telescope a work environment temperature, humidity and air pressure change larger factors such as the effects of sensors in real-time monitoring, in the high stability.

Sensor signal circuit design into the lead, reduce the influence of capacitance. At the same time, the head of the sensors from the traditional cylindrical into plate coating form, further make distributed capacitance greatly reduced, and to ensure the reliability of the measurement.

- The design for the parameters of the electric displacement detector

The accuracy and stability of the signal source influence the measuring accuracy stability. Using AC-DC-AC-D second high frequency inverter method, the stability reaches 20 x 10.6m. At the same time, in the precision of using LM385 oscillator stability of low temperature coefficient characteristics, used as a diode as voltage amplifier provide stabilized voltage power supply, improve the signal source of the ability to adapt to temperature changes. USES the import new devices MAX297 class 8 low pass elliptic function capacitance switch type filter, improve the ability of the dynamic response of the circuit

- Positioning accuracy

The application of nano technology in positioning, greatly improve the telescope's performance, the main several parameters are at a high precision:

1) The surface of the mirror surface shape control error of just 50.

2) Capacitive displacement detector resolution to 0. 001, measurement range for + 25, linearity is 0.35%, and the dynamic characteristics of up to 50.

3) Capacitive displacement in normal work under way detector of drift lower than 5, every day the effects of the interference to less than 50.

References

1. Jiao, C., Li, S., Xie, X., et al.: Controllability of removal function in the ion beam figuring process for optics mirrors. Optical Technique 34(5), 651–654 (2008)
2. Zhou, L., Dai, Y., Xie, X., et al.: Model and method to determine dwell time in ion beam figuring. Nanotechnology and Precision Engineering 5(2), 107–112 (2007)
3. Pahk, H.J., Lee, D.S., Park, J.H.: Ultra precision positioning system for servo motor-piezo actuator using the dual servo loop and digital filter implementation. International Journal of Machine Tools & Manufacture 41(1), 123–131 (2001)
4. Uchida, N.: A vertical X-Y stage for x- ray lithography using SOR. Bull. Japan Soc. of Prec. Eng. 22(2), 102–108 (1988)

A High P1dB GSM RF Switch in Electronic Engineering*

Shen-Whan Chen[1], Shuming T. Wang[2], Zhao-Wei Han[2], Wei-Nung Lee[2], Chien-Hsun Chen[2], and Sheng-Feng Su[2]

[1] Communication Engineering Department, I-Shou University, Kaohsiung City, Taiwan
[2] Electrical Engineering Department, I-Shou University, Kaohsiung City, Taiwan
{jasonchen,smwang,isu9901001m,isu9901032m}@isu.edu.tw,
{vic,edson}@maxiamp.com

Abstract. In order to meet the high RF power capabilities by the requirements of mobile communication system, two types of high P1dB GaAs PHEMT quad-band RF switch are designed and implemented in this paper. One has adopted three-gate-finger method and the other uses only single gate but with extra gate length. Both switches possess qualities of low insertion loss, high degree of isolation, and low noise. Their output P1dB at the bands of 820 to 920 MHz as well as 1720 to 1920 MHz are measured and compared.

Keywords: RF switch, PHEMT, GaAs, FEM, quad-band GSM switch.

1 Introduction

In current wireless communication system, the capability of the power handling as well as the degree of isolation of a RF switch tends to getting higher and higher. Take handset in GSM (Global System for Mobil communication) system as an example, in order to minimizing noise, increasing RF output power, and increasing the range of transmission, a RF switch with high P1dB and isolation is desired in the front-end parts of that system. A typical frond-end module used in a wireless communication system contains PA, LNA, and RF switch.

Nowadays, GaAs (Gallium Arsenide) field effect transistors (FETs) are preferred for the making of RF switches [1-5] because of its better performance, not only can they meet requirements in terms of power and isolation but also they could achieve in comparably smaller size. In this paper, we have designed and implemented two switches and compared their performance. Both switches have the same circuit topology that employs two SPDT structure (as shown in Figure 3 and 5) for application in the GSM quad band frond-end module. As shown in Figure 3, one switch is implemented with the features of three-finger-gate transistors being adopted for reception switch arm. The other switch, as shown in Figure 5, is also implemented in its suitable way.

* This research was supported by the Southern Taiwan Science Park Administration (STSPA), Taiwan, R.O.C., and under the contract of 100CC01.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 123–128.
springerlink.com

2 Method Description

A SDPT (Single-Pole Double-Throw) switch has two switch arms and typically one is used for signal transmission while the other for reception. In order to raise the power capability of a SPDT switch, an increase of power capability in the reception switch arm is the way to achieving it, since when very large amount of RF power that are being carried by the transmission switch arm it is the reception arm that has to endure the intense stress of electrical filed. In other words, when the transmission arm is in the on state and allowing a large amount of RF power to go through, it is the off-state reception arm that must withstand and without being collapsed. That's the reason why the necessary power capability increasing of a SPDT switch is to increase the power capability of the reception arm.

The principle to increase power capability of a switch arm is to decrease the electric filed intensity [6] that are imposed between drain and source of that switch arm. In this paper, two approaches are used to implement the high P1dB RF switches. The first one adopts the method of multiple fingers as shown in Figure 1 and three-finger-gate transistors being used for the reception switch arm. The length of the depletion zone formed under this case, when the reception arm is operated at off state, is about three times longer than the one of standard approach that uses only one gate finger. Hence, in comparison, the electrical field intensity is lower by also a factor of three and the higher power transmission is then achieved. The other approach, as shown in Figure 2, uses only one finger but with very long gate length to spread out the intensity of the imposing electrical field. The typical gate width of a FET is 0.5μm but in this case we 3.7μm for the FETs that are used in series in the reception arm.

Dual SPDT structures are used to implement a single GSM switch that is used in the GSM quad-band frond-end module. One SPDT is for frequency operation from 820 to 920 MHz and the other is for 1720 to 1920 MHz. Hence, two different MMIC chips for GSM application used in quad band front-end module are designed and implemented.

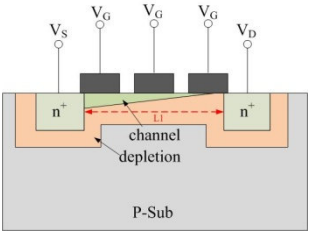


Fig. 1. A three-finger depletion mode PHEMT FET under off condition

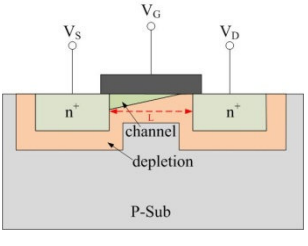


Fig. 2. A single-finger depletion mode PHEMT FET under off condition

3 Experiment

Figure 3 and 4 show the schematic and layout respectively of the three-gate-fingers SPDT switch. The transmission arm is implemented with a single in-series FET structure while the reception arm is with three-fingers and then single-finger L-type structure. The in-series FET in transmission arm uses standard gate length 0.5μm with a gate width of 2000μm (8×250). The two in-series FETs in reception arm use three 0.5μm gate fingers and both with the same gate width of 3000μm (15×200). The separation from drain to source of these three-gate-fingers FETs is 5.7μm. Figure 5 and 6 show the schematic and layout respectively of the long-gate-length SPDT switch. The transmission arm is implemented with a single in-series FET structure while the reception arm is with standard L-type structure. The in-series FET in transmission arm uses standard gate length 0.5μm with a gate width of 2000μm (8×250). The two in-series FETs in reception arm use 3.7μm gate finger and both with the same gate width of 1680μm (8×210). The separation from drain to source of these long-gate-length FETs is as long as it is in those three-gate-fingers FETs. The shunt FET in reception arm uses also standard gate length with a gate width of 600μm (10×60).The shunt FET in reception arm uses also standard gate length with a gate width of 600μm (10×60).

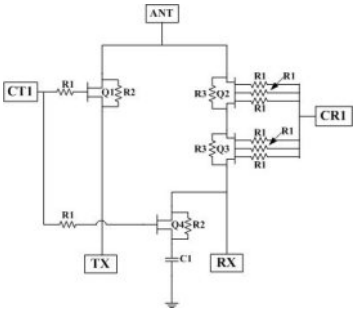


Fig. 3. Schematic of MMIC by three-gate-fingers method

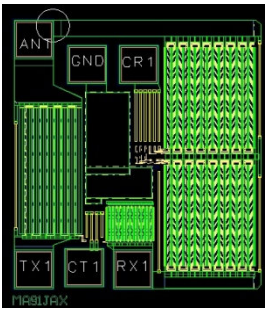


Fig. 4. MMIC by three-gate-fingers method

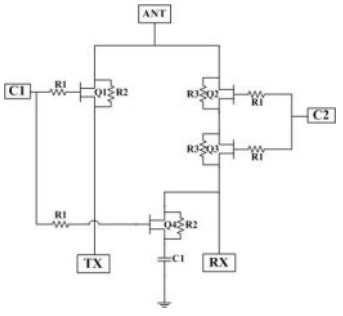


Fig. 5. Schematic of MMIC by long-gate-length method

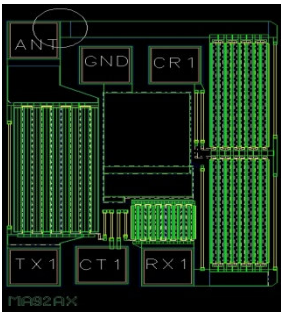


Fig. 6. MMIC by long-gate-length method

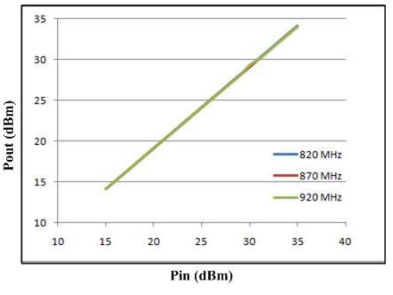


Fig. 7. Measurement results of power out versus power in from 820 to 920MHz for MMIC by three-gate-fingers method

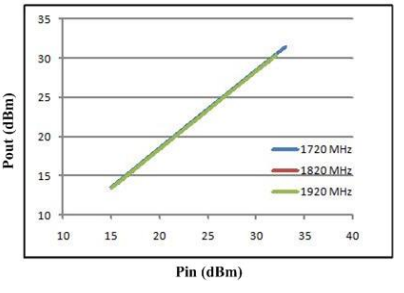


Fig. 8. Measurement results of power out versus power in from 1720 to 1920MHz for MMIC by three-gate-fingers method

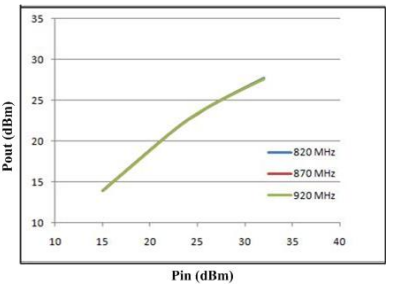


Fig. 9. Measurement results of power out versus power in from 820 to 920MHz for MMIC by long-gate-length method

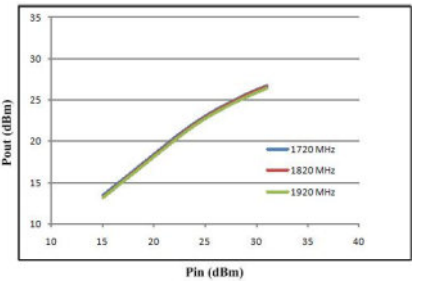


Fig. 10. Measurement results of power out versus power in from 1720 to 1920MHz for MMIC by long-gate-length method

Figure 7 and 8 show the measurement results of power out versus power in from 820 to 920MHz and from 1720 to 1920MHz respectively for MMIC implemented by this method. Table 1 shows the measurement results of insertion loss and isolation. Figure 9 and 10 show the measurement results of power out versus power in from 820 to 920MHz and from 1720 to 1920MHz respectively for MMIC implemented by this method. Table 2 shows the measurement results of insertion loss and isolation.

Table 1. Measurement results by three-gate-fingers method

	820 MHz	870 MHz	920 MHz	1720 MHz	1820 MHz	1920 MHz	Unit
TX Insertion Loss (Small Signal)	-0.26	-0.27	-0.28	-0.68	-0.69	-0.67	dB
RX Insertion Loss (Small Signal)	-0.51	-0.54	-0.56	-0.84	-0.91	-0.94	dB
TX Insertion Loss (30dBm)	-0.57	-0.55	-0.6	-1.1	-1.2	-1.3	dB
TX to RX Isolation	-30.8	-30.3	-29.9	-26.0	-25.8	-25.4	dB
RX to TX Isolation	-20.3	-19.9	-19.5	-13.5	-12.9	-12.5	dB

Table 2. Measurement results by long-gate-length method

	820 MHz	870 MHz	920 MHz	1720 MHz	1820 MHz	1920 MHz	Unit
TX Insertion Loss (Small Signal)	-0.23	-0.25	-0.28	-0.51	-0.57	-0.59	dB
RX Insertion Loss (Small Signal)	-0.48	-0.46	-0.51	-0.78	-0.83	-0.85	dB
TX Insertion Loss (30dBm)	-0.61	-0.65	-0.66	-1	-1.1	-1.2	dB
TX to RX Isolation	-30.4	-29.9	-29.5	-24.9	-24.6	-24.4	dB
RX to TX Isolation	-20.5	-20	-19.7	-14.2	-13.6	-13.2	dB

4 Conclusion

In this paper, two high P1dB GSM quad-band RF switches are implemented. The switch implemented by three-gate-fingers method has better P1dB than the one by long-gate-length method even though they both have the same separation distance from drain to source. Other than that, both have similar performance. However, for the switch by long-gate-length method, instead of three gate resisters being necessary in three-gate-fingers method, it only needs to use one.

Acknowledgments. This research was supported by the Southern Taiwan Science Park Administration (STSPA), Taiwan, R.O.C., and under the contract of 100CC01.

References

1. Hieda, M., Nakahara, K., Miyaguchi, K., Kurusu, H., Iyama, Y., Takagi, T., Urasaki, S.: High-isolation series-shunt FET SPDT switch with a capacitor canceling FET parasitic inductance. IEEE Trans. Microwave Theory Tech. 49, 2453–2458 (2001)
2. Kohama, K., Ohgihara, T., Murakami, Y.: High power DPDT antenna switch MMIC for digital cellular systems. IEEE Journal of Solid-State Circuit 31, 1406–1411 (1996)
3. Gu, Z., Johnson, D., Belletete, S., Fryklund, D.: Low insertion loss and high linearity PHEMT SPDT and SP3T switch ICs for WLAN 802.11 a/b/g applications. In: Radio Frequency Integrated Circuits (RFIC) Symposium, pp. 505–508 (2004)
4. Lee, K.H., Jin, Z., Koo, K.H.: High linearity SPDT switch for dual band wireless LAN applications. In: Microwave Conference Proceedings, APMC, vol. 2, pp. 4–7 (2005)
5. Roberts, M., Albasha, L., Bosch, W., Gotch, D., Mayock, J., Sandhiya, P., Bisby, A.: Highly linear low voltage GaAs pHEMT MMIC switches for multimode wireless handset applications. In: Radio and Wireless Conference, RAWCON 2001, pp. 61–64 (2001)
6. Gu, Z., Johnson, D., Belletete, S., Frykund, D.: Skyworks Solutions, Inc. “A 2.3 V PHEMT power SP3T antenna switch IC for GSM handsets. In: 25th Annual Technical Digest 2003 Gallium Arsenide Integrated Circuit (GaAs IC) Symposium. IEEE (2003)
7. Sanusi, R., Ismail, M.A., Norhapizin, K., Rahim, A., Marzuki, A., Yahya: 15 GHz SPDT Switch Design using 0.15 _m GaAs Technology for Microwave Applications. In: 2008 International Conference on Electronic Design (2008)

Power Control Routing Algorithm for Maximizing Lifetime in Wireless Sensor Networks

YouRong Chen[1], XiLin Hu[2], HaiBo Yang[1], and LingXiao Ge[1]

[1] College of Information Science and Technology, Zhejiang Shuren University, 8 Shuren Street, Hangzhou, Zhejiang, China
[2] Work Group of Science and Technology Information, Shengzhou Electric Power Company, 18 Huancheng North Road, Shenzhou, Zhejiang, China
Jack_chenyr@163.com

Abstract. To maximize network lifetime when node transmission power is fixed in wireless sensor networks, power control routing algorithm for maximizing lifetime (PCRAML) is proposed. The algorithm analyzes the constraint conditions such as node transmission rate constraint, link maximum transmission rate constraint and node energy constraint. Then, network optimization model is established. Subgradient algorithm is used to solve the network optimization model. It can obtain maximum network lifetime and node data transmission rate when node transmission power is determined by genetic algorithm. Genetic algorithm is used to iteratively select, cross and mutate the power populations. Finally node optimal transmission power, maximum network lifetime and optimal routing scheme are obtained. Simulation results show that PCRAML can obtain the optimal scheme of routing and node transmission power, balance network load and maximize network lifetime. Under certain conditions, PCRAML outperforms subgradient algorithms with fixed transmission power.

Keywords: Wireless Sensor Networks, Power Control, Routing, Lifetime.

1 Introduction

Wireless sensor network (WSN) is composed of many energy-constrained nodes which have sensing, computing and wireless communicating with each other. Recently, low-cost WSN has been developed for many current and future envisioned application, such as environment monitoring, battlefield surveillance, health care and home automation and other fields. It gets more and more attention from academia and industry [1]. The network lifetime which needs to be maximized is one of the non-standard metrics in WSN. It is specifically indented for the networks whose nodes are mostly battery-powered and have limited power. Once one node exhausts energy and is disabled, it may affect network data gathering, even may break up the network. Therefore, some researchers define the network lifetime as the time until the first node depletes its battery. It may just as well be defined as the time until the data source can not reach the data sink [2].

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 129–136.
springerlink.com

For wireless sensor networks, various transmission strategies (routing, power control and scheduling) are used to increase the lifetime of such networks. In the strategies, node transmission power control primarily regulates transmission power of each node. If the premise of network connectivity is meted, the power control can reduce the energy consumption, increase single hop throughput, balance the number of single-hop neighbor nodes, remove unnecessary communication links, and come out an optimal forwarding network topology [3].

At present, the research on lifetime maximization and power control has got some accomplishment. To maximize network lifetime, references [4-7] researched on the lifetime maximization problem under different situations. All problems were formulated as linear programming problems assuming that node transmission power can change at any time. Dual decomposition, subgradient algorithm and other optimal algorithms were used to solve it. Scholars also proposed many new power control algorithms such as VRTPC[8], PCAP[9], LMA and LMN[10], CBTC [11] and etc. Reference [12] proposed power control algorithm based on various nearest-neighbor distances algorithm in wireless sensor network. Nodes transmit data with the same optimal transmission power. But it is only fit for the single cluster network. Therefore, this paper develops the research accomplishment about power control and lifetime maximization problem, proposes power control routing algorithm for maximizing lifetime (PCRAML) in wireless sensor networks to obtain the node optimal transmission power and routing scheme, and maximize network lifetime.

2 Network Optimization Model

In the research, assume that:

- Sink node and other nodes are static with fixed location. Sink node can know the topology information of whole network.
- Sensor nodes have the same performance (such as initial energy, energy consumption parameters, maximum communication radius).
- Sensor nodes have the same energy model.
- Sink node gathers data regularly, and sensor nodes transmit data to sink node directly or in multi-hop way.
- Each sensor node's energy is limited and sink node's energy is not limited.

$G(V, L)$ represents wireless sensor networks where V is a non-empty set of network nodes, $|V|$ is the number of nodes and L is a set of wireless link. $G(V, L)$ is undirected connected graph. In wireless sensor networks, if node j lies in communication radius of node i, it can be called neighbor node of node i. Meanwhile, node i is the neighbor node of node j. Link(i,j) represents the link from node i to node j. $N(i)$ is all neighbor nodes set of node i. The network connection matrix A represents the connected relationship of nodes. The elements of matrix A are as follow.

$$a_{ij} = a_{ji} = \begin{cases} 1 & j \in N(i) \\ 0 & j \notin N(i) \end{cases}. \tag{1}$$

S_i is the data sensed rate of node i. F_{ij} is the data transmission rate from node i to node j. $F(i)$ is the data transmission rate of node i. It is composed of sensed rate and received rate.

$$F(i) = \sum_{j \in V} a_{ij} F_{ij} = S_i + \sum_{j \in V} a_{ji} F_{ji}, i \in V. \quad (2)$$

Because network bandwidth is limited, data transmission rate in the link is also limited. So data transmission constraint is as follow.

$$a_{ij} F_{ij} + a_{ji} F_{ji} \le R_{\max} \ \forall i, j \in V. \quad (3)$$

Where, $R_{\max}$ is the maximum data transmission rate of link.

In the data transmission process, Friss free-space model is used to get the relationship of node transmission power and distance. The formula is as follow.

$$P_i = \frac{(4\pi d_{i\max})^2 P_r D_s}{G_t G_r b^2}. \quad (4)$$

Where, P_i is the transmission power. P_r is the received power. G_t is the power magnification of transmission antenna. D_s relates to the loss rate of wireless communication system. G_r is the power magnification of received antenna. $d_{i\max}$ is the maximum communication distance of node i. b is wavelength. Similarly, $j \in N(i)$ can be expressed as $G_t G_r b^2 P_i / (4\pi d_{i\max})^2 L \ge P_{\min}$. P_{min} is the minimum required received power of node i.

In the network, transmission energy consumption E_{Tx} contains electronic energy consumption of transmission circuit $E_{\text{Tx-elec}}$ and energy consumption of signal amplifier $E_{\text{Tx-amp}}$. $E_{\text{Tx-elec}}$ is fixed at $g_1 E_{\text{elec}}$, where g_1 represents the amount of transmission data, and E_{elec} is electronic energy consumption constant. $E_{\text{Tx-amp}}$ relates to transmission power P_i. Receiver only considers the electronic energy consumption $g_2 E_{\text{elec}}$, where g_2 represents the amount of received data. Therefore, the energy consumption rate of received data for node i is $\sum_{j \in V} a_{ji} F_{ji} E_{elec}$. The energy consumption rate of transmission data for node i is $\sum_{j \in V} a_{ij} F_{ij} (E_{elec} + \varepsilon_{\text{ft}} * P_i^{\gamma})$. Where ε_{ft} is amplifier energy consumption constant and γ is the amplification factor. Then total energy consumption of node i must be not larger than its initial energy E_i.

$$\sum_{j \in V} a_{ji} F_{ji} E_{elec} + \sum_{j \in V} a_{ij} F_{ij} (E_{elec} + \varepsilon_{\text{ft}} * P_i^{\gamma}) \le E_i / T_{\text{net}}, i \in V. \quad (5)$$

Where, T_{net} is the network lifetime. When the lifetime of node i is T_i, according to the definition, network lifetime is

$$T_{net} = \min_{i \in V} T_i. \quad (6)$$

The goal is to solve P_i and F_{ij}, maximize network lifetime and obtain the node optimal transmission power and routing scheme, namely

$$\max \ T_{net}(P, F). \quad (7)$$

Led $q = 1/T_{\text{net}}(P,F)$, formula(7) transforms to $\min(q)$. Then the optimization problem can convert into

$$\min(q). \tag{8}$$

s.t.: constrain conditions (2)-(5)

$$q \geq 0\,,\ F_{ij} \geq 0. \tag{9}$$

$$P_i \geq 0. \tag{10}$$

3 PCRAML Algorithm

Genetic algorithm can be used to solve the nonlinear optimization model. But it involves too many variables, has intensive calculation and slow convergence. In the constraint (5), it has $F_{ij}P_i$ nonlinear term, so subgradient algorithm can not be dual decomposed to solve the optimization model. Therefore, genetic algorithm and subgradient algorithm are used to solve the model. Genetic algorithm only determines the node transmission power. Subgradient algorithm solves the network lifetime optimization model with fixed node transmission power.

3.1 Subgradient Algorithm with Fixed Node Transmission Power

Assuming sink node has known the network topology information and node transmission power is determined by genetic algorithm, to solve the network model (10), the main solution algorithm is as follow:

$$\min(q). \tag{11}$$

s.t.: constraint conditions (2)-(5), (9)

Because model (11) is not convex for all variables, q^2 and quadratic regularization term of F_{ij} are introduced. Then the optimization model (12) is a strictly convex function.

$$\min(q^2 + \sum_{i\in V}\sum_{j\in V}\theta a_{ij}F_{ij}^2). \tag{12}$$

s.t.: constraint conditions (2)-(5), (9)

where, θ is regularization factor and $\theta \geq 0$. When θ tends to 0, the value of formula (12) tends to $\min(q)$. It guarantees that the value of q is the smallest. The proof is given in references [6].

Then, the subgradient algorithm is used to solve the optimization model (12)

$$\begin{aligned} L(q, F_{ij}, \lambda, \nu, \mu) = &-\sum_{i\in V}\lambda_i S_i - R_{\max}\sum_{i\in V}\sum_{j\in V}\nu_{ij} + q^2 - q\sum_{i\in V}\mu_i E_i + \\ &\sum_{i\in V}\sum_{j\in V}\theta a_{ij}F_{ij}^2 + a_{ij}F_{ij}(E_{elec}\mu_j + \mu_i(E_{elec} + \varepsilon_{fs}P_t^{\gamma}) + \lambda_i - \lambda_j + \nu_{ij} + \nu_{ji}). \end{aligned} \tag{13}$$

The value of Lagrangian dual function $G(\lambda, \nu, \mu)$ is the minimum value of $L(q, F_{ij}, \lambda, \nu, \mu)$ when q and F_{ij} are known. It can be expressed as

$$G(\lambda, \nu, \mu) = \min_{(q, F_{ij})}\{L(q, F_{ij}, \lambda, \nu, \mu)\}. \tag{14}$$

The optimization model (12) can be converted into Lagrangian dual problem.

$$\begin{aligned} &\text{maximize}_{(\lambda,\nu,\mu)}\ G(\lambda,\nu,\mu). \\ &\text{s.t.}\ \ \nu_{ij} \geq 0\ \forall i, j \in V. \\ &\mu_i \geq 0\ \forall i \in V. \end{aligned} \tag{15}$$

The objective function (15) is concave differentiable function. Subgradient method can be used to solve it. If the k-th iteration step $\theta^{(k)}(k>0)$ satisfies

$$\lim_{k\to\infty} \theta^{(k)} = 0,\ \sum_{k=1}^{\infty} \theta^{(k)} = \infty. \tag{16}$$

As iteration times increases, the solution of (15) converges to optimal solution. The dual variables update in $(k+1)$-th iteration are as follows.

$$\lambda_i^{(k+1)} = \lambda_i^{(k)} - \theta^{(k)}(S_i + \sum_{j\in V} a_{ji}F_{ji} - \sum_{j\in V} a_{ij}F_{ij}), \forall i \in V. \tag{17}$$

$$\nu_{ij}^{(k+1)} = \max\{0, \nu_{ij}^{(k)} - \theta^{(k)}(R_{\max} - a_{ij}F_{ij} - a_{ji}F_{ji})\}, \forall i, j \in V. \tag{18}$$

$$\mu_i^{(k+1)} = \max\left\{0, \mu_i^{(k)} - \theta^{(k)}\left(qE_i - \sum_{j\in V} a_{ji}F_{ji}E_{elec} - \sum_{j\in V} a_{ij}F_{ij}(E_{elec} + \varepsilon_{fi}P_t^{\gamma})\right)\right\}, \forall i \in V. \tag{19}$$

Where, $\theta^{(k)} = w/\sqrt{k}$, w is step select factor and $w>0$. It satisfies the formula (16).

The main variables q and F_{ij} in the k-th iteration are solved as follows.

$$q^{(k+1)} = \arg\min_{(q>0)}\left(q^2 - q\sum_{i\in V}\mu_i E_i\right). \tag{20}$$

$$F_{ij}^{(k+1)} = \arg\min_{\substack{F_{ij}\geq 0, if\ j\in N(i) \\ F_{ij}=0, if\ j\notin N(i)}}\left(\theta a_{ji}F_{ij}^2 + a_{ij}F_{ij}(E_{elec}\mu_j + \mu_i(E_{elec} + \varepsilon_{fi}P_t^{\gamma}) + \lambda_i - \lambda_j + \nu_{ij} + \nu_{ji})\right). \tag{21}$$

3.2 Algorithm Realization

Sink node knows the network topology information and initials the parameters: node initial energy E_i, transmission power populations g_s, node sensed rate S_i, related constants E_{elec}, $R_{\max}$, G_t, G_r, b, Langrange factor λ, ν, μ and so on. Then the algorithm is realized to calculate maximum network lifetime, the node optimal transmission power and rate. Finally sink node broadcasts to inform sensor nodes. Sensor nodes set their transmission power and rate for data routing. The algorithm is described as follows:

Step 1: Once network starts, sink node transmits information query packets to other nodes in flooding way. When nodes receive the packet from sink node, they forward it to neighbor nodes. Meanwhile, nodes transmit their information to sink node along the same path.

Step 2: Sink node gathers the entire network topology information and initializes the network variables and chromosome codes of transmission power populations.

Step 3: Fitness values of populations are calculated. Subgradient algorithm is iteratively executed to solve the fitness function (22). When the error of network

lifetime is less than ε, the algorithm converges to optimal value. Subgradient algorithm stops. At this point the optimal network lifetime is its population fitness.

$$f(x)=q \tag{22}$$

Step 4: Select, cross and mutation operation.

Then the fitness values sort increasing. The top α percent of populations are selected and directly inherited to the offspring of population. The top α percent of populations are crossed. When cross operation is completed, the new populations have mutation with β probability.

Step 5: Return or end.

Judge whether the iteration number is greater than t. If it is less than t, return step 2, else go to step 6.

Step 6: The sink node broadcasts to inform sensor nodes. Sensor nodes set their transmission powers and rates for data routing.

4 Simulation Result and Analysis

The simulation doesn't consider the energy consumption of routing establishment, routing maintenance, routing failure, timeout retransmission, data calculation and some others, but only considers the energy consumption of wireless communication. The numbers of wireless sensor nodes and one sink node are randomly generated.

In the simulation, the needed parameters are as follows: E_{elec}=50nJ/bit, ε_{ft}= 10.8uJ/W, E_i=1000J, γ=2, α=0.5, β=0.05, G_t=G_i=1, D_s=1, S_i=1k/min, R_{max}=50k/min, b=0.328m, θ=1*10^{-4}, w=5*10^{-8}. Simulation square's side length is 50m. Population size M=100. The iteration times K in subgradient algorithm is 2000. The iteration times t in genetic algorithm is 30. The maximum transmission power P_{max} is 0.37W/bit. The minimum received power P_r is 6.3nW/bit.

First the convergence of PCRAM algorithm is analyzed. Fig.1 is network lifetime convergence curve when the number of nodes is 30. As shown in fig.1, the network lifetime is computed at each iteration number. After about 10 iterations, the network lifetime is close to the optimal value. So PCRAM algorithm is convergence.

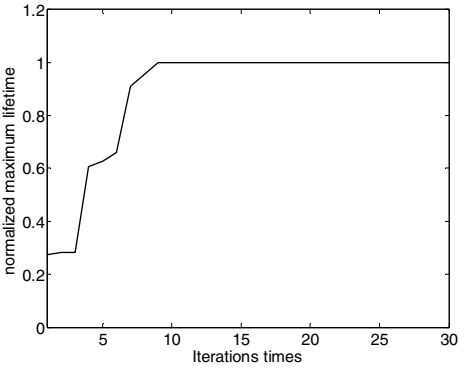


Fig. 1. Network lifetime convergence curve

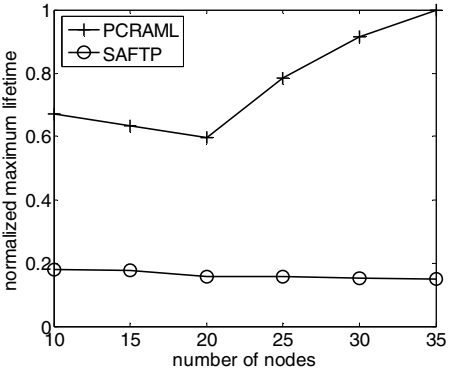


Fig. 2. Network lifetime comparison

To verify the validity of PCRAM algorithm, subgradient algorithm with 3.7mW fixed transmission power (SAFTP) are compared. 10, 15, 20, 25, 30 and 35 numbers of nodes are randomly generated in the simulation area. As shown in fig.2, PCRAML algorithm has very higher network lifetime than SAFTP. It is reason that PCRAML uses subgradient algorithm to obtain the maximum network lifetime and node data transmission rate under determined node transmission power, uses genetic algorithm to iteratively select, cross and mutate transmission power populations, and obtain the optimal transmission power. According to the optimal routing scheme, sensor nodes can balance network load and optimize network lifetime.

5 Conclusion

PCRAML is used to find the node optimal transmission power and routing scheme, and maximize network lifetime. First the network optimization model is established under constraint conditions when node transmission power is fixed. Next, genetic and subgradient algorithms are used to obtain node optimal transmission power, optimal routing scheme and maximum network lifetime. Finally, the convergence of PCRAML is analyzed. The network lifetime of algorithms are compared.

Acknowledgments. This project is supported by a general project, Zhejiang Shuren university of China under grant 2004R002 and a student innovation project, Zhejiang province of China grant 2011R420001.

References

1. Akyildiz, I.F., Su, W.L.: A survey on sensor networks. IEEE Communications Magazine 40, 2–116 (2002)
2. Xu, N., Cassandras, C.G.: A maximum time optimal control approach to routing in sensor networks. In: 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference, Shanghai, P.R. China, pp. 3757–3762 (2009)
3. Gomez, J., Campbell, A.T.: Variable-range transmission power control in wireless Ad hoc networks. IEEE Transactions on Mobile Computing 6, 87–99 (2007)
4. Madan, R., Lall, S.: Distributed algorithms for maximum lifetime routing in wireless sensor network. IEEE Transactions on Wireless Communications 5, 2185–2193 (2006)
5. Gatzianas, M.A., Georgiadis, L.G.: A distributed algorithm for maximum lifetime routing in sensor networks with mobile sink. IEEE Transactions on Wireless Communications 7, 984–994 (2007)
6. He, Y.F., Lee, I., Guan, L.: Distributed algorithms for network lifetime maximization in wireless visual sensor networks. IEEE Transactions on Circuits and System for Video Technology 19, 704–718 (2009)
7. Jadbabaie, A., Ozdaglar, A., Zargham, M.: A distributed Newton method for network optimization. In: 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference, Shanghai, P.R. China, pp. 2736–2741 (2009)
8. Gomez, J., Campbell, A.T.: Variable-range transmission power control in wireless Ad hoc networks. IEEE Transactions on Mobile Computing 6, 87–99 (2007)

9. Wen, K., Guo, W., Huang, G.J.: A power control algorithm based on position of node in the wireless Ad hoc networks. Journal of Electronics and Information Technology 31, 201–205 (2009)
10. Kubisch, M., Karl, H., Wolisz, A., et al.: Distributed algorithms for transmission power control in wireless sensor networks. In: Proceedings of IEEE Wireless Communications and Networking Conference, pp. 132–137. IEEE, New Orleans (2003)
11. Li, L., Halpern, J.Y., Bahl, P., et al.: A cone-based distributed topology control algorithm for wireless multi-hop networks. IEEE/ACM Transactions on Networking 13, 147–159 (2005)
12. Chen, Y.R., Yu, L., Dong, Q.F., Hong, Z.: Power control in wireless sensor network based on nearest-neighbor algorithm. Journal of Zhejiang University (Engineering Science) 44, 1321–1326 (2010)

Analysis of the Influences of Quality Factor on the Resonant Characteristics in Electronic Circuits

Weizhou Hou[1] and Juanling Wang[2]

[1] Physics and Electronics College, Henan University, Kaifeng, Henan 475003, China
hwz204@163.com
[2] Yellow River Conservancy Technical Institute, Kaifeng, Henan 475003, China
kfwjl@163.com

Abstract. By studying the resonant characteristics of electronic circuit, we can obviously see that the quality factor is an extremely important parameter for exploring the voltage ratio of circuits, the frequency-selective property of signal, the efficiency of energy storage and the degree of attenuation of signal, etc. The frequency scanning simulation analysis of resonance circuit by Pspice software further explains that the quality factor affects the frequency-selective property. The result turns out that studies on the quality factor can contribute to a more accurate understanding of resonant characteristics of electronic circuits.

Keywords: Resonance, Quality Factor, frequency-selective, Energy.

1 Introduction

Mankind wouldn't have achieved such amazing feats in radio technology without the development of resonant circuits. It will never go too far to say that resonant circuits are the essential parts in the field of radio technology, while the Quality factor Q is an indispensable parameter of resonant circuits[1]. Therefore, a deep understanding of the Quality Factor Q turns out to be crucial in the basic curriculum of electronic circuits. Based on the example of series resonant circuits, this paper primarily focuses on the concrete meaning of Q in four aspects, which hopefully will be helpful to readers with the understanding of Q and resonance.

2 Characterization of Reactance Component Ends Voltage and Port Voltage Ratio Indicator

See the series circuit showed in Fig1, when it resonates, the resonance frequency is: $\omega_0 = \frac{1}{\sqrt{LC}}$.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 137–142.
springerlink.com

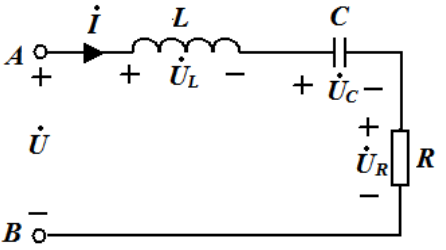


Fig. 1. *RLC* series circuits

Which $\dot{U}$ represents the phase of ports voltage, $\dot{I}$ represents the phase of ports electric current, $\dot{U}_C$ and $\dot{U}_L$ represent phase of the two ports' voltage of capacitors and inductors of reactance element respectively. When the circuit showed in figure 1 resonates, we get formula (1):

$$\dot{U}_C + \dot{U}_L = 0 \tag{1}$$

The voltage ratio of reactance components and port is expressed in Formula (2):

$$U_C / U = U_L / U = 1/(\omega_0 RC) = \omega_0 L / R = Q \tag{2}$$

i.e. when the circuit resonates, the voltage of capacitor or inductance is Q times of the total voltage, which is the first meaning of Q. (Note: U, U_C and U_L are the root mean square-RMS-values of components; ω_0 is the resonant angular frequency. It is same as follows.) For example, in a series resonant circuit, suppose Q=100 and the total alternating voltage is 6V in RMS, the voltage of capacitor or inductance can reach 600V when the circuit resonates. Therefore, the components are likely to be burnt out if the first meaning of Q above is ignored in an experiment.

3 Characterization the Frequency-Selective Property of Signal

One of the most significant applications of resonant circuits in radio technology is the selection of signal. When the resonant frequency is identical to some transmitting frequency, the signal received will reach its peak. When the two frequencies above are greatly different, the signal received will be dim[2-4]. To avoid the "crosstalk" between signals with close and relevant transmitting frequencies, we must need a sharp resonant peak, i.e., a higher value of Q. According to Figure 1, we obtain Formula (3):

$$Z(j\omega) = R + j(\omega L - 1/(\omega C)) = R[1 + jQ(\omega/\omega_0 - \omega_0/\omega)] \tag{3}$$

Further, we obtain Formula (4) and (5):

$$U_R(\omega) = \frac{U}{\sqrt{1 + Q^2(\omega/\omega_0 - \omega_0/\omega)^2}}. \tag{4}$$

$$\frac{U_R(\omega)}{U} = \frac{1}{\sqrt{1 + Q^2(\omega/\omega_0 - \omega_0/\omega)^2}} \cdot \tag{5}$$

In Electronic Circuit Engineering, the difference of ω2 and ω1 corresponding to $U_R(\omega)/U = 1/\sqrt{2} \approx 0.707$ are generally defined as pass band $BW_{0.7}$ of resonant circuits. As the value of frequency is generally positive, the frequency ω in this paper is set positive. By utilizing this condition, we have Formula (6):

$$\frac{1}{\sqrt{1 + Q^2(\omega/\omega_0 - \omega_0/\omega)^2}} = 1/\sqrt{2} \approx 0.707 \cdot \tag{6}$$

We can deduce Formula (6) and obtain Formula (7):

$$\omega^2 \mp \frac{\omega_0}{Q}\omega - \omega_0^2 = 0 \cdot \tag{7}$$

Solve the equation above and we get two positive roots as shown in Formula (8):

$$\omega_1 = \frac{-\frac{\omega_0}{Q} + \omega_0\sqrt{\frac{1}{Q^2} + 4}}{2} \quad \omega_2 = \frac{\frac{\omega_0}{Q} + \omega_0\sqrt{\frac{1}{Q^2} + 4}}{2} \cdot \tag{8}$$

Therefore, pass band $BW_{0.7}$ is as follows in Formula (9):

$$BW_{0.7} = |\omega_2 - \omega_1| = \omega_0/Q \cdot \tag{9}$$

i.e., the pass band $BW_{0.7}$ of resonant circuits has inverse proportion with Q. The higher the Q value is, the better signal selection property we obtain. To help our readers deepen the understanding of the second meaning of Q, we analyze the frequency scanning of the circuit in Figure 1 by Pspice software. The results are shown in Fig2. The value of R is 500Ω, 100Ω and 50Ω, and the value of Q is 0.2,1 and 2, corresponding to Curve 1,2 and 3. Here, L=10mH and C=1μF remain constant.

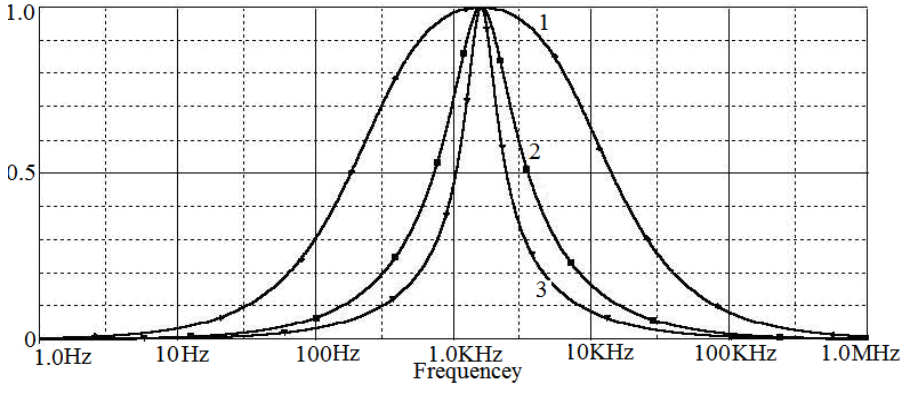


Fig. 2. Curves with different Quality Factor

4 Index Represented 2π Times of the Ratio between the Stored Energy in the Resonance Circuit and the Consumption Energy in Every Cycle

According to the first and second meaning of Q above, we have Formula (10):

$$Q = 2\pi W_S / W_R . \tag{10}$$

Here, W_R represents the energy released from resistor components. W_S represents the energy stored in reactance components. In *RLC* circuits, resistor consumes energy and transforms electromagnetic energy into heat energy. Capacitor and inductance store, instead of consuming, energy and then release it in the process of transforming electricity and magnetic energy. In a period of alternating current, the energy consumed by resistor components is described in Formula (11):

$$W_R = I^2 RT . \tag{11}$$

Here, I is the RMS value of current. The stored energy in capacitor and inductor in the series resonance circuit as forlum (12):

$$W_S = \frac{1}{2} Li^2(t) + \frac{1}{2} Cu_c^2(t) . \tag{12}$$

For $i(t) = I_0 \cos \omega t$, then $u_c(t) = \frac{I_0}{\omega C} \cos\left(\omega t - \frac{\pi}{2}\right) = \frac{I_0}{\omega C} \sin(\omega t)$, we get W_S as forluma (13):

$$W_S = \frac{1}{2} I_0^2 \left[L \cos^2(\omega t) + \frac{1}{\omega^2 C} \sin^2(\omega t) \right] . \tag{13}$$

It implies that in non-resonance condition, WS is a periodically varying quantity, which shows that the resonance circuit exchange reactive power with environment. But in the resonance condition, we have Formula (14):

$$W_S = \frac{1}{2} LI_0^2 \left[\cos^2(\omega t) + \sin^2(\omega t)\right] = \frac{1}{2} LI_0^2 = LI^2 = I^2/(\omega_0^2 C) . \tag{14}$$

Here, W_S is a constant. The resonance circuit doesn't exchange reactive power with environment any more. So W_S in (14) is the stored electromagnetic energy in the circuit in the stable resonance condition. Those energy is imported by the external circuit during the instantaneously process when the resonance circuit is opening. When the circuit is in an stable resonance condition, to keep the resonance, external circuit should import active power continuously to compensate the loss of WR . But in the resonance condition, external circuit need not import reactive power. It can be seen that the ratio between W_S and W_R represents the efficiency of a resonance circuit. So the Quality factor of an resonance circuit can also be defined as Formula (15):

$$Q = 2\pi W_S / W_R . \tag{15}$$

Which shows that the larger of Q, the less dissipation energy relative to the stored energy, namely the higher efficiency of the resonance circuit. This is the third meaning of Q [5-7].

5 Use It to Characterize the Decay Speed of Amplitude

As shown circuit in fig.3, it is known the swithcher K has be at the place labled as Fig.1 since the infinity long time ago till the moment t=0. At the moment t=0 swithcher is switched to the place labled as fig.2. With the condition of $R < 2\sqrt{L/C}$, the *RLC* circuit will under the undamping oscillation situation.

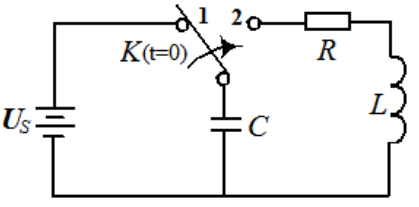


Fig. 3. Under damped oscillation in *RLC* circuit

$e^{-\frac{R}{2L}t}$ is the decaying pattern of the current amplitude, where the time constant is $\tau = 2L/R$, and τ represents the time needed for the amplitude decreasing to e^{-1} of the original amplitude. Because of formula (16).

$$Q = \omega_0 L / R . \tag{16}$$

we can know this value can be expressed by Q. So τ as forluma (17):

$$\tau = 2Q / \omega_0 = QT / \pi . \tag{17}$$

In above equations, T is the period of oscillation, $T = 2\pi / \omega_0$. It is shown: τ equal to Q/π times of the period T, the larger the Q value, the slower the amplitude attenuation. This is the Q's fourth significance. Above priciple can be used in the rough measurement of Q value by showing the damped oscillation curves of *RLC* circuits on the display screen of oscilloscopy(digital ones better).Q value can be obtained by the amplitude ratio of each damped oscillation simply.

References

1. Qiu, G.-Y.: Electric Circuit IV, pp. 210–216. Higher Education Press, Beijing (2006)
2. Zhao, Y., Wang, L.: Design of X-band SIW band pass filter. Electrical Technology, 36–39 (2009)
3. Yu, W.-Y.: The teaching discusses of quality factor of resonant circuit. Journal of Nanchang Vocational & Technical Techers' College, 129–130 (2003)

4. Zheng, F.-J.: Discussion of the Q Factor Concepts in Resonance system. Modern Electronics Technique, 97–99 (2006)
5. Zhao, C., Shi, Y., et al.: Optimal Design Strategy for Three Components Series LLC Resonant Converter. Transactions of China Electrotechnical Society, 65–71 (2008)
6. Han, X.-N.: Caculation and Quality Factor of Resonant Circuits. Shanxi Electric Power, 65–67 (2008)
7. Han, B., Li, J., et al.: Improve Performance of LC Oscillator with Negative-resistor. Journal of Hebei University (Natural Science Edition), 15–17 (2008)

Multi-target Tracking and Communication Based on Digital Beamforming

Dianfei Pan and Naiping Cheng

Department of Optical and Electrical Equipment
Academy of Equipment
Beijing, China
pandianfei_1984@163.com

Abstract. A special monopulse tracking is proposed for multi-tracking, the array of which is split into two azimuth subarrays. By this algorithm can speed up tracking and prevent overshoot and oscillations. In order to actualize multi-target tracking and communication with less DBF (digital beamforming) modules, the space-time multiple accesses method is used in the system. The basic idea is that the tracking processor accomplishes angle acquisition and then gains the accurate azimuth based on monopulse tracking. For the initial data of each module is updated timely and circularly, the tracking time spend to converges the target is decreased significantly. The simulation results show that it can effectively achieve the expectations azimuth angle of each target and the angular error is reduced markedly.

Keywords: multi-target, digital beamforming, antenna array, adaptive.

1 Introduction

With the development of aviation and spaceflight, more and more aircrafts have begun to be seen in space, near space and air, such as satellites, unmanned aerial vehicles (UAVs) and Missiles. In recent years, multiple UAVs or micro satellites constellation are employed to carry out a mission. By this way the system performance can be improved significantly avoiding the system breakdown because of an aircraft fault. The emergence of a constellation of targets has imposed new challenges to objectives detection, tracking and communication. As each target can not be severed by a ground station, if so it will be a huger cost. Therefore the ground station will be analyzed and designed which can carry out search, track and communication for multi-target simultaneity. Considering the advantage of DBF on regional direction and high sensitivity, the adaptive DBF is a wise choice for the system.

Array antenna systems can track and follow more than one targets at a time based on DBF (digital beamforming). Multiple, simultaneous, independent beams can be achieved by the means of DBF modules; either module scans and tracks one target. This would also increase the dwell time which is the time on each target. There are two parts to this paper. The first illustrates the monopulse digital beamforming (DBF) tracking. The array is split into two azimuth subarrays with coordinate symmetry. By

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 143–149.
springerlink.com

this way not only the sum and difference beams can be obtained simultaneously, but also avoiding the difference-to-sum voltage ratio changes with different scan angle of the antenna, as in [1]. Due to the common methods used for multi-target tracking have a good performance only when the targets are no more than DBF modules[2-4], In the second portion, a type of adaptive DBF has been investigated, which was based on space-time multiple accesses. Compared with the former method, the technical presented can tracking and communication with multi-target efficiently, no matter the DBF modules is more or less than targets. The applicability of this technique in the receiver system was validated by computer simulations done in MATLAB. It was seen that each UAV of formation flying UAVs can be tracking and communication successfully even in the condition of more UAVs than DBF modules.

2 Monopulse Tracking and Digital Beamforming

2.1 Monopulse Tracking

Monopulse tracking is different from sequential lobing and conical scanning. The later two methods require the processing of multiple measurements over a finite observation time. They utilize multiple signal samples to determine the error signal to effectively steer the antenna beam to the direction of the target. For monopulse tracking, only a single sample is used to track the target. It obtains the location of a target by comparing signals received in two or more simultaneous beams. This form of tracking is considered to have a higher accuracy compared to time-shared tracking systems such as sequential lobing and conical scan. There are two types of monopulse. One is amplitude comparison monopulse and the other is phase comparison monpulse. Amplitude comparison monopulse is the most common, so the monopulse tracking mentioned here refers to amplitude comparison monopulse.

For amplitude comparison monopulse two beams, pointing in slightly different directions, are created simultaneously. This system utilizes the sum and difference of these patterns to track the target. Two beams are overlapped and their main beams are pointed at small positive and negative angles $\pm\theta q$, which are defined as squint angles or offset angles. The ratio of difference pattern to sum pattern is used to generate the angle-error signal so that the null in the center of the voltage ratio can be placed on the target. The voltage ratio is represented as difference-to-sum voltage ratio. The plot of a typical difference-to-sum voltage ratio is given in Fig.1 [5]. In the linear region as shown, the ratio can be approximated as $\Delta/\Sigma \approx K\theta$. Where K is named the monopulse slope constant, which can be determined from measurement or simulation.

From Fig.1 it's obvious that the common tracking technical will take a longer time to converge to the angle of arrival when the angle is out of the linear region, as the slope becomes smaller changes sign. And the slope of monopulse ratio (MR) changes with different scan angle, so the performance of this technical is instable used in moving target. In order to enhance the ability of multi-target tracking, a special sum and difference beamforming of monopulse is proposed.

Consider an antenna array aligned along the x -axis as shown in Fig.2. As the plane wave arrives at each of the array elements, information on the amplitude and phase are captured by the elements. The sum and difference beamforming process is

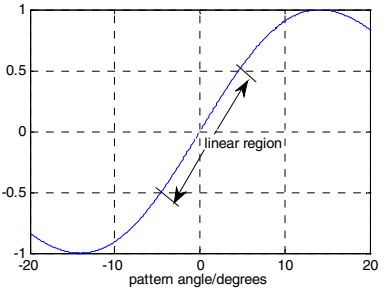


Fig. 1. The normalized $\frac{\Delta}{\Sigma}$ vs.pattern angle

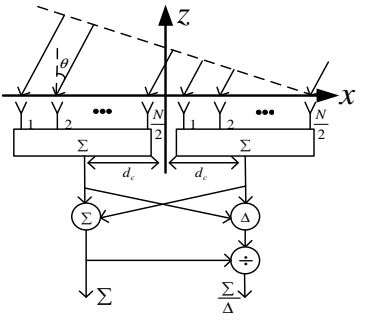


Fig. 2. Sum and difference beamforming

depicted in Fig.2. The array is split into two azimuth subarrays with N /2 elements each. The element is centered at the origin with inter-element spacing d and the locations are symmetrical with respect to the origin of x-axis. Therefore, the location of element n can be expressed as d_n= (2n-(N-1)) d/2. The distance of the centers of the subarrays from the array center is d_c= Nd /2. So the subarray factor for an azimuth subarray is obtained as follows.

$$SF = \sum_{n=1}^{N/2} e^{jkd_n(\sin\theta-\sin\theta_s)} = e^{jkd(\sin\theta-\sin\theta_s)\frac{1-N}{2}} + e^{jkd(\sin\theta-\sin\theta_s)\frac{3-N}{2}} + \ldots + e^{jkd(\sin\theta-\sin\theta_s)\frac{-1}{2}} \quad (1)$$

Then the both sides of equality sign are multiplied by $e^{jkd(\sin\theta-\sin\theta_s)/2}$.

$$e^{jkd(\sin\theta-\sin\theta_s)/2} SF = e^{jkd(\sin\theta-\sin\theta_s)\frac{3-N}{2}} + \ldots + e^{jkd(\sin\theta-\sin\theta_s)\frac{1}{2}} \quad (2)$$

Equation (1) subtracted from (2) gives SF expressed as.

$$SF = \frac{e^{jkd(\sin\theta-\sin\theta_s)\frac{1}{2}} - e^{jkd(\sin\theta-\sin\theta_s)\frac{1-N}{2}}}{e^{jkd(\sin\theta-\sin\theta_s)/2}-1} = \frac{\sin[Nkd(\sin\theta-\sin\theta_s)/4]}{\sin[kd(\sin\theta-\sin\theta_s)/2]} \quad (3)$$

Where θ_s is the direction of arrival (DOA). The subarrays are used to generate the sum and difference beams as following formulas.

$$\sum = SF \cdot e^{-jkd_c(\sin\theta-\sin\theta_s)} + SF \cdot e^{jkd_c(\sin\theta-\sin\theta_s)} = SF \cdot 2\cos[kd_c(\sin\theta-\sin\theta_s)] \quad (4)$$

$$\Delta = SF \cdot e^{-jkd_c(\sin\theta-\sin\theta_s)} - SF \cdot e^{jkd_c(\sin\theta-\sin\theta_s)} = SF \cdot 2j\sin[kd_c(\sin\theta-\sin\theta_s)] \quad (5)$$

The difference over sum ratio is derived as Δ/Σ=jtan(kd_c(sinθ-sinθ$_s$)). Assuming θ-θ$_s$ is a small angle, it's known that sinθ-sinθ$_s$ ≈θ-θ$_s$ and tan(kd_c(θ-θ$_s$))≈kd_c(θ-θ$_s$). So the difference-to-sum ratio can be simplified as $\Delta/\Sigma \approx kd_c(\theta-\theta_s)=K(\theta-\theta_s)$. Where K is the monopulse slope constant can be estimated by K≈kd_c. It can clearly expect that the monopulse slope is not change at different scan angle using this technique.

Suppose that for the tracking of a target (such as a UAV), the maximum azimuth angle is approximately 40° with the antenna tracking the UAV approximately on the horizon. The number of elements was decided to be 10. To prevent grating lobes, the

spacing between the elements is λ/2. In order to send any data collected video, image or audio, the frequency is set at 2.4GHz. So the element spacing is 0.065m. Fig.3 shows the output of the sum and difference beams for scan angle 0° based on the method above. And Fig.4 gives the difference-to-sum ratios at different direction of arrivals.

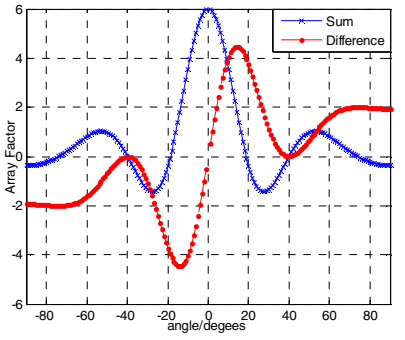


Fig. 3. Monopulse ratio curves at different azimuths

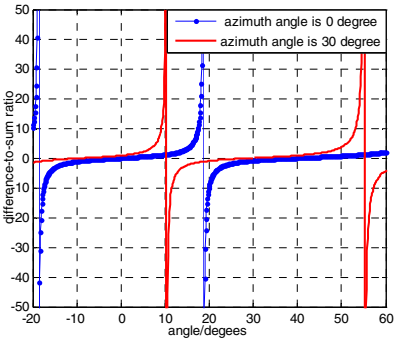


Fig. 4. Sum and difference beams at θs=0°

In Fig.3, it can be seen that at the designed angles of 0° and 30°, the pattern of the sum beam has a maximum while the difference beam has a value of zero. Fig.4 gives the difference –to-sum ratio against θ at different scan angles. It is observed from the Fig.4 that the linear region has a span of approximately 20° (±10° from the azimuth angle). The linear region of each curve near the scan angle is used to compute the correction angle based on (11). If the initial angle of tracking is not at the linear region, the slope of MR is much larger. So more angle distributions could be obtained, which could a less time to converge to the true direction of the signal. In addition this monopluse algorithm can prevent overshoot and oscillations.

2.2 Adaptive Digital Beamforming

Both the sum and the difference beams are used to position the direction of objective but not used for transmission. In [7] the main beam of is used as communication beam, but the performance of anti-noise is not satisfied. The adoption of adaptive algorithm overcomes the defects of beamforming, which makes the system get the adaptive capacity to variable working conditions. The digital phased array can point its main beam electronically to a desired azimuth without actually moving or rotating an antenna. As the array directs its main beam to a target, the other targets can be considered as interferences in multi-target tracking and communication system, Based the adaptive algorithm, steering the main beam to the signal source can reduce interference and improve signal-to-noise ratio (SNR).

Least Mean Square (LMS) has been widely used in practice for the advantages of less computation and easier to implement. Here LMS is employed to decide complex weight vectors, which is multiplied with the signal received by antenna elements. Hence the array forms a transmit beam in the desired direction and minimizes the

output in the interfere direction. The signal received by antenna elements can be expressed as.

$$x(n) = \alpha(\theta_d)s_d(n) + \sum_{k=1}^{M} \alpha(\theta_i)s_i(n) + N(n) \tag{6}$$

Where $\alpha(\theta_d)$ and $\alpha(\theta_i)$ are the steering vectors of desired signal and interferences and M is the number of inter interferences. N is the AGWN (Additive Gaussian White Noise). Multiplied by complex weight vectors, the output of array is expressed as.

$$y(n) = W^T s_d(n) + \sum_{k=1}^{N-1} W^T s_i(n) + W^T N(n) \tag{7}$$

Where $W= [w_1\ w_2 \ldots w_N]^T$. The LSM algorithm avoids matrix inversion operation by using instantaneous gradient vector ∇_J to update the weight vector. The gradient vector can be obtained by the autocorrelation matrix and cross correlation matrix. From LSM algorithm we can get

$$y(n) = W^H(n)x(n); e(n) = d(n) - y(n); W(n+1) = W(n) + \mu x(n)e^*(n) \tag{8}$$

Where μ is a small positive constant, called the step size whose value is between 0 and 1. Employed this algorithm to form beams, the weight vectors can be update adaptively according to the different condition.

Assuming the number of the targets is 6, whose angles of arrival are [0°, 20°,40°, 60°, -20°, -45°] respectively. The other parameters are the same as those in Section A. The digital and synchronous 6 receive modules maintain the data link between the each target and the tracking system. For each target, the others are interferences. The adaptive algorithm can form narrow null in the directions of these interferences. Based on the algorithm proposed, the beam formed for the target of 20° is presented as follow.

It's obvious that both beam patterns have their main beam pointed to the desired direction and the LMS algorithm generates nulls toward the interferences. The ratio of main lobe to the first side is about 12dB and 20dB for sum and LSM respectively. So

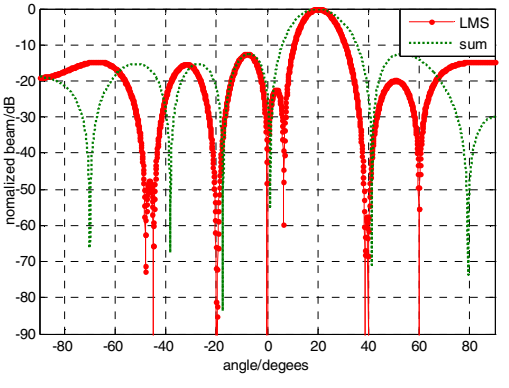


Fig. 5. Beam pattern Comparision of the LSM

the adaptive DBF has better performance of anti-jamming in multi-target system. However, before LMS algorithm converges to its optimum values, the directions of interferences and desired signal will change. The iterative speed should be much faster than the objectives change.

3 Multi-target Tracking and Communication System

As mentioned above each DBF module maintains a main beam point to a target. In this condition the number of objectives is known and no more than DBF modules. However, in practice the desired target is unpredictable and moving randomly for the receiver. In this part a method based on space-time multiple accesses will be proposed. The system structure is shown in Fig.6.

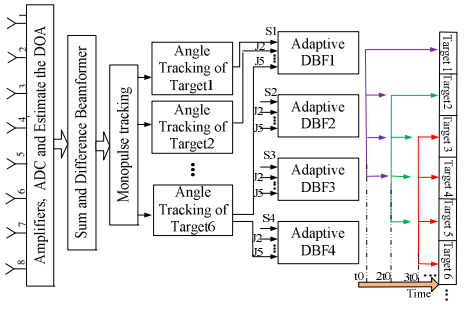


Fig. 6. System structure with 4 DBF and sum modules for 6 objectives

The basic idea is that firstly the tracking processor accomplishes angle acquisition and then gains the accurate azimuth based on monopulse tracking. In this paper the desired objectives are more than adaptive DBF modules. So at initial time t_0 the four modules are employed for objectives from target1 to target4 respectively. In a processing time t_0, the DBF module can finish the desired target tracking in the worst situation (with the largest angle error 10°). In next time $2t_0$ DBF module1 works for target2. At the same time the tracking data of DBF module2 in t_0 was provide to module1. The rest can be deduced by analogy. This process is repeated ceaselessly for the purpose that to sieves more objectives with a fixed number of adaptive DBF modules. For the initial data of each module is updated timely and circularly, the tracking time spend to converges the target is reduced significantly. Take target1 as an example, in a cyclic period T the time t spend to converge its angle can be expressed as t=NT/M. Where N=4 is the number of DBF modules, M=6 is the number of targets and $T=6t_0$. Compared with the normal way the performance will be improved by N/M times. The algorithm can perform better with more targets than DBF modules.

Based on the principle, the following simulation results are given to find the speed and accuracy of our system. Suppose the azimuth angle of target1 changes from 0° to 20°uniformly in 10 seconds. The antenna parameters are the same as above. By simulation the speed and accuracy of tracking are presented as follows.

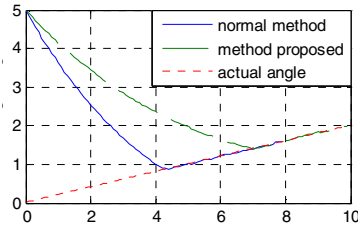


Fig. 7. Comparison of tracking speed

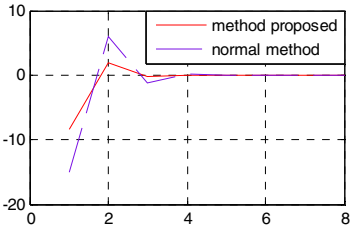


Fig. 8. Comparison of tracking error

As shown in Fig.7 and Fig.8, in the same situation, the space-time multiple accesses method can achieve faster convergence than time multiple accesses. Also the angular error of this technology is much smaller. As can be observed from Fig.8, the initial azimuth angle is more close to actual value for the data is updated timely and circularly. On the contrary, normal method will spend much time to converge the target and possible loss of track with the increase of targets.

4 Conclusions

In this paper, a novel and less complex adaptive beamforming method was introduced for multi-target tracking and communication. In this method, a Special monopulse tracking is presented to avoided overshoot and oscillations. And the two subarrays are communal for each target. Due to the number of target is unknown and increase DBF modules is not wise, the space-time multiple accesses method is employed for the condition of more targets than DBF modules. Simulation results show that the multi-target tracking system provides remarkable improvements, in terms of interference suppression, convergence rate and angular error over that of normal ways.

References

1. Jiangmei, L., Shiyong, C., Man, C.: Study on influence of inter-component-beam angle on sum-difference monopulse sysytem. Journal of Electronic Measurement and Instrumeant 21(5), 26–30 (2007) (in Chinese)
2. Shizhong, Y., Lisheng, Y., Xiaoheng, T., Xiaoqin, D.: Study on the TT & C for the multi-vehicles. Journal of Astronautics 23(6), 12–19 (2002)
3. Gebauer, G., Gockler, T.: Channel-Individual Adaptive Beamforming for Mobile Satellite Communications. IEEE Journal on Selected Areas in Communications 13(2), 439–448 (1995)
4. Yusuf, A., Berke, C., Kevin, M., Gabriel, M.: A Two-Channel 8–20-GHz SiGe BiCMOS Receiver with Selectable IFs for Multibeam Phased-Array Digital Beamforming Applications. IEEE Transactions on Microwave Theory and Techniques 59(3), 716–726 (2011)
5. Levent, G., Robert, B., David, J., Gert, B.: Digital tracking array using off-the shelf hardware. IEEE Antennas and Propagation Magazione 50(1), 107–114 (2008)
6. Cross, F.: Smart antennas for wireless communications. McGraw-Hill Education, Asia (2009)
7. Mohanman, Zainol: MI-NLMS adaptive beamforming algorithm for smart antenna system applications. Journal of Zhejiang University Science 7(10), 1709–1716 (2006)

Study of Linearization Technology of LTE RF Power Amplifier Based on Pre-distortion

Jun Chen

Institute of Physics and Information Engineering of Fuzhou University, Fuzhou 350002, China

Abstract. This paper intends to study the nonlinearity of LTE RF power amplifier and compensate for nonlinearity and memory effect of the amplifier, and compare the digital pre-distortion lookup table-based Digital pre-distortion and memory polynomial pre-compensation effect of the device. It can draw a conclusion that Digital Pre-Distortion can improve ACLR well, the memory effect of the amplifier in the single-carrier FDD-LTE signals can also be compensated at a certain degree.

Keywords: LTE, RF power amplifiers, Pre-distortion.

1 Introduction

With the vigorous development of the third generation mobile communication as well as the growing demand for mobile network rates, LTE has become research and commercial hotspots of B3G and super 3G. In order to improve the spectral efficiency, modulation mode of high PAPR are usually adopted in LTE, such as 16QAM, 64QAM, OFDM, MIMO and so on. It will set a higher demand on the linearity of final stage of RF power amplifier so as to avoid adjacent channel interference to ensure the modulation narrowband characteristics. Now experts at home and abroad has presented many solutions for linearity of RF power amplifier, such as Feed forward technology, feedback technology, predistortion technology and so on. Digital predistortion technology is a very promising and effective method in various kinds of linearity technology. And digital predistortion technology has a very good linear verified results in 3G communication system. The FDD-LTE signal based on 20MHz single-band will be used to compensate the non-linearity of LTE RF power amplifier through digital predistortion technology. At the same time, the non-linearity pre-compensation capacity of the two predistorter between LUT and memory polynomial for LTE RF amplifier will be compared.

2 Predistorter

Now many digital pre-distortion model are studied at home and abroad, such as LUT, memory polynomial, the Hammerstein, improved Hammerstein model and so on. Considering their linearity capacity and memory compensation capacity, the

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 151–154.
springerlink.com

predistorter based on LUT and memory polynomial model are designed. The realization diagram are shown in fig.1 and 2.

In fig.1, the pre-compensation Ic and Qc for RF power amplifier are found by digital baseband input signal u(n) according to the address index., and then the value is sent to the complex multiplier synchronized with the input signal u(n). The results of complex multiplier is pre-distored with no memory effect.

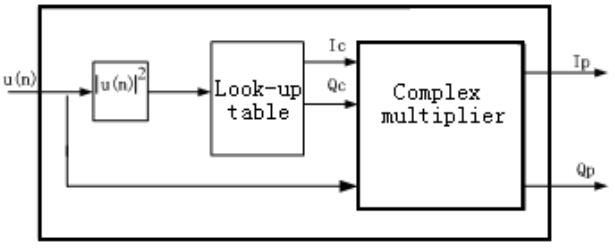


Fig. 1. Digital predistorter based on LUT

In fig.2, baseband signal u(n) is divided into two. One gets into every polynomial units after delay, the other is added by the output of every polynomial units. At last, the combined result z(n) is pre-distored with memory effect.

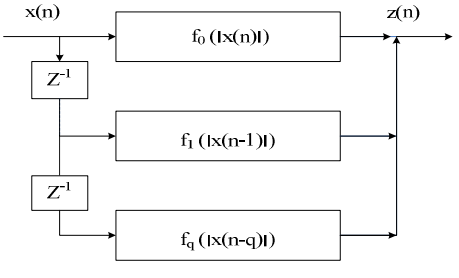


Fig. 2. Digital predistorter based on memory polynomial

3 Test Platform

In order to verify the performance of pre-distortion algorithm quickly, the test platform is shown in fig3. The platform includes Agilent's vector signal generator E4438C, spectrum analyzer E4445A, vector signal analysis system VSA9600, power meter, attenuator and computers. First, the digital baseband IQ signal of single-band FDD-LTE are downloaded into the E4438C to generate RF signal in txt text format. Second, the RF signal are sent to the power amplifier. Third, a small part of signal is coupled from the output of amplifier, which are sent to the attenuator and then sent to the E445A to do the power sepctrum analysis. Last, the output data are captured by VSA9600. The principle is to use the analog to digital converter module for data

acquisition by E4445A firstly, and then the data are sent to the computer to analyze and process through GPIB or LAN interface. Finally, the captured data are saved as txt format. The data file is digital predistored by DSP, then the processed results are downloaded again into E4438C to generate RF signal to test system. The effect of pre-distortion can be seen after several cycles.

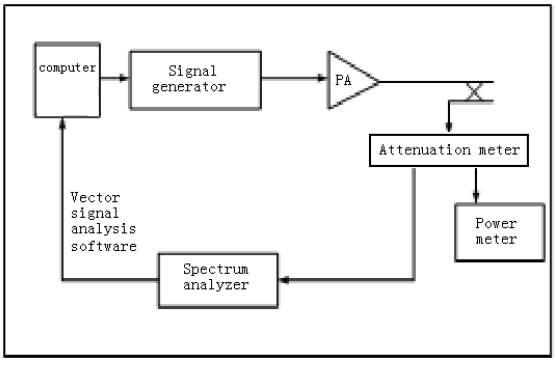


Fig. 3. DPD test platform

4 Test Results

The single-band FDD-LTE of 20MHz are used as test signal in the article. The test amplifier includes two levels. Freescale MHVIC2115NR2 are used in previous amplifier whose drain voltage is 27V, gate voltage is 6.7V, gain is 27dB and the work frequency range is from 2110 to 2170MHz. Freescale MRF6S21050are used in post amplifier whose drain voltage is 27V, gate voltage is 3V and gain is 14dB. The input power is 2dBm and the output power is 39.5dBm in the test. The test results are shown in fig4. Line a is the spectrum without pre-distortion. Line b is the spectrum after pre-distortion based on LUT. Line c is the spectrum after pre-distortion based on memory polynomial with 3-order and 5-tap. From figure 4, the intermodulation distortion is significantly reduced after pre-distorter based on LUT or memory

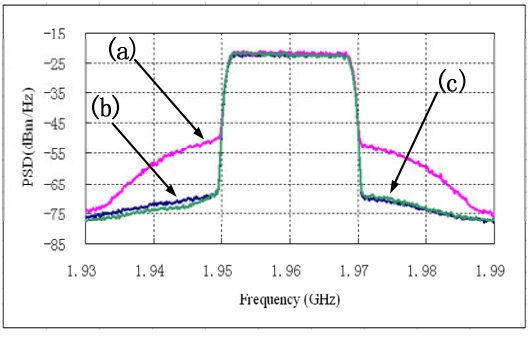


Fig. 4. Test results comparison between LUT DPD and memory polynomial DPD

polynomial. The ACLR improvement of both pre-distorter can reach more than 14.8dBc. The compensation of memory polynomial pre-distorter is more than that of LUT by 2.2dBc, which is due to the memory effect of wideband power amplifier.

5 Conclusions

After the study of non-linearity compensation of FDD-LTE RF amplifier, the digital pre-distorter based on LUT and memory polynomial both can be used to pre-distor the FDD-LTE signal under the RF power conditions in the lab. They both can make ACLR improve a lot and the memory effect of power amplifier can be compensated for FDD-LTE signal in a certain degree.

Acknowledgement. Fund:
Fujian major projects (2010HZ0004-1)
Fuzhou City, school science and technology cooperation (2011-G-105)
Fuzhou University Technology Development Fund (2011-XY-23).

References

1. Zhu, A., Wren, M., Brazil, T.J.: An efficient volterra-based behavioral model for Wideband RF power amplifiers. In: 2003 IEEE MTT-S International Microwave Symposium Digest, Philadelphia PA, pp. 787–790 (2003)
2. Bogya, R.I., Magana, M.E.: Linear Radio Frequency Power Amplifier Design using Nonlinear Feedback Linearization Techniques. In: IEEE 60th Vehicular Technology Conference, vol. 3 (May 2004)
3. Ding, Y., Ohmori, H., Sano, A.: Adaptive predistortion for high power amplifier with linear dynamics. In: The 47th IEEE International Midwest Symposium on Circuits and Systems, Hiroshima, Japan, pp. 121–124 (2004)
4. Liu, T., Boumaiza, S., Ghannouchi, F.M.: Pre-compensation for the dynamic nonlinearity of wideband wireless transmitters using augmented wiener predistorters. In: APMC 2005 Proceedings, Suzhou, China, pp. 2978–2981 (2005)
5. Luo, X., Liu, T., Ye, Y., Cen, L., Zhao, Y., Sun, C.: A 460MHz three-stage Doherty power amplifier with uneven power driving. In: 2011 International Conference on Electronics, Communications and Control, ICECC 2011, pp. 2404–2407 (2011)
6. Sun, C., Liu, T., Ye, Y., Zhao, Y., Luo, X., Li, J.: Optimization design of a broadband three-stage Doherty power amplifier. In: 2011 International Conference on Electronics, Communications and Control, ICECC 2011, pp. 364–367 (2011)
7. Atanaskovic, A., Males-Ilic, N., Milovanovic, B.: Harmonic Controlled Three-Way Doherty Amplifier with Improved Efficiency and Linearity, pp. 36–39 (2009)
8. Kim, I., Moon, J., Jee, S., Kim, B.: Optimized Design of a Highly Efficient Three-Stage Doherty PA Using Gate Adaptation. IEEE Transactions on Microwave Theory and Techniques 58(10), 2562–2574 (2010)

Cross-Polarization Characteristics of Bandpass FSS Structures

Lifeng Wang[1], Yang Wang[1], Zhijun Meng[1], Hao Li[2], and Chen Zhao[2]

[1] School of Aeronautic Science and Engineering, Beihang University, Beijing, China
wanglifeng@ase.buaa.edu.cn, zhuoyue771988@126.com
[2] Military Representative Office of Aviation Army in Beijing, China

Abstract. A study on the insertion phase delay (IPD) characteristics of frequency selective surface (FSS) in the solid half-wavelength wall was carried out. An emphasis was laid on cross polarization, and the depolarization effect at different polarization angles and incidence angles was studied. The result shows that to maintain consistency of the resonant frequency in two principal polarizations (TE & TM) is the solution to solve the problem of cross polarization of band-pass FSS structure, which therefore reduces the IPD difference between the two TE and TM polarization.

Keywords: Cross polarization, Frequency selective surface, Insertion phase delay (IPD), Electromagnetic wave.

1 Introduction

Frequency selective surfaces (FSSs) are spatial filters that are typically used in high performance radomes for radars and antennas. Band-pass FSS structures can pass the electro-magnetic wave in certain frequency ranges and shield others, which are especially suitable for out-of-band stealth of radomes [1, 2]. Airborne radome has several kinds of wall structures among which the solid half-wavelength wall structure is in wide use for its good electric and mechanics performance. Thus, to study the half-wavelength wall bandpass FSS structure is of great significance for the application of stealth radomes.

However, common studies on radome-wall structure with FSS are limited to the influence of factors as dielectric thickness, permittivity, and means of loading on some parameters like resonant frequency, transmission bandwidth and insert loss [3, 4]. There are few studies on other importance characteristics of FSS radome. Insertion Phase Delay (IPD), which is closely related to aiming error, has the same importance to radome as insert loss. In cross polarization, the insert loss of TE polarization is generally different from that of TM polarization, which could even leads to additional undesirable results such as polarization loss, higher side lobe level and larger aiming error. Characteristics of different dielectric structures have already been clearly studied, but current literature on FSS is only involved in the analysis of cross polarization loss due to the change of element shape [5]. Few studies are conducted on IPD and cross polarization characteristics of FSS radomes.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 155–161.
springerlink.com

In this paper, a study on the FSS structure in solid half-wavelength wall, which could be applied to the radome of fire control radar, is carried out. The analysis of IPD characteristics of principal and cross polarization will be performed on the single layer FSS structure in solid half-wavelength wall, and the influence of some factors such as dielectric and polarization will be analyzed. The achievements would do some help to the design of FSS radomes.

2 Model and Method

2.1 Modal Matching Method

Modal matching method, elaborated in a number of literatures [2], is a kind of moment method. The expansion of electric fields to Floquet modes is carried out in all areas, with continuous boundary condition at the interfaces. The integral equation describing the frequency response can be conducted, whose unknown variables are the tangential electric fields,

$$\sum_{r=1}^{2}\sum_{pq}(Y_{pqr}^{1}+Y_{pqr}^{(1)})\boldsymbol{\psi}_{pqr}\iint_{S}\boldsymbol{E}_{t}\cdot\boldsymbol{\psi}_{pqr}^{*}ds = 2\sum_{r=1}^{2}(A_{00r}\xi_{00r}\prod_{k=1}^{N}U_{00r}^{k}\boldsymbol{\psi}_{00r}) \tag{1}$$

where p and q are the grating harmonics because of the periodicity present in the array, r refers to TE or TM mode, $\boldsymbol{\psi}$ is the vector Floquet mode function, ξ_{00} is the admittance of free space, A00 is the incidence coefficient, Et is the tangential electric field in apertures, S is the aperture area, Y_{pqr}^{1} and $Y_{pqr}^{(1)}$ respectively refer to the equivalent admittance from FSS to the direction of $-\hat{z}$ and $+\hat{z}$. Method of moment is applied to solve the integral formula (1). The tangential electric fields are expressed by a set of basis function. After the weighting coefficients of the basis function have been solved, transmission and reflection coefficients therefore could be finally obtained with the help of transmission line theory.

2.2 Solution to IPD

By calculating the ratio of transmitted field and incidence filed in free space at the same observation point, the transmission coefficient containing the phase information could be obtained as

$$T_{H}=\frac{\hat{\boldsymbol{h}}\cdot\boldsymbol{H}(\boldsymbol{R}^{t})|_{k=n=0}}{\hat{\boldsymbol{h}}\cdot\boldsymbol{H}_{0}^{i}e^{-j\beta_{0}(\boldsymbol{R}^{t}-\hat{\boldsymbol{y}}\sum_{k=1}^{M+N}d_{k})\cdot\hat{\boldsymbol{r}}_{0+}}}=|T|e^{-j\varphi_{t}} \tag{2}$$

Where $\hat{\boldsymbol{h}}$ refers to the vector of polarization direct, |T| is the field transmission coefficient and φ_t is the value of IPD.

2.3 Cross Polarization

The electromagnetic waves from the antenna in the radar cabin are generally in principal polarization, but the curved radome wall would probably change the actual polarization of incidence field, defined as polarization angle. As shown in Fig.2, the incident field $\boldsymbol{E}_i$ is decomposed into vertical and parallel components. Then the polarization angle is defined as the angle between parallel component and principal polarization field

$$\phi_i = \arctan | E_{\perp}^{i} / E_{//}^{i} | \tag{3}$$

The transmitted field of parallel and vertical components would be

$$\begin{aligned} E_{//}^{t} &= T_{//} E_{//}^{i} = | T_{//} | e^{-j(IPD_{//})} E_{//}^{i} \\ E_{\perp}^{t} &= T_{\perp} E_{\perp}^{i} = | T_{\perp} | e^{-j(IPD_{\perp})} E_{\perp}^{i} \end{aligned} \tag{4}$$

The polarization angle φ_p of the transmitted field could be obtained as the form of equation (3). It is easy to find that φ_p is closely related to the IPD of both two polarizations. Define the difference between TM polarization and TE polarization as

$$\Delta IPD = IPD_{//} - IPD_{\perp} \tag{5}$$

Only when ΔIPD is 0, φ_p is equal to φ_i.

The direction of the transmitted field in Fig.1 is the original direction of antenna polarization, defined as the direction of principal polarization. The direction of cross polarization is perpendicular to the principal direction. The transmitted filed in principal polarization direction (E_M^t) and in cross polarization direction (E_C^t) could be obtained from the projection of vertical and parallel components in the principal and cross direction. And the cross polarization loss is defined as

$$C_{L-C} = 20 \cdot \log(| E_C^t | / | E_i |) \tag{6}$$

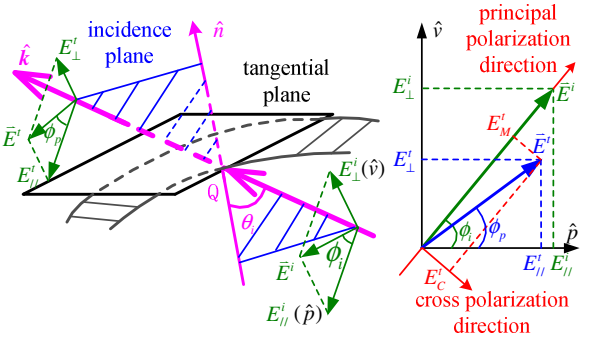


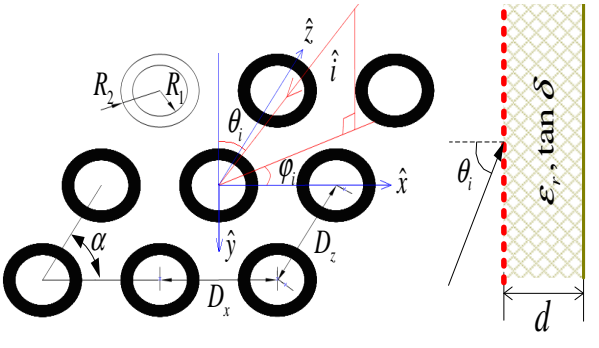
Fig. 1. The cross polarization of electromagnetic wave on radome surfaces

3 Analyses and Discussions

The FSS array of circular rings shown in Fig.2 is taken as the example. The dielectric slab is loaded on one side. Default parameters are given. The calculation principally works on the oblique incidence of 60º. The default dielectric thickness is 6.9mm, corresponding to the half-wavelength of the resonant frequency 12GHz at the incidence angle of 60º.

The polarization of transmitted field and the cross polarization loss are firstly investigated, taking the incident polarization angle $\phi_i = 45°$ as the example, as shown in Fig.3. As to pure dielectric structure, the polarization angle of transmitted filed of 6.9mm dielectric is $\phi_p = 45°$ at 12GHz, the same as φi, while that of 6mm dielectric is $45°$ at 13.9GHz. So the polarization angle of transmitted filed is always equal to the incident polarization angle when the dielectric thickness is half-wavelength, without depolarized effects and cross polarization loss. As to the FSS structure, the polarization angle of transmitted field approaches 45º and the polarization loss also gets close to zero at the resonant frequency. The polarization angles of FSS structure at out-of-band frequencies are more different from φi than that of dielectric, which means stronger depolarization effect and higher cross polarization loss.

The transmission characteristics of FSS structure in half-wavelength wall at different incident polarization angle are illustrated, shown in Fig.4. Results show that transmission curves change gradually from TM polarization to TE polarization when the incident polarization angle varies from 0º to 90º. Since the resonant frequency keeps stable and it corresponds to the half-wavelength thickness of the dielectric, ΔIPD maintains zero. So the cross polarization loss in Fig.4 (b) maintains 0 at resonant point. Moreover, the cross polarization loss is largest at 45º.



$R_1 = 2.7mm, R_2 = 3mm, D_x = D_z = 7mm, \alpha = 90°,$
$\theta_i = 60°, \varphi_i = 0°, \varepsilon_r = 4.0, \tan\delta = 0.015, d = 6.9mm.$

Fig. 2. FSS structure and parameters of the array

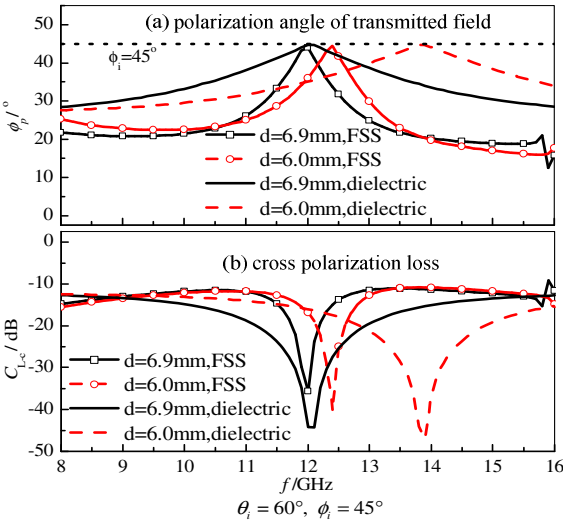


Fig. 3. The comparison of cross polarization between FSS structure and pure dielectric structure

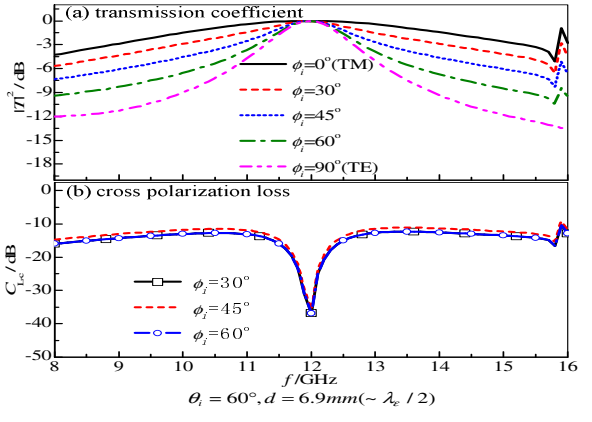


Fig. 4. Transmission characteristics of FSS in different incident polarization angle

Influence of incidence angle on the transmission characteristics of FSS structure in cross polarization is also studied, still taking $\phi_i = 45°$ as the example. As shown in Fig.5, there are transmission coefficient, cross polarization loss, polarization angle of transmitted filed and ΔIPD. Results show that the stability of transmission curves to incidence angle when $\phi_i = 45°$ is better than that at TE or TM polarization and the resonant frequencies shift little. At resonant point in all incidence angles, $C_{L\text{-}C}$ approaches zero, φ_p approaches 45º, and ΔIPD approaches zero, i.e., there is no depolarization effect. When the incidence angle gets smaller (30º), the changes of φ_p, ΔIPD and $C_{L\text{-}C}$ are a bit slight, while the depolarization effect is quite strong when the incidence angle is large (60º).

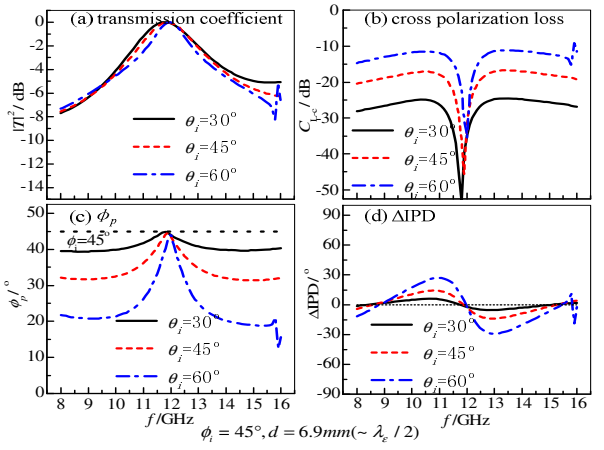


Fig. 5. Influence of incidence angle on the transmission characteristics of FSS in cross polarization

4 Conclusion

In this paper, the cross-polarization characteristics of FSS in the solid half-wavelength wall were studied. The influence of factors such as dielectric thickness and polarization were analyzed. An emphasis was laid on cross polarization, and the depolarization effect at different polarization angles and incidence angles was investigated. The depolarization effect of FSS structure is closely related to ΔIPD characteristics. At the resonant frequency, ΔIPD approaches zero, so the polarization angle of transmitted field approaches the incident polarization angle and the polarization loss also gets close to zero. When the incidence angle keeps constant, the depolarization effect is strongest at the incident polarization angle equal to 45º; when the incident polarization angle keeps constant, larger incidence angle leads to stronger depolarization effect.

A solution to the cross polarization problem of bandpass FSS structure can be found base on above results. That is to maintain consistency of the resonant frequency in the two principal polarization (TE & TM), which therefore reduces the IPD difference between TE polarization and TM polarization. The achievements would do some help to the design of FSS radomes.

References

[1] Munk, B.A.: Frequency Selective Surfaces: Theory and Design. Wiley-Interscience, New York (2000)

[2] Vardaxoglou, J.C.: Frequency Selective Surfaces: Analysis and Design. Research Studies Press, London (1997)

[3] Luebbers, R.J., Munk, B.A.: Some effects of dielectric loading on periodic slot arrays. IEEE Trans. Antennas Propagation AP-26, 536–542 (1978)
[4] Callaghan, P., Parker, E.A., Langley, R.J.: Influence of supporting dielectric layers on the transmission properties of frequency selective surfaces. IEE Proceedings-H 138, 448–454 (1991)
[5] Luebbers, R.J., Munk, B.A.: Cross polarization losses in periodic arrays of loaded slots. IEEE Trans. Antennas Propagation AP-23, 159–164 (1975)

A Low-Power Dual-Modulus Prescaler in 90nm CMOS Technology

Shibin Lu[1], Ying Meng[2], Feifei Wang[1], and Xianwei Jiang[1]

[1] School of Electronics and Information Engineering, Hefei Normal University, Hefei 230601, China
[2] School of Sciences, Anhui University of Science and Technology, Huainan 232001, China
{Shibin.Lu,Ying.Meng,Feifei.Wang,Xianwei.Jiang,LNEE}@Springer.com

Abstract. In PLL design, dual-modulus prescaler is one of the bottlenecks in achieving a higher operation speed. D flip-flop is the key factor limiting the speed of dual-modulus prescaler. By analyzing and improving the structure of Yuan-Svensson TSPC D flip-flop, the dual-modulus prescaler with a high speed, low voltage and low power is designed using TSMC 90nm 1P9M CMOS technology. It can be suitable for WLAN IEEE 802.11a communication standard. The circuit was simulated by Mentor Graphics Eldo. The simulated results show the working frequency is 5.8 GHz with supply voltage 1.2V at room temperature and power dissipation is only 0.8 mW. Its maximum operating frequency reaches 6.25 GHz.

Keywords: low power, dual-modulus prescaler, low voltage, CMOS.

1 Introduction

With the development of the new generation communications system, the demands of multi-mode interoperability wireless transceiver solutions are stronger[1-3].The high-speed frequency synthesizers and phase-locked loops (PLL) are desired in the broadband wireless communications[4]. Within the PLL, the voltage-controlled oscillator (VCO) and the frequency dividers are the key components for the highest operating frequency. Therefore, VCO and frequency divider are required to be high speed, low voltage and low power. Frequency integrated circuits and systems at the range of gigahertz using low-cost CMOS process[5-7]. Low Power consumption helps to increase the battery life time and to reduce the operating temperature[8,9].

Typically, a dual-modulus prescaler is constructed with two parts-a synchronous counter and an asynchronous counter to reduce the number of flip-flops operating at high frequency and the power consumption. In this two-stage approach, the synchronous counter is the critical part in determining the speed of a prescaler. The optimization of D-flip-flop (DFF's) in the synchronous counter is essential to increase the operating frequency of the prescaler. The high-speed operation of MOS transistors is limited by their low transconductance. Therefore, new circuit techniques such as

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 163–168.
springerlink.com

dynamic and sequential circuit techniques must be used in designing the synchronous counter. True-single-phase-clocked (TSPC) dynamic DFF's and differential latches are examples[10,11].

This work presents the dual-modulus prescaler with high speed, low voltage and low power, using TSMC 90nm COMS technology. The frequency division ratio of the dual-modulus prescaler can be changed. It can be suitable for WLAN IEEE 802.11a communication standard. The maximum operating frequency of the proposed prescaler is up to 5.8 GHz at the supply of 1.2V. The circuit consumes on only 0.8 mW at input frequency of 5.8 GHz.

2 Designing Circuit

The proposed dual-modulus prescaler divide-by-32/33 is constructed with three parts: a synchronous counter divide-by-2/3, a cascaded asynchronous counter divide-by-16 and feedback circuit as shown in Fig. 1. This two-stage topology is the result of optimization at architecture level, which reduces the number of flip-flops operating at high frequency and power consumption. In such a topology, the synchronous counter is the only part of the circuit operating at the maximum frequency, and the rest of the circuit runs at maximum one-fourth of the input frequency. A five input OR gate is used to generate the modulus control signal (MC) that decides the division ratio (2 or 3) of synchronous counter, while the fractional division ratio of the prescaler is selected according to Mode signal. When the Mode is high (low), the division ratio is set to 32 (33).

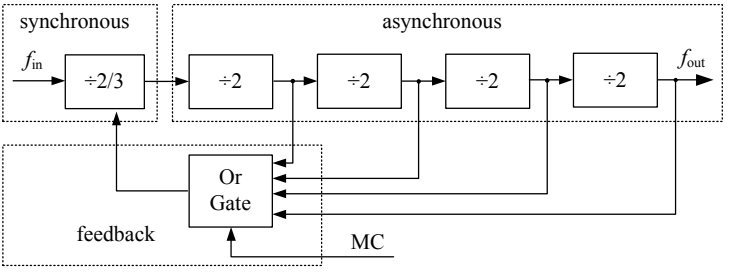


Fig. 1. Divide-by 32/33 dual-modulus prescaler

Fig. 2 shows the topology of divide-by-2/3 prescaler, which has the highest operation frequency of the whole divider. It is the key parts to influence the speed and power consumption of prescaler. The prescaler is divided by two when the mode control signal MC is set to high, and is divided by three when the mode control signal MC is set to low. Using asynchronous divide-by-2/3 counter reduce the number of MOSFET operating at high frequency, so the power consumption of asynchronous circuit is decreased. The D-flip-flops with AND gate avoid the influence of parasitic parameter to improve operation speed and decrease power consumption.

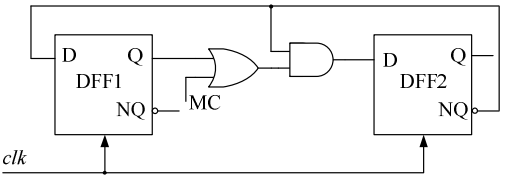


Fig. 2. Typical topology of divide-by-2/3 prescaler

Fig. 3 is the operating speed of the Yuan-Svensson TSPC DFF[12]. The circuit consist of the N-C^{2}MOS stage and the N-precharge stage. RC delay time is big due to each stage having three MOSFETs. The circuit schematic of the ratioed latch shows in Fig. 4. Two cascaded pseudo-nMOS inverters in the ratioed latch replace the N-precharge stage and the N-C^{2}MOS stage in Fig. 3.

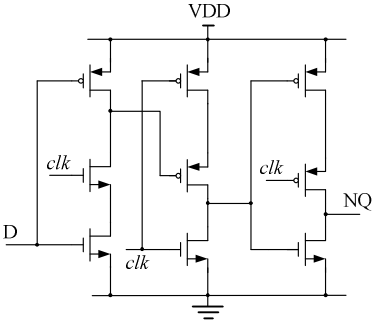


Fig. 3. Diagram of Yuan-Svensson D-flip-flop

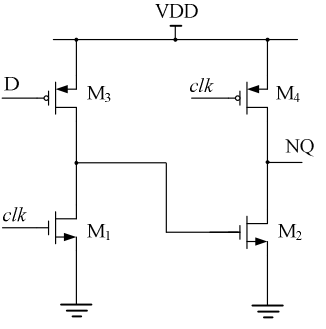


Fig. 4. Schematic of a ratioed latch

Fig. 5 shows the circuit schematic of the proposed negative edge-triggered DFF. When the clock is low(high), the ratioed latch is in the evaluation mode(hold mode).

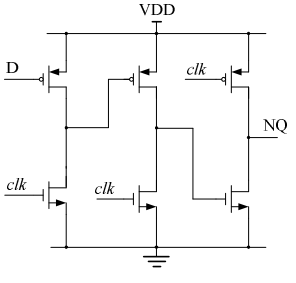


Fig. 5. Proposed dynamic D-filp-flop

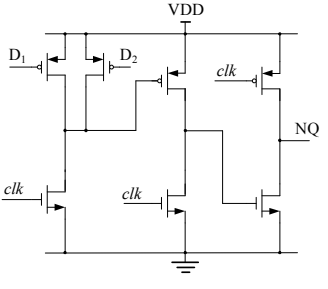


Fig. 6. D-flip-flop integrated AND gate

In contrast to the TSPC DFF in Fig. 3, only six transistors are necessary for the basic flip-flop function, while nine transistors are needed to perform the same function in the case of the TSPC DFF. The number of transistors driven by the clock signal is also reduced, which increases the driving capability of the clock driver and decreases the power consumption substantially.

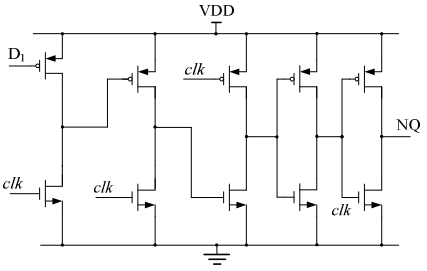


Fig. 7. Asynchronous divider

As shown in Fig. 7, the signal frequency is decreased by divide-by-2/3. In order to avoid logic chaos, a buffer is added at the output Q. The asynchronous chain which composed of DFF attains better compromise among speed, frequency and power consumption.

3 Simulation and Analysis

The hold mode operation is as follows. The W/L ratios of M_1 and M_3 are determined so that M_2 and M_4 remain in cut-off during the high period of the clock. Therefore, the signal path from the input *D* to the output *Q* is not transparent and the latch enters into the hold mode. When the clock changes its state from high to low, the latch enters into the evaluation mode. If the input D is low, M_2 and M_4 is breakover. The pull-down strength of M_2 must be sufficiently larger than the pull-up strength of M_4, so that the output low voltage is lower than the input low voltage of the following gates.

Based on TSMC 90nm CMOS technology, the circuit is simulated using Eldo of Mentor Graphics Corporation. Fig. 8 shows divide-by-2/3,4/5 operation. Fig. 9 shows divide by-8/9, 16/17, 32/33 operation.

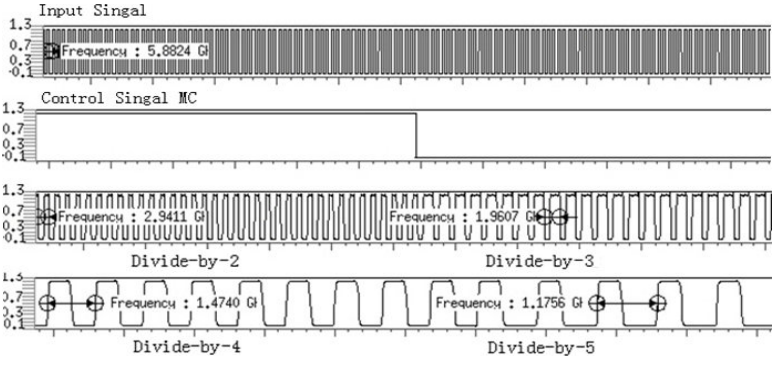


Fig. 8. Divide-by-2/3,4/5

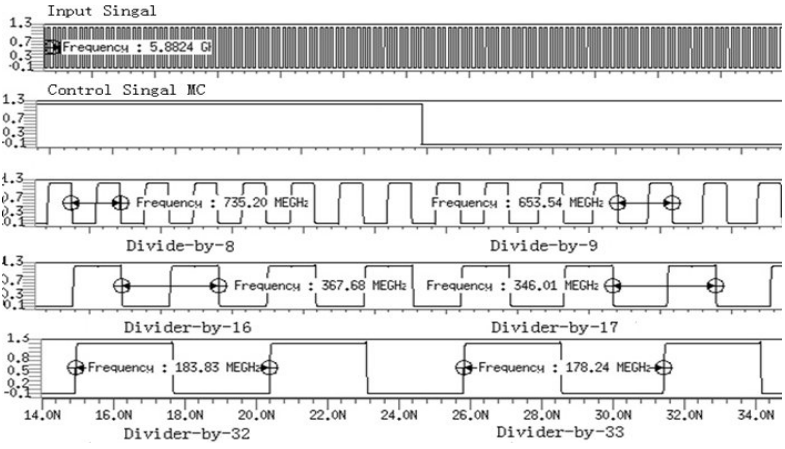


Fig. 9. Divide by-8/9, 16/17, 32/33

4 Conclusion

A high-speed dual-modulus prescaler has been designed using TSMC 90nm COMS technology. The modification of the conventional Yuan-Svensson TSPC D flip-flop is made. The structure of dynamic ratioed D flip-flop is presented and the number of MOSFET decreased. This flip-flop is simple and can work at very high frequency, low power. The simulated results show power dissipation is 0.8 mW and the operation frequency is 5.8GHz with supply voltage 1.2 V at room temperature.

Acknowledgments. This work was supported by Natural Science Foundation of Anhui High Education (KJ2012B143), Foundation for Outstanding Young Talents in Anhui Higher Education (2011SQRL130), and Fund of Key Science Research Base, Hefei Normal University (2012jd16).

References

1. Ray, M., Souder, W., Ratcliff, M., Dai, F., Irwin, J.D.: A 13GHz Low Power Multi-Modulus Divider Implemented in 0.13μm SiGe Technology. In: IEEE Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems, pp. 1–4 (2009)
2. Changhui, H., Khanna, R., Nejedlo, J., Kangmin, H.: A 90 nm-CMOS, 500 Mbps, 3-5 GHz Fully-Integrated IR-UWB Transceiver With Multipath Equalization Using Pulse Injection-Locking for Receiver Phase Synchronization. IEEE J. Solid-State Circuits 46, 1076–1088 (2011)
3. Yu, X.P., Do, M.A., Ma, J.G., Yeo, K.S.: A new 5 GHz CMOS dual-modulus prescaler. In: IEEE International Symposium on Circuits and Systems, pp. 5027–5030 (2005)
4. Alioto, M., Mita, R., Palumbo, G.: Design of high-speed power efficient MOS current-mode logic frequency dividers. IEEE Trans. Circuits Syst. II, Exp. Briefs 53, 1165–1169 (2006)

5. Ng, A.W.L., Leung, G.C.T., Kwok, K.-C., Leung, L.L.K., Luong, H.C.: A 1-V 24-GHz 17.5-mW phase-locked loop in a 0.18-μm CMOS process 41, 1236–1244 (2006)
6. Kuo, Y.F., Weng, R.M.: 5 GHz low power frequency synthesiser with dual-modulus counter. Circuits, Devices & Systems, 365–375 (2009)
7. Henzler, S., Koeppe, S.: Design and Application of Power Optimized High-Speed CMOS Frequency Dividers. IEEE Transactions on Very Large Scale Integration Systems 16, 1513–1520 (2008)
8. Razavi, B.: A study of injection locking and pulling in oscillator. IEEE J. Solid-State Circuits 39, 1415–1424 (2004)
9. Johansson, H.O.: A simple precharged CMOS phase frequency detector. IEEE J. Solid-State Circuits 33, 295–299 (1998)
10. Foroudi, N., Kwasniewski, T.A.: CMOS high-speed dual-modulus frequency divider for RF frequency synthesis. IEEE J. of Solid-State Circuits 30, 93–100 (1995)
11. Yuan, J., Svensson, C.: High-speed CMOS circuit technique. IEEE J. Solid-State Circuits 24, 62–70 (1989)
12. Hanafi, B., Sharaf, K.: 7.5-GHz injection-locked dual modulus prescaler with 1-GHz locking range. In: Internatonal Conference on Microelectronics, pp. 307–310 (2007)

Design of a Novel Leak Detection System for Water Network in Electronic Communication

Lingjuan Zhang[1], Yun Wang[2], and Jianqiang Ren[2]

[1] Department of Physics and Electronic Information,
Langfang Teachers College,
Langfang, Hebei, China
zhanginguan@163.com
[2] Department of Computer Science and Technology,
Langfang Teachers College,
Langfang, Hebei, China
renjianqiang@163.com

Abstract. In order to solve the shortcoming of the traditional leak detection of water supply network, a novel detection system is designed based on ZigBee wireless sensor network. The system is divided into the front-end monitoring nodes, the relay nodes and the central node. All nodes are designed based on the embedded system and interconnected based on the ZigBee CC2430 chip, and the display of the parameter data is based on MZL04-12864 LCD display module. In the monitoring nodes, the three-terminal linear Hall sensor of UGN3501T is used for leak detection. Experimental results show that the design is reasonable and the system is excellent, the average response time of the system is only 2.95 seconds and the average leak detection rate is as high as 98%.

Keywords: Leak detection, ZigBee, CC2430, Hall, Embedded system, electronic communication.

1 Introduction

Water resources play an important role in our production and daily life. But the conditions of present water resources should not be well in the world. In our country, the water shortage is highly serious and the quantity per capita is only a quarter of the world average. As for the effective control of water, how to reduce the waste of water resources has become widely concern.

The leak detector for water network was widely researched [1-2]. The water outage still happens from time to time. Frequently, the downstream valves are not closed and the leak happens after water resupply. It can cause severe loss. A leakpro automatic alarm system of water supply pipe was designed which is based on the real-time monitoring and the GSM technic in [3]. The system can better solve the leak problem which results from not turning off the upper reaches when the water supply pipe stops working for some reason and continue supplying again in case of no waiting. In order to solve the shortcomings of the leak detection for water supply network, a novel detection system was designed based on UGN3501T and embedded ARM chip of S3C2410A in [4], and the detection system has the advantages of leak detection, automatic alarm, water metering and so on. But these functions have common

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 169–174.
springerlink.com

shortcomings, that is, the GSM is more expensive and the multi node design is unsuitable. To solve these problems, the ZigBee wireless sensor net was used to design the novel leak detection system for water network, and the system performance was further enhanced significantly.

2 Structure of System

The general structure of the novel leak detection system for water network is shown in Fig.1. It mainly consists of monitoring nodes, scope relay nodes and center node. The monitoring nodes are set up on the terminal line of water network. If the leak is detected, the node will close the shut-off valve and send out alert messages to the corresponding scope relay node through ZigBee wireless network. The relay node will enhance the power of the received signal and forward it to the center node or administrators. The center node identifies the leak position through the node ID. The administrator can precisely locate the leak position of the water network and the needed repairs will be done.

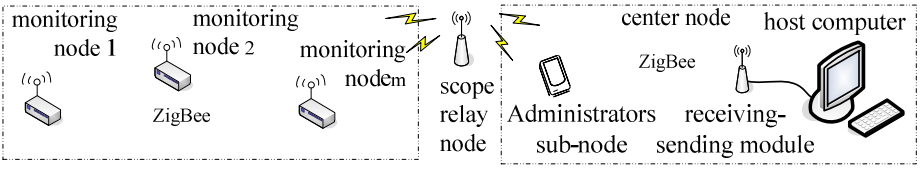
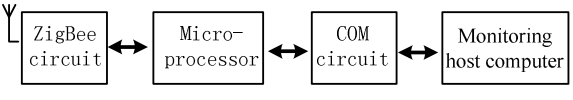


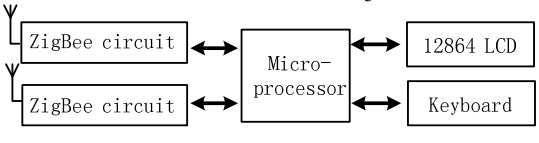
Fig. 1. General structure of the system

3 Design of Hardware Structure

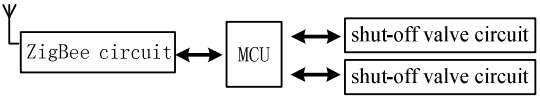
The hardware structure of the novel leak detection system is shown in Fig.2. In hardware design process, the ZigBee circuit, leak detection circuit, shut-off valve control circuit and data display circuit are the key tasks.



（a） Hardware structure of monitoring centre node



（b） Hardware structure of relay nodes



（c） Hardware structure of monitoring node

Fig. 2. Block diagram of hardware structure

3.1 Design of ZigBee Communication Circuit

The three parts of novel detection system communicate with each other through the ZigBee wireless network, so the design of ZigBee circuit is the primary task. In view of existing situation, the ZigBee technology realization mainly includes ZigBee RF + MCU mode, single-chip SOC mode and single-chip internal protocol stack + adscititious chip mode. The single-chip SOC mode has the merits of needing a little additional circuit, which is simple and easy to operate. So, the single-chip SOC mode is used in this paper, and based on TI/CHIPCON chip of CC2430, the ZigBee circuit is designed.

The CC2430 circuit in this system is shown in Fig.3. The CC2430 comes in three different versions: CC2430-F32/64/128, with 32/64/128 KB of flash memory respectively. The CC2430 is a true System-on-Chip solution specifically tailored for IEEE 802.15.4 and ZigBee™ applications. It enables ZigBee nodes to be built with very low total bill-of-material costs. The CC2430 combines the excellent performance of the leading CC2420 RF transceiver with an industry-standard enhanced 8051 MCU, 32/64/128 KB flash memory, 8 KB RAM and many other powerful features. Combined with the industry leading ZigBee™ protocol stack (Z-Stack) from Figure 8 Wireless/Chipcon, the CC2430 provides the market's most competitive ZigBee™ solution. The CC2430 is highly suited for systems where ultra low power consumption is required. This is ensured by various operating modes. Short transition times between operating modes further ensure low power consumption [6].

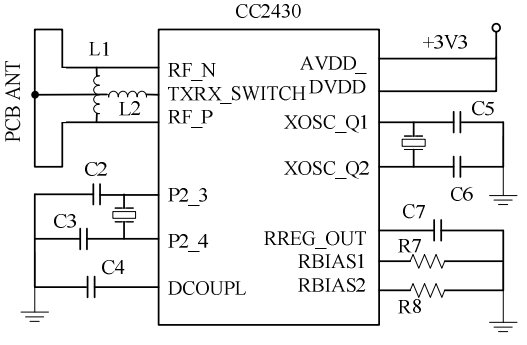


Fig. 3. Circuit of ZigBee CC2430

3.2 Design of Leak Detection and Valve Control Circuit

The leak detection circuit is designed based on UGN3501T, which is an integrated Hall switch-sensor made by the SPRAGUE company [4]. The circuit of leak detection and valve control is shown in Fig.4. A gear disk is attached to the inside of the water pipe. The fast-moving stream of water drives the gear disk. The gear disk turned round the UGN3501T and UGN3501T outputs pulse voltage peaking at 20mV. Firstly, the output voltage is amplified by amplifier μA741 and adjusted by triode 2N5812. Then, the output of 2N5812 is input into an input terminal of MCU. When the pipe leak should happen, MCU will close the shut-off valve by the small relay

JZC-23 and send out alert messages to the corresponding scope relay node through ZigBee wireless network.

The water-supply status detection circuit is used to assess the situation of water leakage. The detection circuit is based on a Schmitt trigger and two electrodes attached to the inside of the water pipe [3]. When the water outage happens, the water level in pipe lowers and the two electrodes cut off. When the water resupply, the two electrodes are connected again and the Schmitt trigger outputs a positive pulse into MCU. Now, if the flow of water detected by UGN3501T exceeds the threshold, the MCU will adjudge the pipe leakage.

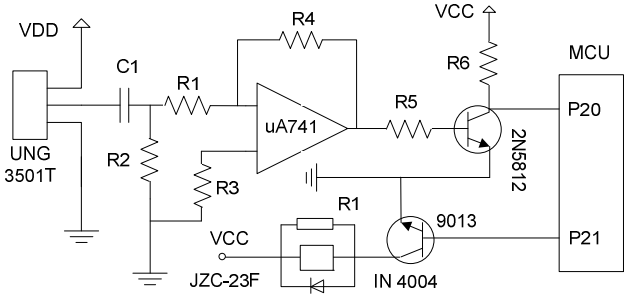


Fig. 4. Circuit of leakage monitoring and switch off

3.3 Design of Data Display Circuit

The design of data display circuit is another key focus for the system. In this paper, MZL04 12864 LCM is used to design the data display circuit. Visual region of the MZL04 12864 LCM is 33.26mm×20.46mm and the power voltage is 3.3V. In the design, the MZL04 12864 LCM is connected with MCU through SPI interface, and the connection circuit is shown in Fig.5.

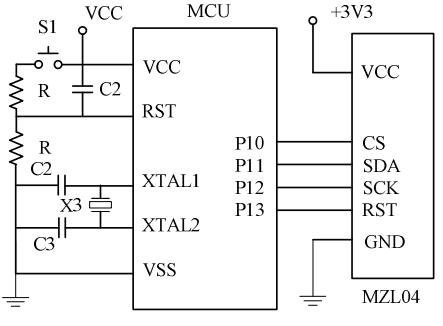
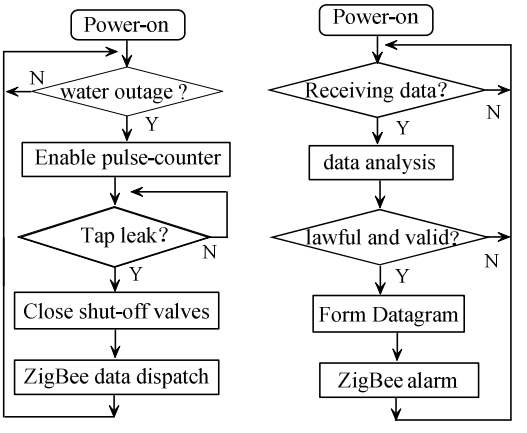


Fig. 5. Circuit of the data display

4 Design of Software Structure

The software design task includes design of the monitoring nodes software, scope relay nodes software and center node software. The monitoring nodes software and

scope relay nodes software are designed based on embedded C language, and the flow charts are shown in Fig.6. The monitoring nodes software is used to realize the water outage monitoring, the leak detection, the valve shut-off and the ZigBee alarm message sending. The scope relay nodes software is used to realize the message receiving and relay, which is sent from the monitoring nodes. After receiving the message, the cope relay nodes software check its validity. Only the proper message can be relayed to the center node. The center node software for PC is designed based on VB language and ACCESS database programming, and it is used to realize the recording, collecting and reporting the alarm messages. The flow chart of it is very simple. No more about this will be mentioned in this paper.



(a) Flow chart of monitoring node (b) Flow chart of relay nodes

Fig. 6. Flow chart of detection system software

5 Experiment and Analysis

In order to testify the performance of the proposed scheme, the experiment was carried out and 30 monitoring nodes were made in it. The system prototypes were tested in one residential area. The experimental results show that the proposed project has higher precision and better performance. The average response time of the system was only 2.95 second and the average leakage detection rate was 98%.

6 Conclusion

A novel leak detection system for water network based on zigbee was studied in this paper and the aim of this work was to solve the shortcomings of existing leak detection system. In our design project, the system mainly consists of monitoring nodes, scope relay nodes and center node.

The various parts are connected with each other via a ZigBee wireless network. When the control nodes monitor the current leak occurring, the following domain

sends the alarm information through the jurisdiction to control center node and the administrator terminal nodes. ZigBee technology effectively improves system performance, enhances networking capabilities, and significantly reduces system cost. Application case has already proven the rationality and use value of the system.

Acknowledgment. Funding for this research is provided by the support of science and technology of Hebei province (P.R. China) (the project number is 11213590) and the education department of Hebei province (P.R. China) (the project number is Z2011118) and the Natural Science Fund for Young Scholars of Lang-fang Teachers College (the project number is LSZQ201006).

References

1. Kong, D.M., Wang, Y.R.: Research on leakage localization for underground water pipeline network and development of leak detection system. Chinese Journal of Scientific Instrument 2, 421–424 (2009)
2. Liu, Y.: Leak Detection of Water Supply Pipeline Based on the Sound Wave Theory. Nondestructive Testing 6, 355–356 (2008)
3. Zhang, L.J.: A Leakpro Autocontrol and Autoalarm System of Water supply Pipe Based on Real-Time Monitoring. Microcomputer Information 23, 29–31 (2007)
4. Fan, Y.S.: Design of Leak Sensing Detector for Water Supply Network. Automation & Instrumentation 6, 12–14 (2010)
5. Zhang, Y., Zhang, G.Q.: Wireless Sensor Network System Based on ZigBee for Aircraft Condition Monitoring. Computer Measurement & Control 11, 2481–2484 (2010)
6. Ma, Y.Q., Li, J.Q., Feng, L.Y.: RF chip CC2430 Based on ZigBee Technology. Microcontrollers & Embedded Systems 3, 45–47 (2006)

Minimization of Binary Decision Diagrams by Adaptive Hierarchy Genetic Algorithm and Its Application in Circuit Test Generation

Zhongliang Pan and Ling Chen

School of Physics and Telecommunications Engineering, South China Normal University, Guangzhou 510006, China
panz@scnu.edu.cn

Abstract. The binary decision diagram is a graph representation of Boolean functions, it also can efficiently represent the logic functionality of digital circuits. The number of nodes in binary decision diagram depends on the variable ordering. A new method is presented in this paper for the variable ordering and the minimization of binary decision diagram, the method is based on adaptive hierarchy genetic algorithm, which uses two populations with hierarchy to evolve in parallel, and the operation parameters in the evolution procedure are adjusted dynamically as the changes of environment. The individual representation scheme corresponding to binary decision diagram, and the implementation steps of adaptive hierarchy genetic algorithm, are given in detail. Besides, the application of the method for the test generation of digital circuits is given. A lot of experimental results show that the binary decision diagrams with smaller number of nodes can be obtained by the method proposed in this paper, and the test vectors of faults in digital circuits can also be produced.

Keywords: Digital circuits, binary decision diagrams, variable ordering, adaptive genetic algorithms, test generation, fault detection.

1 Introduction

The binary decision diagram (BDD) is a graph representation of Boolean functions, one important property of binary decision diagram is that it can give the canonical form of Boolean functions, and the size of the graph, i.e. the number of nodes in the graph, is feasible for many practical Boolean functions [1,2]. The operations of Boolean functions can be implemented as graph algorithms operating on BDDs [3].

The BDD is very sensitive to the variable ordering, the variable ordering used has a significant impact on the number of nodes in the BDD graph, from linear to exponential in the number of variables. For example, Ebendt [4,5] presented the approaches for the minimization of BDD by the exact branch and bound algorithm, or by the combination of the ordered best-first search with a classical branch-and-bound algorithm. Wei [6] proposed a new global method to the variable ordering problem of pseudo-symmetric binary decision diagrams (PSBDDs).

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 175–180.
springerlink.com

Drechsler [7] studied the effect of lower bounds being computed, and made the lower bounds to incorporate with dynamic BDD minimization algorithms. Hung [8] studied the BDD minimization problem by using scatter search approach. Suguo [9] discussed an ordering method named as progressive neighbor first ordering, which emphasizes the logical relationship of variables. Ebendt [10] proposed an approach for the variable ordering of BDD, which is based on a goal-directed best-first search.

In this paper, a new method is presented for the variable ordering and the minimization of binary decision diagram, the method is based on the adaptive hierarchy genetic algorithm, besides, the application of the method for the test generation of digital circuits is studied.

2 Structure of Binary Decision Diagrams

The binary decision diagram (BDD) is a representation of a Boolean function by using a directed acyclic graph. For a BDD, there are two type of nodes: terminal nodes (also called leaf nodes) and nonterminal nodes. Each nonterminal node is labeled with a variable, it has two edges directed toward two children, where two edges are named 0-branch and 1-branch. The 0-branch corresponds to the case where the variable is assigned 0, and the 1-branch corresponds to the case where the variable is assigned 1. Each leaf node is labeled 0 or 1 which is corresponding to the value of the Boolean function for a specific assignment to the input variables.

A path in BDD from the root node to a leaf node corresponds to a truth table entry of Boolean function, where the edges in the path corresponds to the value assignment of Boolean variables, and the value of the leaf node is the value of the Boolean function under that assignment. For a given logic function, the number of nodes in its BDD depends on the variable ordering used. Here we take the BDD of logic function $h=x_1x_2+x_3x_4+x_5x_6$ as an example. For the variable ordering $x_1<x_2<x_3<x_4<x_5<x_6$, the number of nodes in the BDD of function h is 8. For the variable ordering $x_1<x_3<x_5<x_2<x_4<x_6$, the number of nodes in the BDD of function h is 16.

Therefore, the choice of variable ordering largely influences the size of BDD. In the following Section 3, we present a method for the minimization of BDD, which is based on the adaptive hierarchy genetic algorithm.

3 Minimization of BDD by Adaptive Hierarchy Genetic Algorithm

In this Section 3, the minimization of BDD by using adaptive hierarchy genetic algorithm is presented. The genetic algorithm with hierarchy is a nested evolutionary algorithm composed by two evolutionary levels: the lower level evolution and the higher level evolution. The two evolutionary levels work on different populations and fitness functions. There is individual migration between the two evolutionary levels to share genetic materials. The genetic algorithm with adaptive mode has following features: the parameters of population size, the number of points of crossover, and mutation rate for each population, are adaptively updated.

There are two main aspects in the implementations of the minimization of binary decision diagrams by adaptive hierarchy genetic algorithm: (a) The representation of individual, and the definition of fitness function. (b) The implementation steps of adaptive hierarchy genetic algorithm. They will be given in detail in the following.

3.1 Representation and Fitness of Individuals

Suppose all logic Boolean functions are expressed over the variables $x_1, x_2, \cdots, x_n$ for the operations of BDD. We use an integer string of length *n* as the coding of an individual corresponding to a variable ordering of BDD, the *n* denotes the number of variables of the BDD being considered. For example, for a given circuits, we use (3 4 7 6 1 2 5) as the individual representation corresponding to the BDD under the variable ordering $x_3<x_4<x_7<x_6<x_1<x_2<x_5$. A population is a set of these individuals.

The fitness $f(X_i)$ of an individual X_i is related to the number of nodes of the BDD corresponding to the chosen variable ordering. The fitness of an individual is defined by the following equation: $f(X_i)=K_0/m$. Where the K_0 is a positive constant, the *m* is the number of nodes of the BDD corresponding to the individual (i.e. the chosen variable ordering of the BDD).

3.2 Minimization of BDD

The genetic evolution shown in Algorithm 1 is used for the minimization of BDD, which is based on adaptive hierarchy genetic algorithm. Let U(*t*) be the population in higher level evolution, let L(*t*) be the population in the lower level evolution.

Algorithm 1

Step 1. Set the value of parameter *t* is equal to 0. Set the number of individuals in U(0) and L(0) is *N* respectively. Randomly generate every individual in the initial population U(0) and L(0).

Step 2. Carry out the genetic operations for U(*t*) and L(*t*) respectively; Repeat this step in *M* iterations.

Step 3. Migrate a lot of individuals between the populations U(*t*) and L(*t*).

Step 4. Randomly remove one individual from the current population U(*t*), add the best individual in the previous population in higher level evolution to the current populations U(*t*). Similarly, randomly remove one individual from the current population L(*t*), add the best individual in the previous population in lower level evolution to the current population L(*t*). Therefore, the best individual in each type of populations always survives to the next population by the operations of this step.

Step 5. If the termination condition is reached, then the algorithm is stopped; otherwise, set $t := (t+1)$, go to Step 2.

In the Algorithm 1, the *N* and *M* are positive integers, the evolution modes of U(*t*) and L(*t*) are different. The U(*t*) makes use of adaptive genetic evolution, the L(*t*) uses the evolution mode in conventional genetic algorithms. The evolution process in Algorithm 1 for two populations U(*t*) and L(*t*) takes parallel competition and isolating mechanism. The detail implementations of the steps in Algorithm 1 are following.

(1) For the Step 2 in Algorithm 1, the adaptive evolution strategy is used for the higher level evolution of population U(t).

First of all, the number of individuals in initial population U(0) is set to N. For t=0, 1,2,···, the population size of each population U(t+1) is adaptively updated as follows:

$$n_{t+1} = n_t + \frac{f_{\max}}{f_{avg}} - 1$$

The n_t and n_{t+1} is the population size of population U(t) and U(t+1), respectively. The population size is the number of individuals in the population. The f_{max} is the fitness of the best individual in population U(t). The f_{avg} is the average fitness of the individuals in population U(t). Therefore, by taking this updating mode for population size, the population size is dynamically varied with the fitness of the best individual in the population compared to the mean fitness of the individuals in the population.

Secondly, the number of crossover points for reproduction operation is adaptively updated by the following updating mode:

$$C_{t+1} = C_t + \frac{n_{avg}}{n_t} - 1$$

Where the C_t and C_{t+1} is the number of crossover points used for reproduction operation in populations U(t) and U(t+1), respectively. The n_t is the population size of U(t), the n_{avg} is the average number of individuals in previous three populations: U(t), U(t-1) and U(t-2). By using this updating mode, the number of crossover points for reproduction operation will be increased when the number of individuals in current population is less than the mean number of individuals in previous three populations.

Thirdly, the mutation probability for mutation operation is adaptively updated by the following equation:

$$Q_{t+1} = Q_t + \xi \cdot \left(\frac{n_{avg}}{n_t} - 1 \right)$$

The Q_t and Q_{t+1} is the mutation probability being used for the populations U(t) and U(t+1), respectively. The ξ is a small constant such as ξ=0.001. By using this updating mode for mutation probability, the number of individuals being mutated is increased for relatively small population sizes.

(2) For the Step 2 in the Algorithm 1, the evolution strategy in conventional genetic algorithms is used for the lower level evolution of population L(t).

(3) For the Step 3 of the Algorithm 1, the approach of migrating individuals between populations L(t) and U(t) is: choose λ best individuals with highest fitness from population L(t) to substitute the λ individuals being chosen randomly in population U(t). Where the λ is a positive integer and it is less than N/4, the N is the population size of population L(t). For the Step 5 of the Algorithm 1, the termination condition is defined as the optimal individual has been found or the maximal number of evolution iterations has been reached.

The above method for BDD minimization by adaptive hierarchy genetic algorithm has been implemented in C++, the experimental results will be given in Section 3.4.

3.3 Test Generation of Digital Circuits Based on BDD

In general, a lot of faults may be introduced into a circuit or device during the manufacturing process, thus the circuit testing is needed to provide a measure of the quality or reliability of the finished circuit. If there is a fault in a digital circuit, then the test vectors of the fault are the input assignments that cause the faulty circuit and normal circuit (there are not faults in the circuit) produce different output values.

The BDD for the normal circuit is constructed, the BDD is called as normal BDD. The BDD for the faulty circuit is constructed, the BDD is called as faulty BDD, the faulty BDD describes the functionality of the faulty circuit. For a fault, when its normal BDD and faulty BDD are isomorphic, this shows the fault isn't testable, that is there d isn't test vectors for the fault.

The BDD called as test BDD is constructed, which is the XOR operation of the normal BDD and faulty BDD. For the test BDD, each input assignment that leads to the leaf node labeled 1 is a test vector of fault. The flowchart is shown as follows.

Test generation based on BDD
{ Build the normal BDD of the normal circuit;
For each fault in the fault table of circuit, do
{ Build the faulty BDD of the faulty circuit;
IF the normal BDD and faulty BDD are isomorphic, THEN the fault isn't testable, RETURN 0;
ELSE { Build test BDD;
For the test BDD, search for all input assignments that lead to leaf node labeled 1, RETURN test vectors. } } }

3.4 Experimental Results

The method in this paper for the minimization of BDD by using adaptive hierarchy genetic algorithm has been implemented in C++ language, and the method is applied to the generation of the test vector for digital circuits. A lot of experiments have been performed on a personal computer with 3.0GHz and 512MB memory under Windows operation system. The following ISCAS'85 benchmark circuits are used: C432, C499, C1908, C2670 and C5315.

The parameters used in the Algorithm 1 are as follows. The maximal number of evolution generations is 500, the population size of L(t) is 40, the initial population size of U(t) is 40. The values of parameters ξ, μ, τ and β is 0.001, 8, 0.9 and 0.95, respectively. The number of migrating individuals between L(t) and U(t) is 5.

For the minimization of BDD, we have also performed another experiments by using conventional genetic algorithms (GA) in order to compare the method in this paper with GA. The parameters used in conventional genetic algorithm are: the maximal number of evolution generations is 500, population size is 80, crossover rate is 0.85; mutation rate is 0.0001, the roulette wheel selection scheme and two-point crossover are used. Each method (Algorithm 1 and conventional genetic algorithms) is performed 20 times repeatedly for a benchmark circuit. For the circuits C432, C499, C1908, C2670 and C5315, the sizes of the BDDs obtained by the Algorithm 1 and conventional genetic algorithm are following: 1392, 28164, 6614, 2807, 2126, and 2521, 43072, 7683, 4729, 3841, respectively. For the generation of the test vector for the stuck-at faults in digital circuits, we have carried out a lot of experiments for

the ISCAS'85 benchmark circuits C432, C499, C1908, C2670 and C5315. The variable orderings obtained by Algorithm 1 is used for building the BDDs of these circuits. The experimental results show that if there are test vectors for a given stuck-at fault, then the test vectors can be produced by using the BDDs.

Summarize these experimental results, it is shown that the method proposed in this paper can get the good variable ordering of BDD, and is able to obtain the BDD with smaller number of nodes than conventional genetic algorithms. For the generation of test vectors for stuck-at faults in digital circuits, one advantage of the method is that all test vectors of a fault can be obtained.

4 Conclusions

The number of nodes in BDD depends on the variable ordering. In this paper, a new method is presented for the variable ordering of BDD by using adaptive hierarchy genetic algorithm. The application of the method for the generation of test vectors for digital circuits is studied. The experimental results show the BDD with smaller number of nodes can be obtained by the method in this paper, the test vectors of faults can also be obtained. In the future, some work needs to be done for the better selections of parameters values in adaptive hierarchy genetic algorithm.

Acknowledgments. This work were supported by National Natural Science Foundation of China (No.61072028), Guangdong Province & Chinese Ministry of Education Cooperation Project of Industry, Education, and Academy (No.2009B090300339).

References

1. Bars, J.M., Viola, A.: Equivalence classes of Boolean functions for first-order correlation. IEEE Trans. on Information Theory 56(3), 1247–1261 (2010)
2. Chou, K.C.: BDD-based synthesis for mixed CMOS/PTL logic. International Journal of Circuit Theory and Applications 39(9), 923–932 (2011)
3. Lhotak, O., Curial, S., Amaral, J.N.: An optimal encoding to represent a single set in an ROBDD. IEEE Trans. on Computers 59(4), 574–575 (2010)
4. Ebendt, R., Gunther, W., Drechsler, R.: An improved branch and bound algorithm for exact BDD minimization. IEEE Trans. on CAD 22(12), 1657–1663 (2003)
5. Ebendt, R., Gunther, W.: Combining ordered best-first search with branch and bound for exact BDD minimization. IEEE Trans. on CAD 24(10), 1515–1529 (2005)
6. Wei, W., Chrzanowska, J.M.: A global approach to the variable ordering problem in PSBDDs. In: IEEE International Symposium on Circuits and Systems, vol. 5, pp. 117–120 (2001)
7. Drechsler, R., Gunther, W., Somenzi, F.: Using lower bounds during dynamic BDD minimization. IEEE Trans. on CAD 20(1), 51–57 (2001)
8. Hung, W.N., Song, X., Aboulhamid, E.M., Driscoll, M.A.: BDD minimization by scatter search. IEEE Trans. on CAD 21(8), 974–979 (2002)
9. Suguo, D., Yan, S.: A novel ordering method of binary decision diagram. In: International Conference on Management Science and Engineering, pp. 299–304 (2007)
10. Ebendt, R., Drechsler, R.: Approximate BDD minimization by weighted A*. In: IEEE International Symposium on Circuits and Systems, pp. 2974–2977 (2009)

Structure Analysis on High Frequency Conversion Circuit of Switching Power Supply

Xuemei Hu* and Yuanyuan Liu

Opto-electronic Engineering Department, Henan Polytechnic Institute, Nanyang Henan Province, P.R. China 473000
guxiaoyuemei@sina.com

Abstract. Switching power supply can keep the output voltage constant by controlling the on-off time ratio of the switch device by the modern power electronics, it is widely used in the electronic devices such as computer, camera, VCD and so on. Firstly, the article introduces characteristics of the switching power supply, presents that the core of switching power supply circuit is the high frequency conversion. Secondly, it describes principle and application of four types of high frequency conversion circuits. And then contrast analysis is given on four types of structure. Finally, it provides solid theory foundation for structure selection on conversion circuit.

Keywords: Structure, Switching Power Supply, High Frequency, Conversion Circuit.

1 Introduction

Switching power supply is called switching constant voltage power supply, it is named because the components that stabilize voltage work in a switching mode all the time. Switching power supply is the type of power which keeps the output voltage constant by controlling the time ratio of the switching device on-and-off by the modern power electronics. Its good characteristic is smaller volume, lighter weight, lower wastage, higher efficiency, smaller ripple, lower noise, higher intelligence, and it is easier to enlarge the capability, so it is widely used in the electronic devices such as computer, color TV set, camera, VCD, electronic games and so on. Along with the development of power electronics, especially the rapid development of high power MOS technology, the frequency of the switching power supply has been increased to 150-200 kHz, the switching power supply has the high stability and performance-to-price ratio, so the linear controlling manostat power supply of frequency transformer will be replaced by switching power supply day by day.

The key part is the high frequency conversion and controller of the switching power supply in the switching power supply circuit. But the controller of integrated switching power supply has been applied generally, the controlling mode on the switching component is decided by the circuit structure of high frequency conversion.

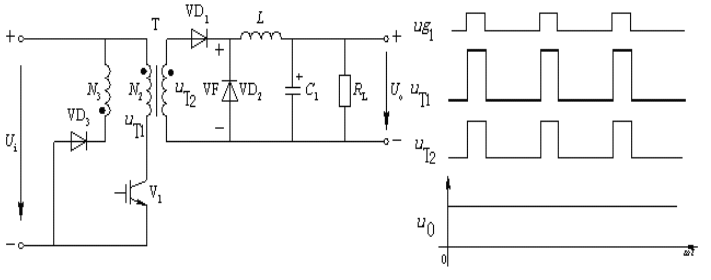
* Corresponding author.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 181–186.
springerlink.com

Therefore we thought that the core of switching power supply circuit is the high frequency conversion circuit, that is, the DC/DC conversion. There are five styles of fundamental modes on high frequency conversion in the input-output-isolation switching power: the single end forward, the single end fly-back, the semi-bridge, the push-pull and the full-bridge. The structure and the principle of the high frequency conversion circuit will be analyzed separately in the following.

2 The Single End Forward Conversion Circuit

TEFCC (the structure of single end forward conversion circuit) is shown as Fig. 1 (a). It is also called the series conversion circuit because its stored energy component series with the load resistance R_L. The DC voltage U_i in the circuit is obtained when frequency AC power supply pass through the power supply filter and rectifier filter. Power switch-tube V_1 is insulated gate bipolar transistor IGBT or MOSFET; T is the high-frequency transformer; L and C_1 compose the LC filter; Diode VD_1 is the half-wave rectification component; VD_2 is the flow-wheeling diode; R_L is the load resistance; U_o is the DC voltage of constant output. Driving signal ug_1 of V_1 is the square-wave whose output is controlled by PWM circuit. Various voltage waveforms are shown as Fig. 1 (b).



(a) schematic diagram (b) waveforms

Fig. 1. Forward conversion circuit

When ug_1 is the high level which controls V_1 on, the transformer obtains the input voltage, that is $u_{T1}=U_i$, the diode VD_1 is on, VD_2 is off. At this time the power supply transports energy to the load by the transformer coupling, the load obtains the voltage, the filter inductance L stores energy. When the controlling circuit causes V_1 off, voltage that switch-tube V_1 withstands is equal to the input voltage, that is $u_{V1}=U_i$, output voltage of the transformer in the primary and the vice side is zero. Now, the stored energy of transformer in the primary side when V_1 is on is returned to the power supply by coil N_3 and diode VD_3. But in the vice side of transformer the output voltage is zero, therefore diode VD_1 is off, L emits energy to load by the diode VD_2, because of the filter effect of capacity C_1, at this time the obtained voltage in the load keeps invariable, its output voltage is

$$U_o = (N_2/N_1)DU_i = kDU_i \tag{1}$$

in the equation (1) k is transformation ratio of transformer, D is the duty cycle of square-wave, N_1 and N_2 is the number of windings of the primary and the vice side of the transformer. It can be seen from the above formula that output voltage U_o is only decided by U_i and D.

The structure of the forward conversion circuit is quite simple and easy to realize. It is suitable for switching power supply from several watts to several hundred watts.

3 The Single End Fly-Back Conversion Circuit

TSEFBCC (the single end fly-back conversion circuit)is shown as Fig. 2, its high-frequency transformer T acts as isolation and current limitation like a inductance L, the stored energy component L is parallel with load R_L, therefore it is also called the parallel conversion circuit. Simultaneously, different from the forward conversion circuit, the core of transformer works in the other side of the magnetic hysteresis loop, thus it is also called the fly-back conversion circuit.

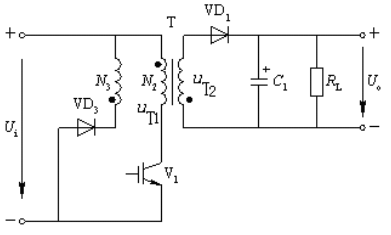


Fig. 2. Fly-back conversion circuit

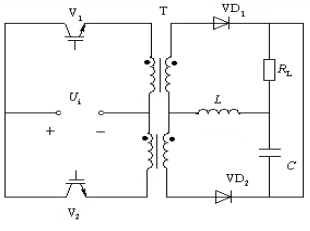


Fig. 3. The Push-Pull converter circuit

When the control circuit causes V_1 on, VD_1 does not become on at the same time as a result of dot convention diode. When V_1 cuts off, induced electromotive force by the vice side of transformer reverses, which causes VD_1 on, the power supply charges the capacitor, producing the voltage on the load R_L simultaneously. In this circuit, the control of the base and design of the vice side winding should follow the fly-back principle.

Similarly, both the voltage endurance of V1 and the input voltage of transformer are equal to the input voltage of power supply, therefore it is the same as forward conversion circuit, the structure is quite simple, easy to realize, which is suitable for switching power supply from several watts to several hundred watts.

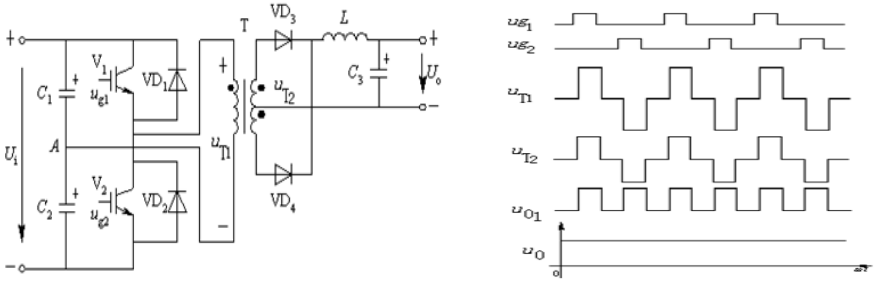
4 The Push-Pull Conversion Circuit

In fact, TPPCC (the push-pull conversion circuit) works in a push-pull way, which combines two single ends forward conversion circuits, the circuit is shown as Fig. 3, two power switching component V_1 and V_2 conduct alternately. Its working process is: when V_1 conducts but V_2 cuts off, it can be conclude that only when VD_2 conducts, current can give R_L power supply by L; when V_1 cuts off while V_2 conducts, only VD_1 can conduct, and current gives power supply to the R_L by L, therefore the current of R_L is continual.

Similarly, it can be seen that both voltage withstand of the switch-tube and the input voltage of transformer are equal to the input voltage of power supply, the core of transformer works in the both sides of magnetic hysteresis loop. The structure of the push-pull conversion circuit is relatively quite complex, it requires highly driving circuit, but the output power is bigger, so it is suitable for power supplies from several hectowatt to several kilowatt switching. Therefore, this type of conversion circuit obtains the widespread application.

5 The Semi-bridge Converter Circuit

TSBCC (the schematic diagram of the semi-bridge converter circuit) is showen as Fig. 4 (a), the waveforms of each spot output voltage are shown as Fig. 4 (b). Input voltage U_i is transformed into the double power supply through two equal input capacities (C_1,C_2), $U_{C1}=U_{C2}=U_i/2$, that is, A's voltage U_A is the half of input voltage U_i. The driving signal ug_1 and ug_2 of switch-tube V_1, V_2 are produced by the control circuit, they are mutually for opposition PWM signal. In order to prevent the power supply from short-circuit when switch-tube V_1, V_2 conduct, there must be a certain dead time between driving signal ug_1 and ug_2, that is the time when the two driving signals are zero.



(a) circuit diagram (b) waveforms

Fig. 4. The Semi-Bridge conversion circuit

When ug_1 is the high level while ug_2 is the low level, V_1 conducts, V_2 cuts off. The voltage of capacity C_1 is exerted by V_1 to the primary side of high frequency transformer, now $u_{T1}=U_i/2$, diode VD_3 conducts, VD_4 cuts off, the rectified output voltage U_{o1} is in the same direction with the graphical representation U_o, then U_{o1} passes through the LC_3 filter to obtain output voltage U_o again. When ug_2 is the high level while ug_1 is the low level, V_2 conducts, V_1 shuts off. The voltage of capacity C_2 is exerted to the primary side of high frequency transformer, now $u_{T1}=-U_i/2$. The diode VD_4 conducts, VD_3 cuts off, the rectified output voltage U_{o1} is also in the same direction with the graphical representation U_o. When V_1 and V_2 cut off simultaneously, the voltage of the transformer's primary side and vice side is zero, that is $u_{T1}=0$, $u_{T2}=0$. When VD_3 and VD_4 are conducted simultaneously, inductance L starts to store energy. When V_1 and V_2 are cut off simultaneously,

although the voltage on the transformer vice side is zero, L can also emit energy, the output voltage can constant invariable because of capacity C_3.

The semi-bridge converter circuit is different from the forward and fly-back converter circuit. In a switching period T, the current that passes the high frequency transformer in the first half is equal to the second half in an opposite direction. Therefore, compared with the first two kind of circuits, the core of transformer is in the both sides of magnetic hysteresis loop B-H, the core obtains the full use and is prevent from the magnetic saturation, therefore the high frequency transformer may be designed slightly but with bigger power. When a switch-tube conducts, the other switch-tube which is in the cut-off state withstands the same voltage with the input voltage. The moment when switch-tube turns conduction into shutdown, the peak voltage that is caused by the leakage inductance will influence V_1 and V_2. Therefore each side of V_1 and V_2 has diode VD_1, VD_2 in parallel, which will clamp the peak voltage by leakage inductance, so voltage that switch-tube withstands will not surpass the input voltage absolutely, VD_1, VD_2 will also have the flow function as flow-wheeling diode, but voltage exerted on high frequency transformer is only the half of input voltage.

The structure of the semi-bridge converter circuit is relatively complex, it has high requirement to driving circuit, but the output power is bigger, which is suitable for the switching power supplies from several hectowatt to several kilowatt.

6 The Full-Bridge Conversion Circuit

Other two switch-tube V_3, V_4 are used to replace the two capacity C_1 and the C_2 in the semi-bridge conversion circuit, which equipped with corresponded driving circuit can be composed TFBCC (the full-bridge conversion circuit), that is shown as Fig.5. The rectification diode VD_5 and VD_6 in the vice side of transformer can realize the full-wave rectification. The driving signal ug_1 of V_1 is the same as ug_4 of V_4, the driving signal ug_2 of V_2 is the same as ug_3 of V_3, but both ug_1, ug_4 and ug_2, ug_3 are opposite mutually.

When ug_1 and ug_4 are the low levels, ug_2 and ug_3 are the high levels, V_2 and V_3 conduct, V_1 and V_4 cut off, the supply voltage is exerted in the primary side of high-frequency transformer by V_2 and V_3, at this time $u_{T1}=U_i$. When ug_1 and ug_4 are the high levels, ug_2 and ug_3 are the low levels, V_1 and V_4 conduct, V_2 and V_3 cut off, $u_{T1}=-U_i$. Comparing semi-bridge circuit, the voltage on the primary side increases one time, but the voltage withstand of each switch-tube is same still with the input voltage.

Collector and the emitter of switch-tube V_1, V_2, V_3 and V_4 clamps with clamping diodes VD_1, VD_2, VD_3 and VD_4. As a result, when switch-tube works from conduction to closure, the maximum of peak voltage caused by the transformer primary magnetized current's energy as well as the stored energy of leakage inductance will not surpass supply voltage U_i, meanwhile the energy of magnetized current can be feedback to the power supply, thus efficient of the complete machine can be enhanced.

Although the the full-bridge conversion circuit has adopted four switch-tubes, the structure is complex, and requires driving circuit highly, the comprehensive performance is the best. Under the equal situation of input voltage U_i, it may put out a bigger

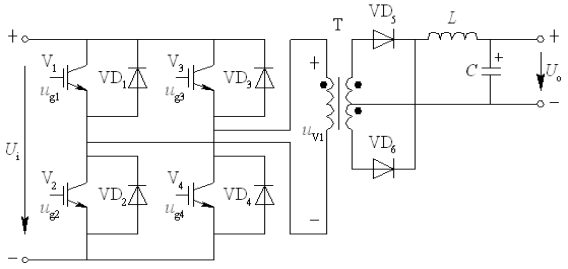


Fig. 5. The Full-Bridge conversion circuit

Table 1. Contrast analysis and conclusion

Circuit Type	Input Voltage	Efficiency of Transformer	Requirement of Driving	Synthesizes Performance
TEFCC	Ui	Low	Low	Ordinary
TSEFBCC	Ui	Low	Low	Ordinary
TPPCC	Ui	Low	High	Good
TSBCC	Ui／2	High	High	Good
TFBCC	Ui	High	High	Good

power if compared with the semi-bridge conversion circuit, thus it is suitable for the switching power supplies from number hectowatt to several kilowatt and even larger.

7 Conclusion

The performance comparison of the above five circuits under the same situation of input voltage U_i is shown as the following table 1. When selecting the conversion circuit of switching power supply, various factors must be considered according to the actual requirements.

References

1. Meng, Y., Wang, K., Pan, J.: A DC/DC Switching Power Supply Based on Push-Pull Forward Conversion Technique. Low Voltage Apparatus 17, 57–60 (2007)
2. Dong, J., Chen, W.: Modeling and Analysis of Capacitive Effects in High-frequency Transformer of SMPS. Proceedings of the CSEE 27, 121–126 (2007)
3. Sun, H., Jiang, H., Cui, L.: Design of high frequency transformer in fly-back. Switching Power Supply Technicques 10, 66–68 (2007)
4. Guo, Y., Ma, Y.: Design of fly-back and two output switch power. Journal of Shaanxi University of Technology 24, 18–21 (2008)
5. Sun, X., Li, G.: The Transformer of the High-Frequency Switching Power Supply. Techniques of Automation and Applications 27, 52–55 (2007)
6. Xu, C., Wei, Z., Ye, L.: Analysis of Output Voltage Ripple for Boost Conversion in Switching Mode Power Supply. Telecom Power Technology 27, 15–18 (2010)

Principle and Application Analysis on Soft-Switching Circuits

Xuemei Hu* and Han Lian

Opto-Electronic Engineering Department, Henan Polytechnic Institute, Nanyang Henan Province, P.R. China 473000
guxiaoyuemei@sina.com

Abstract. The soft-switching technology is widely used on power electronic technology device for decreasing circuit switch wastages and noise effectively by controlling the ON and OFF of the switch device under zero voltage(zero current) condition. Firstly, this paper analyzes the basic principle of soft-switching circuits in details, emphasizes on some kinds of circuits. Finally, necessary analysis of the characteristics on the key circuits is given.

Keywords: Principle, Analysis, Soft-switching, PWM Control.

1 Introduction

In recent years, with the rapid development of power electronic technology, power electronic device, which is widely used in computer, aviation, communication and so on, is required to be small, light and effective, consequently, its efficiency and electromagnetic compatibility (EMC) is more highly required. In the power electronic devices, the bulk and the weight of the magnetic components of the filter capacitance and inductance and transformer is in a high proportion. It can be learned from "Circuit" and "Electrical engineering" that increasing the frequency of the switch can reduce the parameters of the filter and make the voltage transformer smaller, thereby the bulk and the weight of the device can be reduced effectively. So to increase the frequency of the circuit is the best way to make the device smaller and lighter.

Thus it becomes the goal of power electronics. However, traditional switch devices work in a hard condition, when the frequency of the switch devices is increased, the wastage and the EMI are increased severely along with it, and the efficiency of the circuit is reduced, so it is not feasible to increase the frequency of the switch simply. At present, adopting the soft switch technology is a common way to solve the problem.

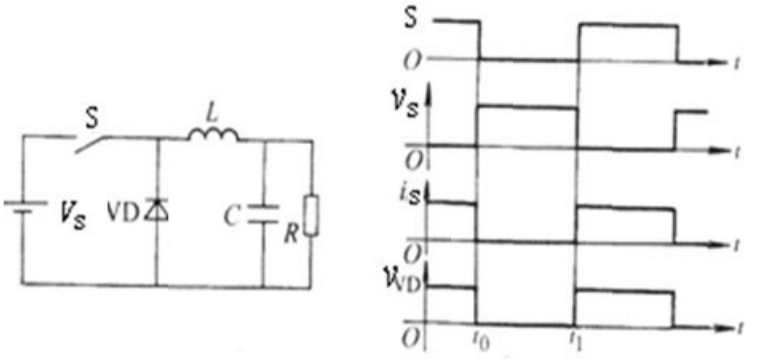
2 Working Principle of Soft-Switching

The name "soft-switching" is relative to "hard switch". Hard switch produces large wastage and noise when controlling the on and off of the circuit because of the acute

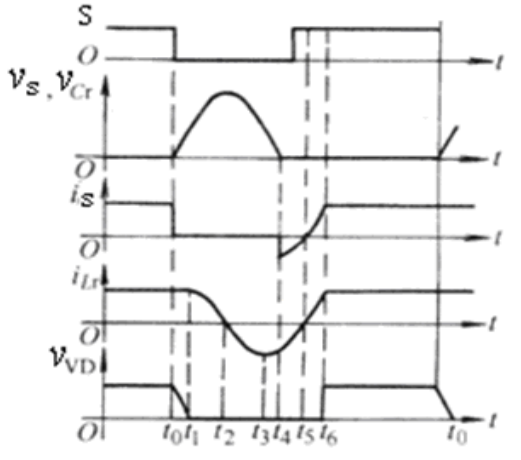
* Corresponding author.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 187–192.
springerlink.com

change of the voltage and current, the wastages increases with the increase of the frequency and decreases the efficiency of the circuit; the noise interferes the circuit seriously and bothers the work the other devices around it. Fig. 1 (a) shows the hard-switching circuit and its wave. Fig. 1(b) shows the soft-switching circuit and its wave.



(a) hard-switching

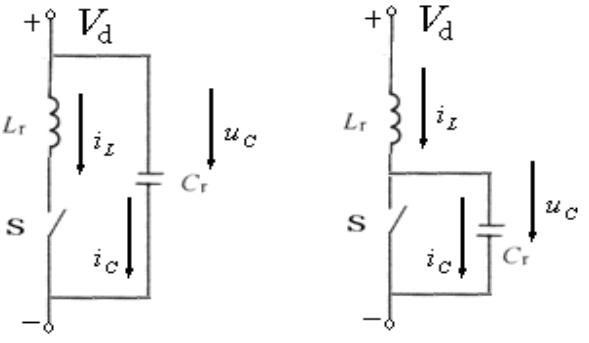


(b) soft-switching

Fig. 1. Comparison of circuit and wave

Soft-switching adds resonant components such as small electrical inductance and capacitance etc.to a hard-switching circuit, it makes up a network of auxiliary change, because of the resonant introducing during the switching, the switch becomes ON when the both ends of the voltage is zero, or it becomes OFF when the current is zero, the switching condition is improved, so the wastage and noise of traditional hard-switching switch are reduced, thus the efficiency of the circuit is increased. Because the ON and OFF of the switch depends on the resonant circuit, soft-switching is also called the resonant soft-switching.

Soft-switching includes the process of soft ON switching and soft OFF switching. The ideal process of soft ON switching is that after the voltage decreases to zero, the current increases slowly to a ON state value, so wastage and noise will not be produced when the switch is on, soft ON switching is called zero-voltage-switching (ZVS). The ideal process of soft OFF switching is that after the current decreases to zero, the voltage increases to a OFF state value, so wastage and noise will not be produced when the switch is off, soft OFF switching is called zero-current-switching (ZCS).



(a) ZCS (b) ZVS

Fig. 2. ZCS and ZVS

Fig. 2 shows the simplified equivalent circuit of Zero-Current-Switching (ZCS) and Zero-Voltage-Switching (ZVS). Fig. 2 (a) is ZCS circuit, and Fig. 2 (b) is ZVS circuit. *Lr* is the resonant inductance and *Cr* is the resonant capacitance. To make the soft-switching work, it is crucial that the related current and voltage comes to zero when the switch transformations. In the ZCS, the current that flows through the switch (S) is the current flowing through the inductance (*Lr*), the i_L in the Fig. 2, it can be expressed as

$$i_L(t) = -C\,du/dt = V_d \sin \omega_0 t / Z_0 \tag{1}$$

While in the zero-voltage-switching (ZVS) circuit, the voltage of the switch S is the voltage (u_C) of the capacitance (*Cr*), it can be expressed as

$$u_C(t) = V_d(1-\cos \omega_0 t) + I_{L0} Z_0 \sin \omega_0 t \tag{2}$$

equation in 2, $\omega_0 = 1/\sqrt{LC}$, $Z_0 = \sqrt{L/C}$.

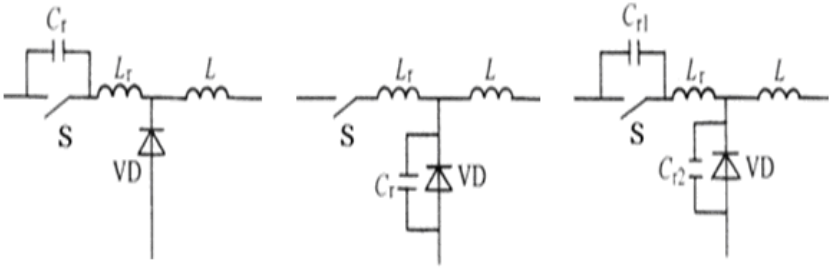
3 Classification and Several Kind of Model on Soft-Switching Circuit

Soft-switching technology has experienced constant development and improvement since the advent. There has been many kinds of soft-switching circuits, and now the

new soft-switching circuits are being appearing unceasingly. They can be divided into three kinds according to the development, the first kind is quasi-resonant circuit represented by resonant technology, it adopts frequency conversion control (PFM) generally; the second kind is zero switching circuit; and the third one is zero conversing circuit, the second and the third kinds adopt PWM commonly.

3.1 Quasi-resonant Circuit

Wave of the voltage or the electric current in the quasi-resonant circuit are the sine half-waves, therefore it is called the quasi-resonant circuit. Quasi-resonant circuit can be divided into ZVS quasi-resonant circuit, ZCS quasi-resonant circuit, ZVS multi-resonant circuit. Fig. 3 shows the basic units of quasi-resonant circuit. Fig. 1 (b) shows the ZVS quasi-resonant circuit and its working wave.



(a) ZVS (b) ZCS (c) ZVS multi—resonant

Fig. 3. Basic switching units of quasi-resonant circuit

The increased resonant inductor *Lr* and resonant capacitor *Cr* in the circuit are much smaller than the filter capacitor *C* and the filter inductance *L*. When the soft-switching circuit S is shut off, the resonant inductor *Lr* resonates with resonant capacitor *Cr*, waves of voltage or current in the circuit are similar to the sine half-wave. The voltage across the switch S has been reduced to zero before the opening, so Fig. 1 (b) shows ZVS quasi-resonant circuit. Fig. 1 (b) shows that switching wastage and switching noise have been falling after the introduction of resonant, but also some negative issues are occurred. Resonant voltage peak is higher, so voltage press of power electronic devices must be enhanced. Valid values of resonant current is large, there is a lot of exchange of reactive power in the circuit, then on wastage is increased, its resonant cycle is changing with the input voltage and load.

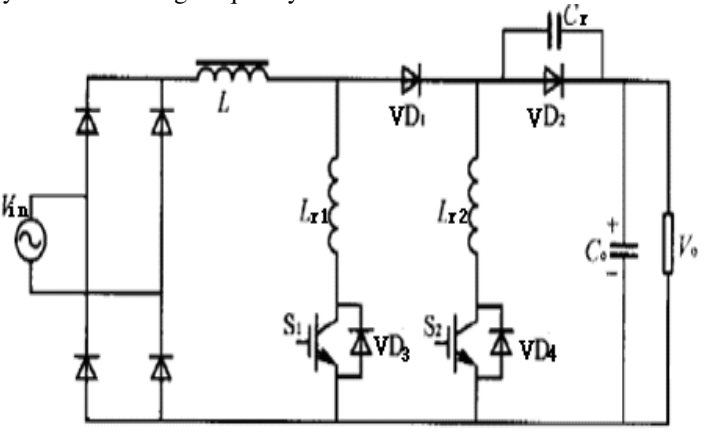
3.2 Zero Switching PWM Circuit

Auxiliary switch is introduced to control the time of resonance, it makes the resonance only start in the process of being ON and OFF, zero switching PWM circuit can be divided into zero-voltage-switching PWM circuit (ZVS-PWM) and zero-current-switching circuit PWM(ZCS-PWM).

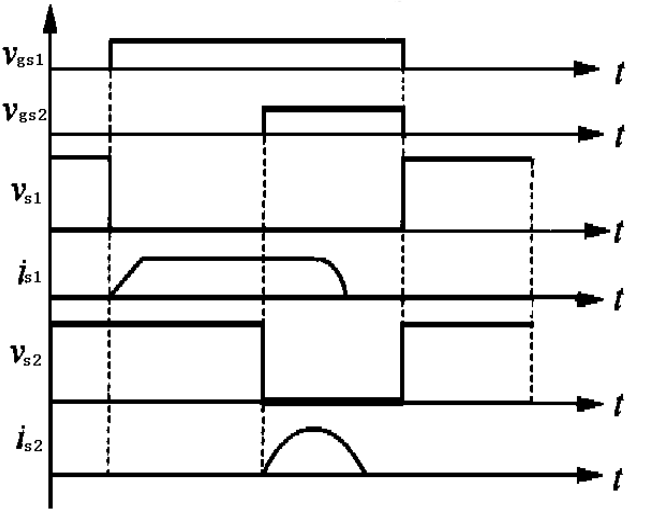
Fig. 4 hows the circuit and waves, auxiliary switch S_2 is introduced to make the main switch S_1 OFF in zero current condition. The main switch S_1 is ON firstly, the

current through the switch S_1 increases gradually to the input current, now both diode VD_1 and VD_2 are shut off, capacitance Cr is charged reversely to Vo. After the auxiliary switch S_2 switch on, *Cr* resonates with *Lr*, when the voltage in the *Cr*'s ends reduces to zero, the diode VD_1 is on, and the capacitance *Cr*, the inductance Lr_1 and Lr_2 resonate till the diodes VD_3,VD_4 switch on, then S_1,S_2 switch with a ZCS mode. This circuit makes S_1,S_2 switch with a ZCS mode, but the two switches are hard switching, the resonance circuit of *Cr* and Lr_2 and the circuit of *Cr* and Lr_1, Lr_2 must get through the end of output, the wave of voltage can be increased.

It can be seen from Fig. 4 that, zero switch circuit has many advantages compared with resonant circuit. Voltage and current is basically square-wave, rising and falling slowly, the voltage that the switch stands is significantly smaller, its circuit can be controlled by fixed switching frequency PWM.



(a) circuit



(b) waves

Fig. 4. Circuit and waves of ZCS PWM

3.3 Zero Conversing PWM Circuit

Auxiliary switch circuit is introduced to control the start time of resonance, the resonance circuit is parallel to the main switch, thus the input voltage and the current influence the resonance little, the circuit is in a broad range and it can stay in a soft working from zero load to full. And the exchange of wattless power in the circuit is reduced to the minimum; it enhances the efficiency of circuit further. This circuit can be divided into ZVT-PWM and ZCT-PWM.

4 Conclusion

Soft-switching technology has been widely used in power electronic devices, adopting soft-switching technology, switch frequency is not restricted by switch wastage, frequency of switching device is improved, the size of the cooling device is reduced and the working reliability and efficiency of switching devices is increased. Phase-shifted Full-bridge ZVS pulse width soft-switching technology has been adopted in intelligent high-frequency switch power supply charging module of ATC series in Guangzhou metro, the switch component is ZVS and ZCS switch, it is ideal soft-switching because there is no voltage and current overshoot or rush.

Compared with the hard-switching, wastage of soft-switching charging module is reduced by 40%, the efficiency is increased to 94% ~ 96%, the pressure of power switching device is smaller, so the reliability of electric power switching device is improved. Because of the reduction of the voltage variation (dv/dt) and the current rate (di/dt), the interference of battery module of electromagnetic is reduced significantly, so the EMC performance is improved.

Due to its advantages, soft-switching technology has been a important means to reduce switch wastage, improve the system efficiency, ameliorate the electromagnetic interference and improve the system reliability, and it has become one of the hotspots in the study of power electronic technology and it is receiving more and more attention.

References

1. Liu, Z.: Journal of Lanzhou University, Lanzhou 42, 115–119 (2006)
2. Hu, H.: Power Electronics 43, 12–14 (2009)
3. Wang, T.: Power Electronics, Beijing, pp. 195–199 (April)
4. Wu, W.: Transactions of China Electrotechnical Society 23, 47–51 (2008)
5. Wang, Q.: Proceedings of the CSEE 28, 15–21 (2008)
6. Yang, Y.: Proceedings of the CSEE 28, 50–54 (2008)
7. He, R., Wei, W.: Research on a New Natural Soft-switching for Variable Current Technology. Power Electronics 42, 1–3 (2008)
8. Yu, L.: On the Soft-switch Technology and Its Application. Shandong Textile Economy 4, 100–102 (2010)

The Characterization of the Trivariate Super-Wavelet Wraps and Applications in Electronic Engineering

Hongwei Gao[1,*] and Jun Lu[2]

[1] Yulin University, Yulin 719000, China
[2] Zhengzhou Normal University, Zhengzhou 450044
{sxxa66zxcv,jhnsx123}@126.com

Abstract. Mechanical engineering is a discipline of engineering that applies the principles of physics and materials science for analysis, design, manufacturing, and maintenance of mechanical systems. It is the branch of engineering that involves the production and usage of heat and mechanical power for the design, production, and operation of machines and tools. An approach for designing a sort of orthogonal super-wavelet wraps in three-dimensional space is presented and their orthogonality traits are characterized by virtue of iteration method and time-frequency representation method. The orthogonality formulas concerning these super-wavelet wraps are established. Moreover, it is shown how to draw new Riesz bases of space $L^2(R^3)$ from these wavelet wraps. The trivariate dual frames is also discussed.

Keywords: wavelet transferm, trivariarte super-wavelet wraps, tight wavelet frames, iteration method, time-frequency analysis approach, semi-orthogonality.

1 Introduction

Although the Fourier transfer has been a major tool in analysis for over a century, it has a serious lacking for signal analysis in that it hides in its phases information concerning the moment of emission and duration of a signal. What was needed was a localed time-frequency representation which has this information encoded in it. Transform and Gabor Transform were used for harmonic studies of nonstationary power system waveforms which are basically Fourier Transform-based methods. Wavelet analysis has become a developing branch of mathematics for over twenty years. The main feature of the wavelet transform is to hierarchically decompose general functions, as a signal or a process, into a set of approximation functions with different scales. The last two decades or so have witnessed the development of wavelet theory[1]. Wavelet packets, owing to their nice traits, have attracted considerable attention. They can be widely applied in science and engineering [4,5]. Coifman R. R. and Meyer Y. firstly introduced the notion for orthogonal wavelet packs which were used to decompose wavelet components. Chui C K. and Li C.[6] generalized the concept of orthogonal wavelet wraps to the case of non-orthogonal wavelet wraps so

* Corresponding author.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 193–199.
springerlink.com

that wavelet wraps can be employed in the case of the spline wavelets and so on. The intro-duction for biorthogonal wavelet wraps attributes to Cohen and Daubechies. Tensor product multivariate wavelet wraps has been constructed by Coifman and Meyer. The introduction for the notion on nontensor product wavelet wraps attributes to Shen [7]. Since the majority of information is multidimensional information, many researchers interest themselves in the investigation into multivariate wavelet theory. But, there exist a lot of obvious defects in this method, such as, scarcity of designing freedom. Therefore, it is significant to investigate nonseparable multivariate wavelet theory. Yang [8] constructed a-scale orthogonal multiwavelet wraps that were more flexible in applications. It is known that the majority of information is multi-dimensional information. Thus, it is necessary to generalize the concept of multivariate wavelet wraps to the case of quarternary nonseparable vector-valued wavelets. The goal of this paper is to give the definition and the constructing procedure of orthogonal trivariate super-wavelet wraps and characterize their properties.

2 The Three-Dimensional Multiresolution Analysis

We begin from the following notations. Z and N stand for integers and nonnegative integers, respectively. Let R be the set of all real numbers. R^3 denotes the 3- dimensional *Euclidean* space. By $\mathrm{L}^2(\mathrm{R}^3)$, we denote the square integrable function space on R^3. Set $x=(x_1,x_2,x_3)\in \mathrm{R}^3$, $k=(k_1,k_2,k_3)$, $\omega=(\omega_1,\omega_2,\omega_3)$, $z_l=e^{-i\omega_l/2}$, where $l=1,2,3;$ *and* $a\geq 2, a\in Z$. The inner product for any $\hat{\lambda}(t)$, $\hbar(t)\in \mathrm{L}^2(\mathrm{R}^3)$ and the Fourier transform of $\hat{\lambda}(t)$ are defined as, respectively

$$\langle \hat{\lambda}, \hbar\rangle = \int_{R^3} \hat{\lambda}(t)\,\overline{\hbar(t)}\,dt, \quad \hat{\hat{\lambda}}(\omega)=\int_{R^3}\hat{\lambda}(t)\,e^{-it\cdot\omega}dt,$$

where $\omega\cdot t=\omega_1 t_1+\omega_2 t_2+\omega_3 t_3$, and $\overline{\hbar(t)}$ denotes the conjugate. Denoting $a=|\det(A)|$, the -re exist a elements $\nu_0,\nu_1,\cdots,\nu_a$ in Z_+^3 by algebra theory such that $Z^3=\bigcup_{\nu\in\Gamma_0}(\nu+AZ^3)$; $(\nu_1+AZ^3)\bigcap(\nu_2+AZ^3)=\varnothing$, where $\Gamma_0=\{\nu_0,\nu_1,\nu_2,\cdots,\nu_a\}$ denotes the set of all different representative elements in the quotient group $Z^3/(AZ^3)$ and ν_1,ν_2 denote two arbitrary distinct elements in Γ_0. Let A be a 3×3 integer matrix. Set $\nu_0=\underline{0}$, where $\underline{0}$ is the null of Z_+^3. Let $\Gamma=\Gamma_0-\{\underline{0}\}$ and Γ,Γ_0 be two index sets. By $L^2(R^3,C^v)$ we denote the set of all vector- valued functions $H(x)$, that is, $L^2(R^3,C^v):=\{H(x)=\quad (h_1(x),h_2(x)\cdots,h_v(x))^T:\ h_l(x)\in L^2(R^3),\ l=1,2,\cdots v\}$, where T means the transpose of a vector. Video images and digital films are examples of vector-valued functions where $h_l(x)$ in the above $H(x)$ denotes the pixel on the l th column at the point x.

Definition 1. We say that a vector-valued functions $\hat{\lambda}(x)\in L^2(R^3,C^v)$ are orthogonal, if the following equality is satisfied

$$[\hat{\lambda}(\cdot),\hat{\lambda}(\cdot-k)]=\delta_{\underline{0},k}I_v,\quad k\in Z^3, \tag{1}$$

where I_v denotes the $v\times v$ indentity matrix and $\delta_{\underline{0},k}$ is the Kronecker symbol.

Definition 2. A sequence of vector-valued functions $\{\mathrm{T}_n(x)\}_{n\in Z^3} \subset U \subset L^2(R^3, C^v)$ is called a Riesz basis of U if it satisfies *(i)* for any $G(x)\in U$, there exists a unique $v\times v$ matrix sequence $\{P_n\}_{n\in Z^3} \in \ell^2(Z^3)^{v\times v}$ such that

$$G(x)=\sum_{n\in Z^3} P_n T_n(x),\ x\in R^3, \tag{2}$$

where $^2(Z^3)^{v\times v}=\{P: Z^3 \to C^{u\times u}, \|P\|_2 = \sum_{l,s=1}^{v}\sum_{n\in Z^4} |q_{l,s}(n)|^2)^{\frac{1}{2}} < +\infty\}$, *(ii)* there exist two constants $0<C_1\le C_2<+\infty$ such that, for any matrix sequence $\{M_n\}_{n\in Z^3}$, the following equality follows. i.e.,

$$C_1\|\{M_n\}\|_* \le \| \sum_{n\in Z^3} M_n T_n(x) \| \le C_2 \|\{M_n\}\|_*, \tag{3}$$

Definition 3. Let V be a shift invariant subspace of space $L^2(R^3)$. A subset $\Psi=\{\psi_1,\psi_2,\cdots,\psi_a\}$ is called a normalized tight frame generator for V if

$$\{\psi(x-k): k\in Z, \psi\in\Psi\}$$

is a normalized tight frame generator for the subspace V.

Definition 4. A finite subset $\Psi=\{\psi_1,\psi_2,\cdots,\psi_a\}$ of $L^2(R^3)$ is normalized tight frame wavelet (or super-wavelets) if the affine system

$$X(\Psi):=\{\psi_{j,k}:=\sqrt{a}\psi(ax-k) \mid j\in Z, k\in Z, \psi\in\Psi\}$$

is called a normalized tight frame generator for V if

$$\{\psi(x-k): k\in Z, \psi\in\Psi\}$$

is a normalized tight frame generator for the space $L^2(R^3)$.

The multiresolution analysis method is an important approach to obtaining wavelets and wavelet packs. We introduce the notion of multiresolution analysis of $\mathrm{L}^2(\mathrm{R}^3)$. Let $\hbar(x)\in \mathrm{L}^2(\mathrm{R}^3)$ satisfy the following refinement equation:

$$\hbar(x)=\sqrt{a}\cdot\sum_{u\in Z^3} d(u)\hbar(Ax-u) \tag{4}$$

where $\{d(u)\}_{u\in Z^3}$ is a real number sequence and $\hbar(x)$ is called a scaling function. Taking the Fourier transform for the both sides of refinement equation (1), we have

$$\hat{\hbar}(A^T\omega)=D(\omega)\hat{\hbar}(\omega)\ . \tag{5}$$

Define a subspace $U_l\subset \mathrm{L}^2(\mathrm{R}^3)$ $(l\in Z)$ by

$$U_l = clos_{L^2(R^3)}\left\langle a^{l/2}\hbar(A^l x-k): k\in Z^3\right\rangle, \tag{6}$$

We say that $\hbar(x)$ in (1) generates a multiresolution analysis $\{U_j\}_{j\in Z}$ of $\mathrm{L}^2(\mathrm{R}^3)$, if the sequence $\{U_l\}_{l\in Z}$, defined in (3) satisfies the below: (a) $U_l \subset U_{l+1}, \forall\, l\in Z$; (b) $\bigcap_{l\in Z} U_l = \{0\}$; $\bigcup_{l\in Z} U_l$ is dense in $\mathrm{L}^2(\mathrm{R}^3)$; (c) $f(x)\in U_l \Leftrightarrow f(Ax)\in U_{l+1}$, $\forall\, l\in Z$; (d) the family $\{|\det(A)|^{l/2}\,\hbar(A^l x-k): k\in Z^3\}$ is a Riesz basis for U_l.

Let W_l $(l\in Z)$ denote the orthogonal complementary subspace of U_l in U_{l+1} and order $A=2I_3$, also assume that there exists a vector-valued function $\Lambda(x)=(\lambda_1(x),\lambda_2(x),\cdots,\lambda_7(x))^T$ (see [7]) forms a Riesz basis for W_l, i.e,

$$W_l = clos_{L^2(R^3)}\left\langle \Lambda_\nu(2^l\cdot -k):\ \nu=1,2,\cdots,7;\ \ k\in Z^3\right\rangle,\ \ l\in Z \tag{7}$$

From (4), it is clear that $\Lambda_1(x), \Lambda_2(x)\ ,\cdots,\Lambda_7(x)\in W_0\subset U_1$. Therefore, there exist seven real sequences $\{b^{(\iota)}(u)\}$ $(\iota=1,2,\cdots 7,\ \ u\in Z^3)$ such that

$$\Lambda_\iota(x)=2\sqrt{2}\cdot\sum\nolimits_{u\in Z^3} b^{(\iota)}(u)\,\hbar\left(2x-u\right),\ \iota\in\Gamma, u\in Z^3. \tag{8}$$

We say that $\Lambda_\iota(x)\in L^2(R^3,C^\nu)$, $\mu\in\Gamma$ are orthogonal vector-valued wavelets associated with an orthogonal vector-valued scaling functions $\hbar(x)$, if the family $\{\Lambda_\mu(x-k),\ k\in Z^3,\mu\in\Gamma\}$ is a Riesz basis of subspace U_0, and

$$[\hbar(\cdot),\Lambda_\mu(\cdot-k)]=0,\quad \mu\in\Gamma,\ \ k\in Z^3. \tag{9}$$

$$[\Lambda_\iota(\cdot),\Lambda_\mu(\cdot-k)]=0,\quad \iota,\mu\in\Gamma,\ \ k\in Z^3. \tag{10}$$

3 The Features of a Sort of Vector-Valued Trivarate Wavelet Wraps

We begin with the following notations:

$$\Phi_0(x)=\hbar(x),\ \ \Phi_\iota(x)=\Lambda_\iota(x),\ \ d^{(0)}(u)=d(u),\ \ d^{(\iota)}(u)=b^{(\iota)}(u),\ \ \iota\in\Gamma,\ u\in Z^3.$$

Definition 5. A set of vector-valued functions $\{\Phi_{8n+\nu}(x): n=0,1,2,\cdots,\ \nu=0,1,2,\cdots 7\}$ is called a nonseparable vector-valued trivariate wavelet wraps with respect to the orthogonal scaling function $\hat{\lambda}(x)$, where

$$\Phi_{8n+\nu}(x)=\sum\nolimits_{k\in Z^3} d^{(\nu)}(n)\Phi_\nu(2x-k),\ \ \nu\in\Gamma_0=\{0,1,2,\cdots 7\}, \tag{11}$$

By implementing the Fourier transform for the both sides of (14), we have

$$\widehat{\Phi}_{8n+\nu}(2\omega)=D^{(\nu)}(z_1^2,z_2^2,z_3^2)\cdot\widehat{\Phi}_n(\omega).\ \nu\in\Lambda_0. \tag{12}$$

$$D^{(\nu)}\left(z_1,z_2,z_3\right)=D^{(\nu)}(\omega/2)=\sum_{k\in Z^3} d^{(\nu)}(k)\,z_1^{k_1}z_2^{k_2}z_3^{k_3} \tag{13}$$

Lemma 1[7]. Let $\lambda(x) \in L^2(R^3)$. Then $\lambda(x)$ is an orthogonal function if and only if

$$\sum_{u \in Z^3} | \hat{\lambda}(\omega + 2u\pi) |^2 = 1. \tag{14}$$

Lemma 2. Assuming that $\hbar(x)$ is an orthogonal scaling function. $D(z_1, z_2, z_3)$ is the symbol of the sequence $\{d(u)\}$ defined in (3). Then we have

$$\begin{aligned}\Omega &= |D(z_1, z_2, z_3)|^2 + |D(-z_1, -z_2, -z_3)|^2 + |D(-z_1, z_2, z_3)|^2 + |D(z_1, -z_2, z_3)|^2 \\ &+ |D(z_1, z_2, -z_3)|^2 + |D(-z_1, -z_2, z_3)|^2 + |D(-z_1, z_2, -z_3)|^2 + |D(z_1, -z_2, -z_3)|^2 = 1.\end{aligned}$$

Lemma 3. If $\Lambda_\nu(x)$ ($\nu = 0, 1, 2, \cdots, 7$) are orthogonal wavelet functions associated with $\hbar(x)$. Then for $\iota, \nu \in \{0, 1, 2, \cdots, 7\}$, we have

$$\begin{aligned}\Xi_{\iota,\nu} &= \sum_{j=0}^{1} \Big\{ D^{(\iota)}((-1)^j z_1, (-1)^j z_2, (-1)^j z_3) \overline{D^{(\nu)}((-1)^j z_1, (-1)^j z_2, (-1)^j z_3)} \\ &+ D^{(\lambda)}((-1)^{j+1} z_1, (-1)^j z_2, (-1)^j z_3) \cdot \overline{D^{(\nu)}((-1)^{j+1} z_1, (-1)^j z_2, (-1)^j z_3)} \\ &+ D^{(\lambda)}((-1)^j z_1, (-1)^{j+1} z_2, (-1)^j z_3) . \overline{D^{(\nu)}((-1)^j z_1, (-1)^{j+1} z_2, (-1)^j z_3)} \\ &+ D^{(\lambda)}((-1)^j z_1, (-1)^j z_2, (-1)^{j+1} z_3) \overline{D^{(\nu)}((-1)^j z_1, (-1)^j z_2, (-1)^{j+1} z_3)} \Big\} = \delta_{\lambda,\nu}.\end{aligned} \tag{15}$$

Theorem 1. For every $u \in Z^3$ and $m, n \in Z_+$, we have

$$[\Phi_m(\cdot), \Phi_n(\cdot - k)] = \delta_{m,n} \delta_{0,k} I_\nu. \tag{16}$$

Proof. For the case of $m = n$, (20) follows from Theorem 1. As $m \neq n$ and $m, n \in \Gamma_0$, (20) can be established from Theorem 2, where $\Gamma = \{0, 1, 2, \cdots, 7\}$. In what follows, assuming that m is not equal to n and at least one of $\{m, n\}$ do-esn't belong to Γ_0, rewrite m, n as $m = 8m_1 + \lambda_1$, $n = 8n_1 + \mu_1$, where $m_1, n_1 \in Z_+$, and $\lambda_1, \mu_1 \in \Gamma_0$. **Case 1** If $m_1 = n_1$, then $\lambda_1 \neq \mu_1$. By (14), (16) and (18), (20) holds, since

$$\begin{aligned}(2\pi)^3 \langle \Phi_m(\cdot), \Phi_n(\cdot - k) \rangle &= \int_{R^3} \hat{\Phi}_{8m_1+\lambda_1}(\omega) \overline{\hat{\Phi}_{8n_1+\mu_1}(\omega)} \cdot \exp\{ik\omega\} d\omega \\ &= \int_{[0,4\pi]^3} D^{(\lambda_1)}(z_1, z_2, z_3) \sum_{l \in Z^3} \hat{\Phi}_{m_1}(\omega/2 + 2l\pi) \cdot \overline{\hat{\Phi}_{m_1}(\omega/2 + 2l\pi)} \, \overline{D^{(\mu_1)}(z_1, z_2, z_3)} \cdot e^{ik\omega} d\omega \\ &= \int_{[0,4\pi]^3} D^{(\lambda_1)}(z_1, z_2, z_3) \overline{D^{(\mu_1)}(z_1, z_2, z_3)} \cdot e^{ik\omega} d\omega \\ &= \int_{[0,2\pi]^2} \Xi_{\lambda_1,\mu_1} \cdot e^{ik\omega} d\omega = O.\end{aligned}$$

Therefore, we have $\langle \Phi_m(\cdot), \Phi_n(\cdot - k) \rangle = O.$

Case 2. If $m_1 \neq n_1$, order $m_1 = 8m_2 + \rho_2$, $n_1 = 8n_2 + \mu_2$, where $m_2, n_2 \in Z_+^3$, and $\rho_2, \mu_2 \in \Gamma_0$. If $m_2 = n_2$, then $\rho_2 \neq \mu_2$. Similar to Case 1, (16) follows. As $m_2 \neq n_2$, order $m_2 = 8m_3 + \rho_3$, $n_2 = 8n_3 + \mu_3$, where $\alpha_3, \sigma_3 \in Z_+^3$, $\rho_3, \mu_3 \in \Gamma_0$. Thus, taking finite steps (denoted by κ), we obtain $m_\kappa, n_\kappa \in \Gamma_0$, and $\rho_\kappa, \mu_\kappa \in \Gamma_0$.

$$8\pi^3[\Phi_m(\cdot), \Phi_n(\cdot - k)] = \int_{R^3} \widehat{\Phi}_m(\omega)\widehat{\Phi}_n(\omega)^* \cdot e^{ik\omega} d\omega$$

$$= \int_{R^3} \widehat{\Phi}_{8m_1+\lambda_1}(\omega)\widehat{\Phi}_{8n_1+\mu_1}(\omega)^* \cdot \exp\{ik \cdot \omega\} d\omega = \cdots\cdots\cdots\cdots\cdots\cdots$$

$$= \int_{([0,2\cdot 4^\kappa \pi]^3} \{\prod_{l=1}^{\kappa} \mathcal{D}^{(\rho_l)}(\omega/2^l)\} \cdot O \cdot \{\prod_{l=1}^{\kappa} \mathcal{D}^{(\mu_l)}(\omega/2^l)\}^* \cdot \mathrm{e}^{-ik\cdot\omega} d\omega = O$$

Corollary 1[6]. For $n \in Z_+$, $v \in Z^3$, we have

$$\left\langle \Phi_n(\cdot), \Phi_n(\cdot - k) \right\rangle = \delta_{0,k} I_v \tag{17}$$

Corollary 2[8]. For $u \in Z^3$ and $n \in Z_+$, $\iota \in \{0, 1, 2, \cdots, 7\}$, we have

$$[\Phi_{8n}(\cdot), \Phi_{8n+\rho}(\cdot - k)] = \delta_{0,\rho}\delta_{0,k} I_v . \tag{18}$$

Theorem 2[7]. Let $\phi(x), \tilde{\phi}(x), \hbar_\iota(x)$ and $\tilde{\hbar}_\iota(x), \iota \in J$ be functions in $L^2(R^3)$ defined by (28), (29), (33) and (34), respectively. Assume that conditions in Theorem 1 are satisfied. Then, for any function $f(x) \in L^2(R^3)$, and any integer n,

$$\sum_{u \in Z^3} \left\langle f, \tilde{\phi}_{n,u} \right\rangle \phi_{n,u}(x) = \sum_{\iota=1}^{7} \sum_{s=-\infty}^{n-1} \sum_{u \in Z^3} \left\langle f, \tilde{\hbar}_{\iota:s,u} \right\rangle \hbar_{\iota:s,u}(x) . \tag{19}$$

4 Conclusion

Mechanical engineering overlaps with aerospace engineering, building services engineering, civil engineering, electrical engineering, petroleum engineering, and chemical engineering to varying amounts. The notion for nonseparable semi-orthogonal trivariate super-wavelet wraps is given and a procedure for constructing them is described. The semi-orthogonal property of nonseparable trivariate super-wavelet wraps is studied.

References

1. Telesca, L., et al.: Multiresolution wavelet analysis of earthquakes. Chaos, Solitons & Fractals 22, 741–748 (2004)
2. Iovane, G., Giordano, P.: Wavelet and multiresolution analysis:Nature of ε^∞ Cantorian space-time. Chaos, Solitons & Fractals 32(4), 896–910 (2007)
3. Zhang, N., Wu, X.: Lossless Compression of Color Mosaic Images. IEEE Trans. Image Processing 15(16), 1379–1388 (2006)

4. Chen, Q., et al.: A study on compactly supported orthogo-nal vector-valued wavelets and wavelet packets. Chaos, Solitons & Fractals 31(4), 1024–1034 (2007)
5. Shen, Z.: Nontensor product wavelet packets in $L_2(R^s)$. SIAM Math. Anal. 26(4), 1061–1074 (1995)
6. Chen, Q., Huo, A.: The research of a class of biorthogonal compactly supported vector-valued wavelets. Chaos, Solitons & Fractals 41(2), 951–961 (2009)
7. Li, S., et al.: A Theory of Geeneralized Multiresolution Structure and Pseudoframes of Translates. J. Fourier Anal. Appl. 6(1), 23–40 (2001)
8. Chen, Q., Shi, Z.: Construction and properties of orthogonal matrix-valued wavelets and wavelet packets. Chaos, Solitons & Fractals 37(1), 75–86 (2008)

Research on Determining Subsidiaries' Annual Completed Index of Electric Power Construction Group Company Based on the Grey System Theory

Wei Li[1] and Bo Zhang[2]

North China Electric Power University, 071003 Baoding, China
1587618895@qq.com

Abstract. The aim of this paper is to solve the problem how the electric power construction group company determines the completed amount of each branch every year. It introduced the GM (1, 1) model and gray prediction accuracy requirement, and made some improvement on the model. Then it used the model to predict a subsidiary's annual completed index and the result showed the improved model has higher prediction precision. Finally, used this model to predict the nearly two years of the subsidiary's annual completed indexes.

Keywords: electric power construction group company branch index GM (1, 1).

1 Introduction

In modern society the competition between the companies is becoming increasingly. Many small companies in the competition can not afford the risk of accidents to be eliminated. So the collectivization trend of company is becoming more and more obvious. However, with the increase of the scale the company will generate many branches, and then it faces a problem that how much annual completion amount of branches should be. Now many group company adopts a bargaining way to set completed index of every branch company, namely group chief says an amount, the director of branch says one, and eventually define the amount that is acceptable to both through negotiations. Such index may be completed, but there will always be some losses in the interests of the party. So the group company will reduce the number of income. Therefore it is necessary for group company to adopt a scientific prediction method to determine the completed index of branches. It can improve Credibility and can avoid dispeling positive attitude of branch staff because of blind ration. In short, the aim of company is to enable group company of electric power construction to maximize the benefits.

Grey prediction finds some rules by simply processing the raw data. Grey prediction model is better than other forecasting models, because it needs less historical data. It has the advantage of high prediction precision.GM (1,1) model as the classical model of grey prediction model can complete scientificly and rationally the prediction indexes of every branch company. This model has great significance on the group company of electric power construction especially for developing one, some of whose branches

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 201–205.
springerlink.com

maybe be established for a short time. The previous data is little. The model just can play its role well. Because group company of electric power construction have many branches, we only select one branch to make index forecast and the prediction process of other branches is same.

2 The Construction Process of GM (1,1) Model Theory

2.1 Inspection and Processing of Data

Set basic data for $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \wedge, x^{(0)}(n))$

To ensure that GM (1, 1) grey forecasting model can use the basic data to carry on the forecast, it needs to calculate the series $\lambda(k)$:

$$\lambda(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)} \qquad (k = 2,3.\wedge, n) \tag{1}$$

If all $\lambda(k)$ have fallen on covered capacity $X = (e^{\frac{-2}{n+1}}, e^{\frac{2}{n+1}})$, the series can be used for grey forecasting of GM (1, 1) model. Or need to make the necessary transformation processing of datas $X^{(0)}$ so that they can fallen on covered capacity. This paper adopts weakening buffer operator to amend original data. Process is as follows.

After the average weakening buffer factor role the number is

$$XD_1 = \left\{x(1)d_1, x(2)d_1, \cdots x(n)d_1\right\}$$

and $x(k)d_1 = \frac{1}{n-k+1}\sum_{i=k}^{n} x^{(0)}(i), (k = 1,2,\cdots, n)$

If the first processing data does not fall within $X = (e^{\frac{-2}{n+1}}, e^{\frac{2}{n+1}})$, you should continue to use the above model to process the data until the data meets the requirement.

2.2 GM (1,1) Model

Formula of accumulated generating series is:

$$\hat{x}(k+1) = (x^{(0)}(1) - \frac{b}{a})e^{-ak} + \frac{b}{a} \qquad (k = 1,2,\wedge, n-1). \tag{2}$$

Obtain the generation sequence $\hat{X}^{(1)} = \{\hat{x}^{(1)}(1), \hat{x}^{(1)}(2), \wedge, \hat{x}^{(1)}(n)\}$.

We get forecast data column after LeiJian. $\hat{X}^{(0)} = \{\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \wedge, \hat{x}^{(0)}(n)\}$

And $\hat{X}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1)$.

2.3 Error Inspection

2.3.1 Residual Error Inspection

Calculate $x^{\wedge(0)}(k)$ according to the model. Then calculate the residuals $\varepsilon(k) = x^{(0)} - x^{\wedge(0)}$ and relative error $\alpha(k) = \frac{\varepsilon(k)}{x^{(0)}(k)}$. Generally less than 5% average relative error satisfies the requirement of accuracy.

2.3.2 Posterior Error Inspection

2.3.2.1 Calculation of Variance Ratio: $C = \frac{S_2}{S_1}$

Among them, S_1 is the standard deviation of the original series and S_2 is the standard deviation of residuals.

2.3.2.2 Calculate small error probability $P=\{ | \varepsilon(k) - \bar{\varepsilon} | < 0.6745 \times S_1 \}$

Determine predicted accuracy levels according to the criteria in the following table.

predicted accuracy levels	P	C	predicted accuracy levels	P	C
best （first）	>0.	<0.	good （third）	>0.7	<0.
better （second）	>0.	<0.	bad （fourth）	≤0.7	≥0.

3 Case Analysis

This paper shows the yearly finished amount of Hebei branch in a group company of electric power construction for nearly five years (2003-2010). The specific amount are (23673.38 25213.41 26147.8 35174.89 100640.42 51858.72 135911.33240440.6). The unit is ten thousand yuan.

3.1 Data Processing

All $\lambda(k)$ have fallen on covered capacity (0.8,1.25) through calculating. The series ratio are (0.94 0.96 0.74 0.35 1.94 0.38 0.57), 5 datas of which are hardly ever in the feasible region. Use average weakening buffer operator firstly to deal with the original series, and the first processed data are (79882.57 87912.45 98362.29 112805.19 132212.76 142736.88 188175.96 240440.6). Then series ratio are (0.91 0.89 0.87 0.85 0.93 0.76 0.78), 2 datas of which are hardly ever in the feasible region. Continue to process the data for the second time. The data are (135,316.08 143,235.16 152,455.61 163,274.28 175,891.55 190,451.14 214,308.28 240,440.6), and series ratio are (0.94 0.94 0.93 0.93 0.92 0.89 0.89).Now all data fall on the feasible region, so you can use the GM (1 , 1) model to predict without further data processing.

3.2 Forecasting Precision Analysis

Through the use of the grey modeling software MATLAB7.0, it can obtain a=-0.0886, b=119770.

The prediction model is for

$$\hat{x}(k+1) = 1486941.87e^{0.0886k} - 1351805.87 \quad (k = 1,2,\wedge,n-1)$$

For convenience, it only retains the integer. Then LeiJian generating data and get the forecast data (137780 150550 164490 179720 196370 214560 234430). Table is done as follows:

year	the original GM (1, 1) model			GM (1, 1) model of the second adjustment		
	raw data	fitted values	relative error (%)	The second processing data	fitted values	relative error (%)
2003	23673.38	23673.38	0	135316	135316	0
2004	25213.41	11430	55	143235	137780	3.81
2005	26147.8	17640	55	152455	150550	1.25
2006	35174.89	27210	22	163274	164490	-0.74
2007	100640.4	41980	58	175891	179720	-2.18
2008	51858.72	64770	-24	190451	196370	-3.11
2009	135911.3	99920	26	214308	214560	-0.12
2010	240440.6	1541400	-541	240441	234430	2.5
average relative error	-349			1.41		

The result from the raw data shows the average relative error is -349%,> 5%.The error is far greater than accuracy. The average relative error of the second revised data is 1.41%< 5%,which meet the forecast accuracy.

3.3 Set the Completed Index of 2011

By examining the above model, the prediction accuracy is very good. The model can serve as the prediction model of Hebei branch. Through prediction the indexes of Hebei branch in the next 2 years are (253731, 277109) wanyuan. The prediction process of other branches is the same as that of Hebei branch.

4 Summary

This article uses GM (1, 1) grey forecasting model to predict completed index of Hebei branch in a group company of electric power construction in next 2 years, which has a very important significance on the group company of electric power construction. This paper checks residual error and posterior difference after processing data, and then meet high accuracy. So this model has higher practicality and efficiency. However, this prediction model demands a higher accuracy of the basic data offered by each branch. Each branch must accurately report the amount of actual completion of the previous years. And because electric industry is very volatile, prediction index of each branch needs to be done every year.

References

[1] Niu, D., Cao, S., et al.: Power Load Forecasting Technology and Application. China Power Press (1998)
[2] Deng, J.: Grey Prediction and Grey Decision. Huazhong University of Science and Technology Press, Wuhan (2002)
[3] Liu, S., Dang, J.: Grey System Theory and Application. Science Press (2004)
[4] Li, J.: The Research on Electric Construction Project Management Information System Based on WEB. North China Electric Power University (2010)
[5] Liu, C.: HBO Electric Power Construction Company Development Strategy Research. Huazhong University of Science and Technology (2008)
[6] Wang, Z., Ren, Y.: The Establishment of Electric Power Construction Company and Performance of Integrated Management System. Electric Power Standardization and Technical Economy (3) (2008)

Tri-band Antenna Design with a Novel Slot Structure for Mobile Communication

Ge Zhang and Zhibin Zeng

Engineering Research Center of Digital Audio & Video,
Ministry of Education
Communication University of China
Beijing, China
zg19851985@163.com

Abstract. Multiple mobile communication service requires multi operations to share one antenna, in order to save limited spare in the terminal. In order to achieve miniaturization and multi-band of an antenna internal mobile handset, a tri-band antenna structured with planar inverted F-type is proposed in this paper. A compact microstrip antenna for triple-band of GSM900, DCS1800 and ISM2.4GHz was realized based on PIFA antenna with a novel slot structure. Simulation results showed that our new antenna with excellent frequency and radiation characteristics, the antenna satisfies terminal of modern generation wireless communication system.

Keywords: internal antenna, tri-band, planar inverted F-type, PIEA.

1 Introduction

Antenna in the radio equipment has two main functions: the first is the energy conversion function and the second is the directional radiation function. Energy conversion function is to guide the traveling waves and the free-space wavelength conversion. The transmitting antenna can transform electromagnetic wave (high frequency current) guided by the feeder to space radiation electromagnetic wave and radiate to the distance. The receiving antenna transforms the electromagnetic waves of space into electromagnetic waves (high frequency), and then transmits to the receiver.

At present, with the mobile communication technology continuous development, the handheld wireless terminal device in people's daily life plays an important role in these devices, mobile phones is used as the most frequently communication tools. As the most important RF front-end devices for transmitting and receiving electromagnetic waves, the future development trend of mobile phone antenna is bound to be small, built-in, multi-band and intelligent. The microstrip antenna as a new antenna has the following advantages: small size, light weight, low profile, conformal and the carrier, simple to manufacture [1]. PIFA microstrip antenna (Planar Inverted F-shaped Antenna) is widely used, due to the antenna shaped like a reverse English letters F, so named. This phone antenna has small size, so it can take full

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 207–213.
springerlink.com

advantage of the printed circuit board as the ground plane, and through the short-circuit piece resonator length is reduced to a quarter of wavelength, therefore, since its inception has been the main form of the handset antenna. With communication services surged in recent years, the domestic mobile communication system using both the GSM900 and DCS1800 frequency bands, more foreign is already using the band expanded to a higher PCS1900, ISM2450 band. Nowadays, dual-band, tri-band and multiple frequency technology have become a research hotspot. Multi-frequency techniques using a single multi-mode, multi-layer overlap patch and single load. Gaps in the monolayer loading techniques load, also known as the patch surface slotted design has aroused widespread attention for its simple, practical, easy to process, reduce the antenna size.

2 PIFA Antenna and Multi-frequency Technique

The typical structure of the PIFA consists of a flat rectangular piece of metal, a large ground plane, a narrow short-circuit metal plate [2] (at the length of the rectangular flat metal piece to the shorter side of the edge). As shown in Figure1, L_1 and $L2$ are the length and width of the radiating patch, respectively, W is the width of the short-circuit film, F is the distance between the feed point jumper, and h is the height of the PIFA. The short-circuit metal plate placed in the radiation between the metal ground plane, allows the length of the rectangular radiation metal reduce to half, so as to achieve the goal of smaller antennas, the location of the short-circuit metal plate, $TM10$ of the electric field is zero. When the width of short-circuit metal plate is more narrow than the side of the flat rectangular metal surface L_1, the antenna effective inductance increase and the resonant frequency is lower than the traditional short-circuit microstrip antenna. Therefore, the same size of flat rectangular metal surface to get the same resonant frequency, it must narrow the flat rectangular metal surface to achieve the purpose reducing the size of original antenna.

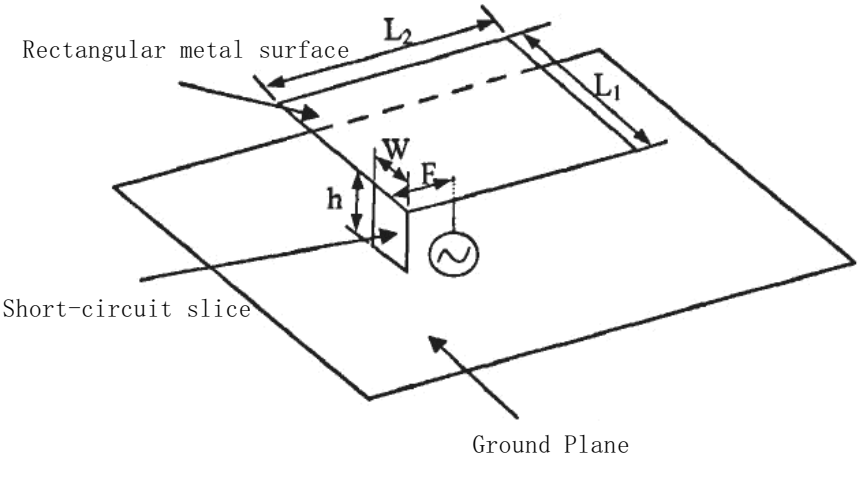


Fig. 1. The typical structure of the PIFA

The approximate resonant frequency of the PIFA antenna can be described as [3]:

$$f_r = \frac{c}{4(L+W)} \tag{1}$$

Here, c denotes the speed of light; L and W represent the long side and the short side of the radiation patch, respectively.

According to the assumption that the PIFA antenna as a half-wavelength ring slot antenna, we can derive this approximate formula to solve the resonant frequency [4]. In addition, the height between the radiator and the ground has a very critical impact on the bandwidth of the antenna.

There are two methods for PIFA antenna to realize multi-band. One is the form of multi-feed point, but the frequency and coordination scope is subject to certain restrictions, usually do not use this technology [5]. The other is the technology to slotting on the radiator and this multi-band technology is mostly used in the practical application [6]. PIFA using the slotting design which comprehensive utilization of the advantages of the patch antenna and slot antenna. Patch antenna is easy to design, processing, multi-frequency work. After slotting in the rectangular radiation patch, the patch structure changes, the natural mode of the surface changed, and the surface current distribution also will be changed. This might be impact the original resonant, but would also produce additional frequency, namely the so-called "multi-frequency resonant antenna". At this point, the new antenna structure is the combination of the patch antenna and the slit antenna.

After slotting, the antenna radiation mechanism can reference the microstrip slit antenna to analysis. The difference is that the gap of the patch antenna is open in the patch, whereas the slot antenna is open in the ground plate. The radiation field of the slot antenna can be algorithm by the electric vector [6]. When the slit width is much smaller than the vacuum wavelength, E and H surface pattern can be obtained by derivation of the Equation (2) and (3).

$$E_0 = \frac{jk_0E_0LW}{4\pi} \bullet \frac{e^{-jk_0}}{r} \bullet \frac{\sin X}{X} \bullet \sin\varphi \tag{2}$$

$$H_\theta = \frac{jk_0E_0LW}{4\pi} \bullet \frac{e^{-jk_0}}{r} \bullet \frac{\sin X}{X} \bullet \cos\varphi\cos\theta \tag{3}$$

$$X = k_0L\sin\theta\cos\varphi/2 \tag{4}$$

In Equation (4), L and W denote the length and width of the gap, respectively. Calculated on the application, we need to know the exact value of the gap transverse electric field. This can be used for the antenna which has rules for the shape of the antenna But with the continual development and progress of the microstrip antenna, a lot of complex structure and process innovation antenna came into being. Such as *C*-groove, Cross recessed high dielectric constant and high-power medium antenna, in order to accurately calculate the gap on the direction of electric field and far-field seeking diagram is very difficult. Therefore, modern antenna design is based on simulation-based test, supplemented by theoretical analysis, many of which are by virtue of the experience. Carefully adjust the location of the slotted, the slot size and the width of the slotted gap, load short-circuit wall in the right position. It is possible to make high-order mode resonant frequency is

reduced to the desired frequency band, without affecting the low mode in the low frequency applications [7]. Therefore, in recent years, this approach is often used to perform the multi-frequency antenna design.

3 Tri-band PIFA Antenna Design

3.1 Antenna Model Size

Rectangular patch antenna to this article as a carrier study, the use of slotted technology in the radiator, based on finite difference time domain method for design, simulation and optimization may be as shown in the antenna model. Its antenna underlying medium height of 1.64mm, the dielectric constant of 4.5, the upper dielectric layer of air, a total height of 10mm, the ground plane size is 87*50mm (PCB board size), patch size is 61.4mm*18.2mm. Specific patch size as shown in Figure 2.

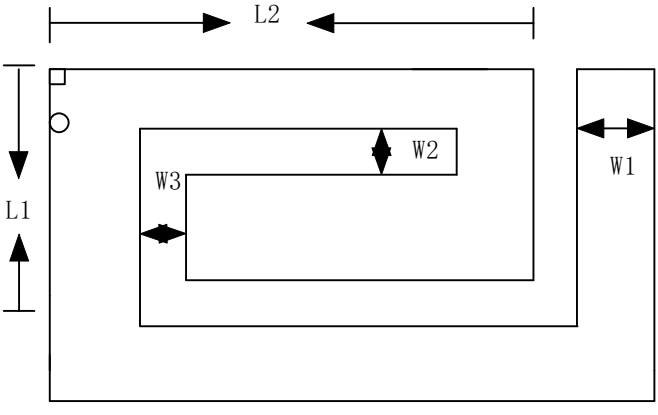


Fig. 2. Antenna patch

After repeated optimization and modification, and we can get ultimately the slot size values which are shown as follows (Unit mm):

L1 = 15.20mm, *L2* = 48.43mm;

W_1 = 5.12mm, *W2* = 2.84mm, *W3* = 3.15mm.

3.2 Antenna Tuning

The resonant frequency is adjusted by the following steps: First, adjust the gap length and width. First determine the resonant frequency of the low-band antenna. The basic principle is that the gap cannot be bent too much in order to avoid cutting too much and affect the bandwidth and efficiency and then adjust the band gap and length.

Second, feed point and short-circuit wall plays the role of transforming the antenna impedance. Their inductive reactance and the capacitance which are formed by the radiator and the ground constitute a resonant circuit. The longer, the more obvious of the change (actual performance in change of the antenna height).when to change their thickness (the cross-sectional size) will change the impedance of the antenna [8]. Adjust the matching of the antenna, changing the feed point and the short-circuit wall, position, length and cross-sectional size.

3.3 Influence of Parameters

Change the slot size, that is fine-tuning the current path, thus changing the impedance characteristics and affect the resonant frequency, each parameter change on the three frequencies (f_1 = 0.88GHz, $f2$ = 1.8GHz, $f3$ = 2.45GHz), as shown in Table 1.

Table 1. The impact of various parameters on the frequency

Parameter	Change			Result
Adjust L1(mm)	15.20	15.11	15.01	L1↓
f1(GHz)	0.88	0.86	0.81	f1↓
f2(GHz)	1.80	1.85	1.89	f2↑
f3(GHz)	2.45	2.51	2.62	f2↑
Adjust L2(mm)	48.43	48.60	48.91	L1↑
f1(GHz)	0.88	0.89	0.91	f1↑
f2(GHz)	1.80	1.78	1.75	f2↓
f3(GHz)	2.45	2.43	2.41	f2↓
Adjust W1(mm)	5.12	5.09	4.80	L1↓
f1(GHz)	0.88	0.86	0.80	f1↓
f2(GHz)	1.80	1.82	1.85	f2↑
f3(GHz)	2.45	2.46	2.48	f2↑
Adjust W2(mm)	2.84	2.80	2.78	L1↓
f1(GHz)	0.88	0.90	0.91	f1↑
f2(GHz)	1.80	1.78	1.76	f2↓
f3(GHz)	2.45	2.42	2.40	f2↓
Adjust W3(mm)	3.15	3.18	3.20	L1↑
f1(GHz)	0.88	0.89	0.91	f1↑
f2(GHz)	1.80	1.78	1.76	f2↓
f3(GHz)	2.45	2.42	2.40	f2↓

PIFA is a small antenna; a little error in the design can make serious effect on the antenna performance and the various parameters of mutual restraint, mutual influence. Fine adjustment is therefore necessary on several main parameters, comprehensive comparisons in order to optimize the antenna size to obtain good performance.

4 Simulation and Analysis

Antenna in the sending and receiving process of the electromagnetic energy will inevitably produce energy losses in the simulation space to run the simulation program, shown in Figure 3, the planar inverted F antenna return loss curve. Three frequency bands cover GSM/DCS/ISM band. The simulation results meet the current operating frequency band of GSM mobile terminal required.

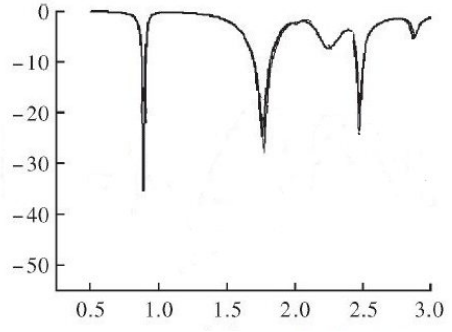


Fig. 3. The tri-band antenna return loss

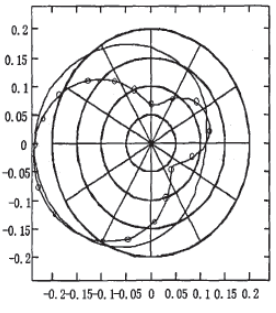


Fig. 4. 900MHz directional diagram

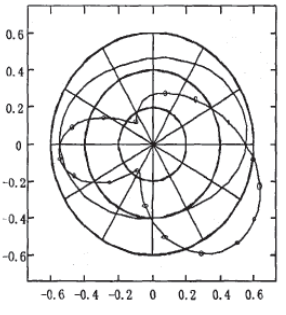


Fig. 5. 1800MHz directional diagram

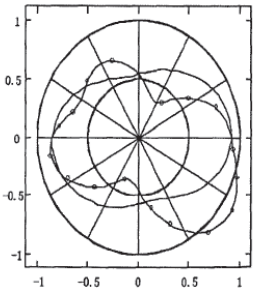


Fig. 6. 2450MHz directional diagram

5 Conclusion

According to the basic principles of the PIFA antenna with a compact and simple shape planar inverted F microstrip antenna design, antenna modeling and simulation, and continue to debug, optimize, the frequency of the antenna can cover the three bands of GSM900MHz, DCS1800MHz and ISM2450MHz. Antenna gain and bandwidth meet the work requirements of modern mobile communication systems. Simulation shows that appropriate selection on the parameters of the open slot on the antenna resonant frequency for the tri-band PIFA antenna design will provide an improved method.

References

[1] Kim, S.W.: Smart antennas at handsets for the 3G wideband CDMA systems and adaptive low-power rake combining schemes. Faculty of the Virginia Polytechnic Institute and State University, Blacksburg Virginia (2002)
[2] Zuo, S., Zhang, Z., Wu, W.J., Yin, Y.: A compact tri-band PIFA antenna for WLAN and WiMAX applications. Microware and Optical Technology Leters 52(4), 919–922 (2010)
[3] Chen, S., Wang, B., Pang, J., Xiao, S.: Compact and wideband PIFA for DCS/PCS/UTMS/WLAN communication system. Microware and Optical Technology Letters 52(5), 1097–1100 (2010)
[4] Wu, C.-H., Wong, K.-L.: Communitions; Ultra-wideband PIFA with a capacitive feed for penta-band folder type mobile phone antenna. In: Antennas and Propagation Society International Symposium, vol. 57(8), pp. 2461–2464 (2009)
[5] Wong, K.-L.: Plannar Antenna for Wireless Communications. Wiley Interscience, Hoboken (2003)
[6] Lev, P., Nikolay, T., Yehuda, L.: Multi-band flat-plate inverted-F antenna for WiFi/WiMAX operation (2008)
[7] Chen, C.F., Huang, T.Y., Wu, R.B.: Design of dual-and triple–passband filters using alternately cascade multiband resonators
[8] Stutzman Warren, L., Thiele Gary, A.: Antenna theory and design (1998)

Miniature Wide-Band Patch Antennas for Radar Network Applications

Dijian Xu and Jiankui Zeng

Chongqing University of Science and Technology,
Chongqing, 401331
Xdj1003@163.com

Abstract. Radar network which is composed of many types of radar gets more and more attention from researcher recently. This technology can greatly improve radar efficiency. The miniature of antennas in this type of radar system is very important because in many applications the space for radar system to install is very small such as in aircraft and missile, etc. In this paper, the principle of radar network based on transmitting diversity is described firstly, and then a miniature wide-band patch antenna for radar network is presented. In this method, the performance of radar is better than that of conventional radar.

Keywords: Signal processing, radar network, patch antennas.

1 Introduction

It has been recently shown that radar network system has the potential to dramatically improve the performance of the communication systems over single antenna system [1],[2]. The probing signal design problem for the narrowband radar network radar has been addressed. Our approach to this design problem is similar to the mathematical approach and is different from the more pragmatically approach. Unlike the traditional beam-forming approach, which employ highly correlated signals of transmitting or receiving antenna arrays to steer a beam towards a certain direction in space, radar network makes use of the independence between signals from different transmitters and receivers to improve the more information received from the target and the robustness of the transmit-receive link. Our main contributions are the following: i) we address the question of determining a desirable transmit beam pattern, and show how to obtain such a beam pattern; ii) we modify the beam pattern matching criterion in several ways; and iii) we outline an efficient Semi-definite.

Microstrip patches are currently being used for many applications due to its several key advantages over conventional wire and metallic antennas. Its major advantages include its low profile, light weight, and ease of fabrication and integration with RF devices. However, there are several applications where even the physically small size of a conventional microstrip patch antenna is still too large. One such scenario is for its use on a mobile communications.

Global System for Mobile communication is a globally accepted standard for digital cellular communication. GSM is the name of a standardization group

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 215–220.
springerlink.com

established in 1982 to create a common European mobile telephone standard that would formulate specifications for a pan-European mobile cellular radio system operating at 900 MHz. It is a type of personal communication service based on wireless technology. In this system the whole communication network is divided into cells that cover a certain geographic area.

Motivated by the radar network technique in communication systems, the two new concepts of radar network radar are introduced. One is transmitting diversity radar network radar [3], the other one is receiving and transmitting diversity radar network radar [4]. The proposed radar network radar enjoys the same benefits that radar network communication systems have[5][6][7]. Specifically, the transmitting diversity radar network radar can greatly improve the radar's performance over traditional radar on anti-intercept of radar signal, weak target detection, etc.[3]

The remainder of the paper is organized as followed. First, the principle of radar network radar based on transmitting diversity was introduced. Then receiving signal processing method of radar network radar based on transmitting diversity was presented. And then, the miniature wide-band Patch Antennas for radar network radar is presented. At last, some conclusions are drawn.

2 The Principle of Radar Network Based on Transmit and Receive and Transmit Separately

To detect small RCS targets, such as cruise missile or stealth fighter, surface-based or ship-based phased radars are usually designed for high peak power levels and large power-aperture-gain (PAG) products. Unfortunately, this will bring some problems. First of all, these radars must operate in the presence of strong clutter, resulting in challenging requirements on system dynamic range fundamentally limited by the "state of the art"' Secondly, strong clutters also result in challenging requirements on phase noise, stability, isolation, spurs, and other hardware-related specification. Thirdly, due to high peak power levels, radar signals will suffer more easily from harmful electromagnetic interference (EMI) and enemy Electronic Intercept (ELINT) system.

To partially address these problems, radar arrays are being designed to use Digital Beam Formation (DBF) on receives [5]. [6].In a DBF system, analog-to-digitization (AID) conversion is performed at each of the receive elements in the array, so can reduce A/D dynamic range by N times (N is the number of receive sub-arrays) and facilitate the formation of multiple simultaneous receive beams (which enables faster search rates).Through such arrays offer many benefits, they still operate much like earlier phased array radar mode and cannot resolve the signals anti-intercept problem. Radar network radar based on transmitting diversity can solve well the problems described above[8]. On surface-based or ship-based phased array, on transmitting, the array aperture would be subdivided into low-gain elements group (or sub-arrays) in elevation, each radiates a unique[9], orthogonally coded waveform (note that the various emitted signals will not be combined to form a single focused beam, instead, the radiated energy will cover a broad angular sector) shown in Fig. 1. Because the array is subdivided into M subarrays, the transmit gain and the transmitted power per channel are reduced by M times.

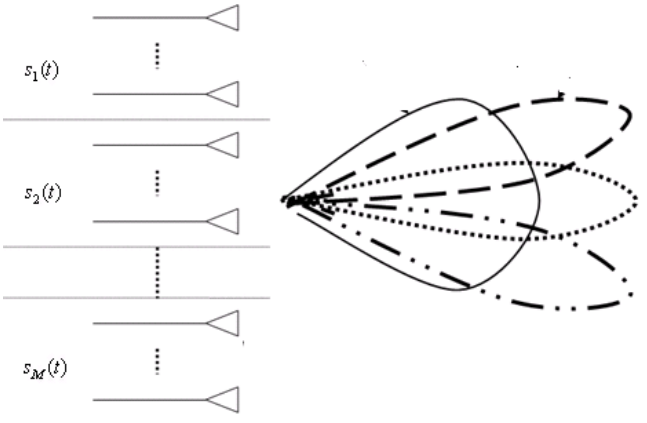


Fig. 1. The principle of radar network radar based on transmitting diversity

On receiving the signal at each individual receiver is processed through a bank of M matched filters. Each filter is matched to one of the transmitted waveforms, thereby recovering the returns due to a single transmit signal. This produces a total of MN matched filter outputs. Since the locations of each transmit and receive element are known, these MN signals can be phased and combined (analogous to normal transmit and receive bean-forming) to form high gain receive beams in one or more directions. By digital time-delay or DBF technique, the beams will cover the volume of space illuminated by the wide transmit beam. Further integration (i.e., Doppler processing) is used to maintain sensitivity, and to achieve the same range, as desired.

System above can be in either the MIIMO or the conventional transmit modes (all subarrays transmit the same signal)[10], as needed, a great deal of flexibility is provided.

3 Radar Network Signal Processing

According to the principle described above, the configuration of radar network radar based on transmitting diversity is depicted. Now consider an radar network system with transmit elements (or sub-arrays) and M receiving elements (or subarrays), transmitting orthogonal waveforms $S_1(t), S_2(t), \cdots, S_M(t)$.

Assuming narrow signals (this is not a necessary requirement, but simplifies the discussion) and isotropic elemental radiators, the aggregate waveform incidents upon any target can be represented as

$$p(t) = \alpha_1 \sum_{m=1}^{M} S_m(t - \tau_m) \tag{1}$$

Where, $\tau_m = \dfrac{(m-1)d\sin\theta}{c}$, d is the delay time from the transmit to the target. α_1 is the attenuate factor, which is same for all transmitter. Then, the summation of all waveform is

$$p(t) = \alpha_1 \sum_{m=1}^{M} S_m(t)\alpha_m(\theta) \tag{2}$$

Where, θ is the target direction of arrival, $\alpha_m(\theta)$ is the transmit response vector of mth element. The wave is reflected back to antenna. The n-th subarray received waveform is then

$$x_n(t) = \alpha_2 \cdot p(t-\tau_n) = \alpha_2 \cdot p(t)e^{-j\phi_n} \tag{3}$$

4 The Miniature Wide-Band Patch Antennas for Radar Network

The problem of radar target tracking, especially with mono-pulse techniques, can be divided into two categories: viz. active tracking and passive tracking. Active radar target tracking uses both range and bearings measurements. This is the most common type of tracking in real applications. Traditionally, the Kalman filter and its variations have been used for active tracking. However, when the target under track can perform maneuvers, multiple dynamic models will be adopted to describe the target. These models usually involve strong nonlinearities. In these cases, Kalman filter based methods will not provide accurate estimations. Passive tracking is different from active tracking, where only the bearings measurement, i.e. the direction of arrival (DOA), is available. Because of this reason, bearings-only tracking is also called DOA tracking. The problem of bearings-only tracking arises in a variety of important practical applications in surveillance, guidance or positioning systems. Typical examples are aircraft surveillance (using a radar in a passive mode or an electronic warfare device), underwater target tracking using passive sonar. The purpose of bearings-only tracking is to estimate the kinematics (the positions and velocities) of a target using its noise-corrupted bearings measurements[11].

Due to the inherent nonlinearities in the observation model, bearings-only target tracking has become a standard nonlinear filtering problem that receives intensive investigations. Target tracking using sensor networks is an emerging new tracking application. Sensor networks are typically composed of one or more central information processing units with a large number of sensor nodes, which have limited communication and computation capabilities and limited power supply. During recent years, the continuing miniaturization of computing and (wireless) communication circuitry as well as sensor devices has enabled mass production of intelligent wireless micro-sensors at a low cost. Tracking targets with geographically dispersed and cooperating sensors is attractive for several reasons[12].

First, it can improve the robustness of tracking algorithms; sensors deployed close to targets would result in more reliable signal readings. Also, it can be more cost effective. The challenge is to design a tracking method which ingeniously reconciles the two defining characteristics, abundance in quantity[13][14]. and inferiority in quality, to realize the desired robustness. One important application of wireless sensor networks is target tracking, where the target of interest ranges from moving objects in civil and military surveillance applications, to changes in light, temperature, pressure, and acoustics in environmental monitoring. The type of signals to be sensed is determined based on the types of objects to be tracked. In this dissertation, a multiple

sensor fusion and tracking algorithm based on particle filtering was developed for ground vehicle tracking using acoustic sensor networks.

Several designs were evaluated with the commercial field simulator based on Method of Moments and parametric study has been performed. Finally optimized designs for wide band GSM applications are chosen. It shows the geometry of full U-slot patch antennas with shorting wall. The full U-slot patch has a length of L = 70 mm and a width of W = 84 mm. The center frequency is observed at 1.06 GHz. The shorting wall is applied between ground and patch at the non-radiating edge of the patch which is made up of the same material. The co-axial feed is located at the center of the patch (x = 42, Y = 35). The radius of the feed conductor of coaxial probe is 2 mm. Its slot dimensions are shown in Fig.1. Air dielectric is used between the patch and the ground having thickness, h = 28 mm. The antenna is placed on ground plane.

In a tracking problem, a target is usually described by a state space model, which contains a state process model and an observation model (or measurement model). The state space approach to time series modeling focuses the attention to the state vector of a system. The state vector contains all relevant information required to describe the system. For example, in tracking problems, the system state could be the kinematic characteristics of the target (i.e. position, velocity, etc.). In the current literature, the target tracking problem is always formulated as a state estimation problem, which is also known as filtering. In state estimation, the Kalman filter is widely used when the model is linear and the additive noise is Gaussian. But the real world dynamic systems are often more complex, typically involving nonlinear and non-Gaussian elements.

5 Conclusion

In this paper, we firstly investigated the radar network signal model and its signal processing steps. The each sub-array of MIMO radar performs detection respectively, and then the result is processed by the nonlinear filter. In this paper, a novel method is presented for increasing the bandwidth of a rectangular patch as well as considerable reduction in size by using half U-slot structure. Various full and half U-slot structures are designed and results are studied. The feed is optimized in such a way that better impedance bandwidth is achieved. Parametric variation carried out for full U-slot with shorting wall and results are compared. The simulated and experimental results have been verified using commercially available method of moment based IE3D software and Agilent Vector Network Analyzer; E5062A respectively. Use of shorting wall at the non radiating edge is found to reduce size of RMSA.

The half U-slot patch antenna maintains similar wide band behavior. Performance of full U slot and half U slot is compared. It is seen that with shorting wall size reduction is 68 to 76% and radiation efficiency is obtained above 84% for operating band as compared to conventional RMSA design for the particular frequency range.

Acknowledgments. This research is funded by construction project of Chongqing science and technology platform (The technique center of Chongqing Metallurgical detection and equipment), the project No. is cstc2011pt-gc70007. This research is also supported by the Chongqing construction project of university Innovation team (Metallurgical detection technique), the project No. is 201028.

References

[1] He, Z.-S., et al.: MIMO radar and its technical characteristic analysis. ACTA Electronic Sinica 33(12A), 2441 (2005)
[2] Foschini, C.J.: Layered space-time architecture for wireless communication in a fading environment when using multiple antennas. Bell Labs Technical Journal 1, 41 (1996)
[3] Foschini, G.J., Gans, M.J.: On the limits of wireless communications in a fading environment when using multiple antennas. Wireless Press Communication 6, 311 (1998)
[4] Doucet, A., de Freitas, N., Gordon, N.: Sequential Monte Carlo Methods in Practice. Springer, New York (2001)
[5] Rabideau, D.J., Parker, P.: Ubiquitous MIMO digital array radar. In: The 37th Asilomar Conference on Signals, Systems and Computers, p. 1057 (2003)
[6] Ristic, B., Arulampalam, S., Gordon, N.: Beyond the Kalman Filter: Particle Filters for Tracking Applications. Artech House Publishers (February 2004)
[7] Fishler, E., Haimowich, A., Blum, R., Cimini, L., Chizhik, D., Valenzuela, R.: Statistical MIMO radar. In: 12th Conf. on Adaptive Sensor Array Processing (2004)
[8] Rabideatu, D.J., Howard, L.C.: Mitigation of digital array nonlinearities. In: Proc. IEEE Radar Conf., p. 175 (2001)
[9] Nardone, S., Lindgren, A., Gong, K.: Fundamental properties and performance of conventional bearings-only target motion analysis. IEEE Transactions on Automatic Control 29(9), 775–787 (1984)
[10] Teneketzis, D.: The decentralized quickest detection problem. In: IEEE International Large-Scale Systems Symposium (1982)
[11] Kovattana, T.: Theoretical analysis of intrusion alarm using two complementary sensors. Final Technical Report, Stanford Research Institute, Menlo Park
[12] Fefjar, A.: Combining techniques to improve security in automated entry control. In: Proceedings of the Carnahan Conference on Crime Countermeasures (1978)
[13] Stearns, S.D.: Optimal detection using multiple sensors. In: Proceedings of the Carnahan Conference on Crime Countermeasures (1982)
[14] Chair, Z.: Ph.D. dissertation in progress, Syracuse University, N.Y.

Grid-Connected Renewable Electricity Storage: Batteries vs. Hydrogen

Rodolfo Dufo-López and José L. Bernal-Agustín

Department of Electrical Engineering, University of Zaragoza,
Calle María de Luna 3, 50018, Spain
{rdufo,jlbernal}@unizar.es

Abstract. The random nature, abrupt variations of electricity generation, and high differences between the peaks and valleys of the demand curve and the generation curves of wind farms and photovoltaic generators can be avoided by means of electricity storage in batteries or hydrogen. The storage would bring many benefits to the electrical grid. However, present acquisition costs of batteries and hydrogen components are too high so that these systems cannot be profitable economically. In terms of energy, the storage in batteries is much better than in hydrogen because the roundtrip efficiency is much higher. In a matter of years, some forms of battery technology could be competitive for these applications.

Keywords: Wind, Photovoltaic, Grid-connected Systems, Storage, Batteries, Hydrogen.

1 Introduction

Electricity generation by means of renewable sources in Spain and other countries has created economic growth and has encouraged the installation of many grid-connected photovoltaic generators and wind farms around the country.

However, there is critical issue involved with the renewable generation: electricity generation does not match the demand curve (peak and valley periods of energy generation do not usually correspond with peak and valley periods of energy consumption), and the demand curve cannot be accurately forecasted. Another important problem related to the wind farms is the large generation gradients, i.e., large variations in electricity generation in very short periods of time. In some cases, the electrical grid cannot absorb these variations properly, and these fluctuations can cause grid problems and the disconnection of the wind farms, as occurs in Spain on windy days.

If the renewable penetration of the electrical grid continues to increase, it is probable that the electrical grid will have many stability problems.

The difference between the generation curve and the demand curve can be reduced by means of electricity storage: during peak periods of renewable generation which do not correspond with peak periods of demand, electricity can be stored, and during peak periods of demand, the stored electricity can be injected into the electrical grid.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 221–225.
springerlink.com

Doing so may enable the electricity generation curve to better match the demand curve. Additionally, storage can absorb abrupt variations in the electricity generation, injecting into the electrical grid a more constant generation and preventing grid stability problems.

The most common electricity storage technique is hydro storage by means of water-pumping, reversible hydro plants [1]. This technology is mature and cheap; however, it can only be applied in mountainous regions, and its environmental impact must be taken into account. Another storage option is compressed air energy storage (CAES), but it is not a mature technology, and it can only be applied in some regions with special geological characteristics [1].

There are other techniques of storage which can be installed anywhere (they do not depend on the characteristics of the terrain): electrochemical batteries [2-4] storage and hydrogen storage [5].

2 Grid-Connected Renewable Electricity Systems with Storage

At the present time in Spain, the most important technologies for renewable electricity production are wind power and photovoltaic power generation.

Currently, the grid-connected wind farms and photovoltaic generators do not use storage: all the electricity generated is instantaneously injected into the electrical grid. As mentioned above, if the power installed in wind farms and photovoltaic generators continues to increase, the electrical grid will have serious stability problems. To avoid these problems, battery and hydrogen storage are good options as they can be installed anywhere.

Fig. 1 shows a renewable generation system composed of a photovoltaic generator (PV), wind turbines, a rectifier (AC/DC converter), an inverter (DC/AC converter), storage (batteries or hydrogen), controllers, and a transformer for the connection to the electrical grid. The system shown in Fig. 1 would be a hybrid system with storage, but it could be a PV-storage system (without wind turbines) or a wind farm with storage (without PV generator).

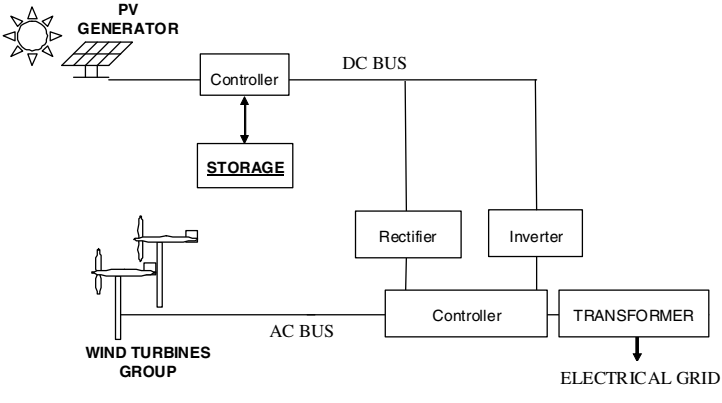


Fig. 1. Grid-connected renewable electricity generation system (photovoltaic generator, wind farm or hybrid system) with storage

Fig. 2 shows the two options of storage: hydrogen or batteries. In both cases, the storage is connected by means of a controller to the DC bus, as both techniques use DC electricity to store the electricity.

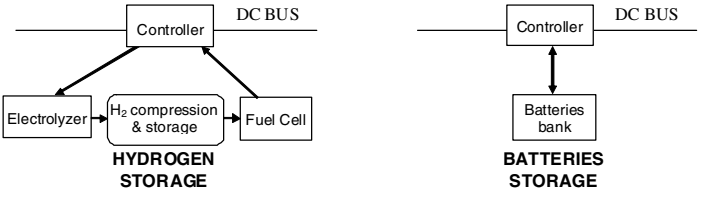


Fig. 2. Storage technologies: Hydrogen (Electrolyzer + compression + storage + fuel cell) or Batteries

2.1 Hydrogen Storage

Hydrogen storage is composed of (Fig. 2):

- Electrolyzer: it consists of two electrodes which are placed in distilled water in order to create electrolysis of water. DC electricity is applied between the electrodes and hydrogen appears at the cathode (negatively charged electrode), while oxygen appears at the anode (positively charged electrode). Different electrolyzer technologies use different electrolytes, membranes, electrodes, and catalyzers. The most important technologies are Alkaline, PEM (proton exchange membrane), and High Temperature Steam.
- Hydrogen compression and storage: hydrogen produced by the electrolyzer is compressed and led to the storage tank.
- Fuel cell: the reverse operation of the electrolyzer operation is performed by the fuel cell (reversing the electrolysis of water): hydrogen from the storage tank is combined with oxygen in the air, and it produces DC electricity. There are many technologies: Alkaline, PEM, among others [6].

Table 1 shows the main characteristics of these components, taking into account that these technologies are not fully mature. The lifespan varies from less than 15,000 h to more than 40,000 h.

Table 1. Typical characteristics of hydrogen components

Component	Capital cost (€/kW)	Efficiency
Electrolyzer	1000-1700	65-75% of HHV_{H2}
Fuel cell	1500-2500	40-50% of LHV_{H2}

The main problems of storing renewable electrical energy in hydrogen are as follows:

1. The roundtrip efficiency (electrolyzer-compression-hydrogen tank-fuel cell) is about 30%. This means that, if the electricity generated from renewable sources is

stored in hydrogen, only 30% of this electricity will be injected into the electrical grid. This is a very low value which cannot compete with batteries.

2. Acquisition cost of the components is enormous.

Results of previous studies [5] show that the electricity stored in hydrogen should be sold to the electrical grid at about 1.7 €/kWh at the present scenario (present costs and efficiencies) and about 0.4 €/kWh for the best future scenario, in which the costs for the hydrogen components are about 5% of present costs.

2.2 Batteries Storage

There are several battery technologies that could be applied in renewable electricity storage, shown in Table 2.

Table 2. Battery technologies

Battery Type	Capital cost (€/kWh)	Roundtrip Efficiency (%)	No. of cycles	Max. DOD (%)	Self-disch. (%/m)	Oper. Temp (ºC)	Energy density (Wh/kg)	O&M cost
Lead-acid (stationary convent.)	80-100	75-85	500-1000	70	1.5-5	-20 to 50	30	Very Low
Lead-acid (OPZS-OPZV)	100-250	80-90	1200-1800	80	1.5-5	-20 to 50	30	Very Low
Nickel-Cadmium	300-700	70-80	1500-3000	100	5-20	-40 to 50	50	~0
Nickel Metal Hydride	600-900	65-70	1000-1500		10-25	-40 to 50	80	~0
Lithium Ion	800-1300	95-100	1500-3000	80	1-5	-30 to 90	90-150	~0
Lithium Polymer	700-1200	90-100	600	80	2-5	-20 to 60	100-150	~0
Sodium Sulphur (NaS)	300-1000	89-92	2500-3500	100	~0	300-350	100	Medium (Heating)
Vanadium Redox (VRB)	200-700	65-75	5000-10000	80	~0	0-40	30-50	High (pumps)
Zinc Bromine (ZnBr)	250-800	70-80	4000	100	~0	0-40	70	High (pumps)
Metal Air	50-200	50	few 100		~0	-20 to 50	450–650	Low

The most used batter technology for the storage of electrical energy is lead-acid. In off-grid systems (PV-systems, hybrid systems), it is common to use the new lead-acid batteries OPZS (tubular plates) or OPZV (tubular plates with gel electrolyte, maintenance free). However, for grid-connected high-power systems, other technologies are more adequate.

Previous studies showed that sodium-sulphur (NaS) batteries are the most promising technology to store electricity in high power systems [4] because their acquisition cost is expected to be reduced to about 100 €/kWh of storage, and their lifespan can increase to 5,000 or 6,000 cycles. With these assumptions for the near future, the electrical energy stored could be sold at about 0.2 €/kWh, which will soon be a competitive price in Spain.

3 Conclusions

At the present time, batteries and hydrogen are technically suitable to store electrical energy from renewable sources; however, their acquisition cost makes these systems unprofitable. Batteries have roundtrip efficiencies from 70 to 90%, much higher than the roundtrip efficiency of the hydrogen conversion system (electrolyzer-compressor-tank-fuel cell), which is about 30%. Therefore, batteries are much better than hydrogen to store electricity. The most promising technology of batteries to store electrical energy in high power systems is NaS batteries. The expected reduction in acquisition costs and increase in lifespan of this kind of battery will enable the renewable systems with storage to compete with the systems without storage in a few short years. This will bring many benefits to the electrical grid: avoiding stability problems and producing generation curves of this type of system similar to the demand curve, compared with the generation curves of the renewable systems without storage.

Acknowledgements. This work was supported by the "Ministerio de Ciencia e Innovación" of the Spanish Government, under Project ENE2009-14582-C02-01.

References

1. Ibrahim, H., Ilinca, A., Perron, J.: Energy storage systems – characteristics and comparisons. Renewable and Sustainable Energy Reviews 12(5), 1221–1250 (2008)
2. Farber De Anda, M., Boyes, J., Torres, W.: Lessons learned from the Puerto Rico battery energy storage system. Sandia National Laboratories Technical Report SAND99-2232 (1999)
3. Kashem, M.A., Ledwich, G.: Energy requirement for distributed energy resources with battery energy storage for voltage support in three-phase distribution lines. Electric Power Systems Research 77(1), 10–23 (2007)
4. Dufo-López, R., Bernal-Agustín, J.L., Domínguez-Navarro, J.A.: Generation management using batteries in wind farms: Economical and technical analysis for Spain. Energy Policy 37(1), 126–139 (2009)
5. Bernal-Agustín, J.L., Dufo-López, R.: Hourly energy management for grid-connected wind–hydrogen systems. International Journal of Hydrogen Energy 33(22), 6401–6413 (2008)

A Method for Designing Orthogonal Trivariate Wave-Functions and Applications in Electronic Engineering

Ruifang Huang* and Zhixia Duan

Dept. of Fundamentals, Jiyuan Vocational and Technical College, Jiyuan 464650, China
{zxs123hjk,zxc123wer}@126.com

Abstract. For years, electronic engineering has donated pagers and paging services for families waiting for an organ donation. Electronic engineering is considered at the same level as electrical engineering, so that general programs are called electrical and electronic engineering. The existence of orthogonal vector-valued trivariate wavelets is discussed. A necessary and sufficient condition is presented by means of vector-valued multiresolution analysis and paraunitary vector filter bank theory. A constructive method for designing a sort of orthogonal vector-valued wavelets with short support is proposed, and their properties are investigated.

Keywords: Electronic engineering, trivariate wave-functions, functional analysis, filter banks, orthogonal, paraunitary vector-valued filters, oblique frames.

1 Introduction

Wavelet analysis is nowadays a widely used tool in applied mathematics. Computer science or computing science designates the scientific and mathematical approach in information technology and computing. The advantages of wavelets and their promising features in various application have attracted a lot of interest and effort in recent years. Already they have led to exciting applications in signal processing [1], image processing [2] and so on. Vector-valued wavelets are a sort of generalized multiwavelets [3]. Vector-valued wavelets and multiwavelets are different in the following sense. For example, prefiltering is usually required for discrete multiwavelet transforms but not necessary for discrete vector-valued wavelet transforms [4]. In real life, video images are vector-valued signals. Vector-valued wavelet transforms have been recently investigated for image processing by scholar W. Li. Chen and Huo studied biorthogonal compactly supported vector-valued wavelets with 2-scale[5]. Chen and Shi researched orthogonal[6] construction and properties of matrix-valued wavelets. Inspired by [5,6], we are ready to investigate the construction of a class of orthogonal compactly supported vector-valued trivariate wavelets with multi-scale. Similar to uniwavelets, it is more complicated and meaningful to investigate orthogonal vector-valued trivariate wavelets with 2-scale.

* Biography: female, Master, engages in applied mathematics and mathematics education.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 227–233.
springerlink.com

2 The Trivariate Multiresolution Analysis

We begin with some notations. Set $Z_+ = \{0\} \bigcup N$, s, $n, v \in N$ and $s, n, v \geq 2$, $Z^3 = \{(z_1, z_2, z_3) : z_r \in Z, r = 1,2,3\}$, $Z_+^3 = \{\{(z_1, z_2, z_3) : : z_r \in z_+, r = 1,2,3\}$. For any X, $X_1, X_2 \subset R^3$, denoting by $aX = \{ax : x \in X\}$, $X_1 + X_2 = \{x_1 + x_2 : x_1 \in X_1, x_2 \in X_2\}$, $X_1 - X_2 = \{x_1 = \{x_1 - x_2 : x_1 \in X_1, x_2 \in X_2\}$. There exist a^3 elements μ_0, $\mu_1, \cdots, \mu_{a^3-1}$ in space Z_+^3 by algebra theory such that $Z^3 = \bigcup_{\mu \in \Gamma_0} (\mu + aZ^3)$; $(\mu_1 + aZ^3) \bigcap (\mu_2 + aZ^3) = \varnothing$, where $\Gamma_0 = \{\mu_0, \mu_1, \cdots, \mu_{a^3}\}$ denotes the set of all different representative elements in the quotient group $Z^3/(aZ^3)$ and μ_1, μ_2 denote two arbitrary distinct elements in Γ_0, M is a 3×3 matrix. Set $\mu_0 = \underline{0}$, where $\underline{0}$ is the null of Z_+^3. Let $\Gamma = \Gamma_0 - \{\underline{0}\}$ and Γ, Γ_0 be two index sets.By $L^2(R^3, C^v)$, we denote the aggregate of all vector-valued functions $\hat{\lambda}(x)$, i.e., $L^2(R^3, C^v) := \{\Lambda(x) = (\hat{\lambda}_1(x), \hat{\lambda}_2(x), \cdots, \hat{\lambda}_v(x))^T : \hat{\lambda}_l(x) \in L^2(R^3),\ l = 1,2,\cdots, v\}$, where T means the transpos of a vector. For $\hat{\lambda}(x) \in L^2(R^3, C^v)$, $\|\hat{\lambda}\|$ denotes the norm of vector-valued function $\hat{\lambda}(x)$, i.e., $\|\hat{\lambda}\| := (\sum_{l=1}^{v} \int_{R^3} |\lambda_l(x)|^2 \, dx)^{1/2}$. In the below * means the transpose and the complex conjugate, and its integration is defined to be

$$\int_{R^3} \hat{\lambda}(x)\, dx = (\int_{R^3} \lambda_1(x) dx, \int_{R^3} \lambda_2(x)\, dx, \cdots\cdots, \int_{R^3} \lambda_v(x) dx)^T.$$

The Fourier transform of $\hat{\lambda}(x)$ is defined as $\hat{\hat{\lambda}}(\gamma) := \int_{R^3} \hat{\lambda}(x) \cdot e^{-ix\cdot\gamma} dx$, where $x \cdot \gamma$ denotes the inner product of real vectors x and γ. For any $\hat{\lambda}(x), \hbar(x) \in L^2(R^3, C^v)$, their *symbol inner product* is defined by

$$\langle \hat{\lambda}(\cdot), \hbar(\cdot) \rangle = \int_{R^3} \hat{\lambda}(x) \hbar(x)^* dx, \tag{1}$$

where * means the transpose and the complex conjugate, and I_v denotes the $v \times v$ identity matrix. A sequence $\{\hat{\lambda}_l(t)\}_{l \in Z^3} \subset U \subseteq L^2(R^3, C^v)$ is called an orthonormal collection of the subspace U, if the following condition is satisfied

$$\langle \hat{\lambda}_j(\cdot), \hat{\lambda}_k(\cdot) \rangle = \delta_{j,k} I_v, \quad j, k \in Z^3. \tag{2}$$

Definition 1. We say that $\hat{\lambda}(x) \in U \subseteq L^2(R^3, C^v)$ is an orthogonal vector-valued function of the subspace U if its translations $\{\hat{\lambda}(x-u)\}_{u \in Z^3}$ is an orthonormal collection of the subspace U, i.e.,

$$\langle \hat{\lambda}(\cdot - n), \hat{\lambda}(\cdot - k) \rangle = \delta_{n,k} I_v, \ n, k \in Z^3. \tag{3}$$

Definition 2[5]. A sequence $\{\hat{\lambda}_v(x)\}_{v \in Z^3} \subset U \subseteq L^2(R^3, C^v)$ is called an orthonormal basis of U, if it satisfies (2), and for any $F(x) \in U$, there exists a unique sequence of $v \times v$ constant matrices $\{Q_k\}_{k \in Z^3}$ such that

$$F(x)=\sum_{u\in Z^3} Q_u \hat{\lambda}_u(x), \quad x\in R^3. \tag{4}$$

Definition 3[5]. A vector-valued multiresolution analysis of $L^2(R^3,C^v)$ is a nested sequence of closed subspaces $\{Y_l\}_{l\in z}$ such that (i) $Y_l\subset Y_{l+1},\forall\, l\in Z$; (ii) $\bigcap_{l\in z} Y_l=\{0\}$; $\bigcup_{l\in z} Y_l$ is dense in $L^2(R^3,C^v)$, where 0 is the zero vector of space $L^2(R^3,C^v)$; (iii) $\phi(t)\in Y_0$ if and only if $\phi(a^l t)\in Y_l$, ; (iv) there is $\hat{\lambda}(x)\in Y_0$ such that the sequence $\{\hat{\lambda}(x-u),u\in z^3\}$ is an orthonormal basis of subspace Y_0.

By Definition 2 and Definition 3, we obtain $\hat{\lambda}(x)$ satisfies the below equation:

$$\hat{\lambda}(x)=\sum_{k\in Z^3} P_k \hat{\lambda}(ax-k) \tag{5}$$

where $\{P_k\}_{k\in Z^3,}$ is a finite supported sequence of $v\times v$ constant matrices, i.e., $\{P_k\}_{k\in Z^3,}$ has only finite non-zero terms, and the others are zero matrices. By taking the Fouries transform for (5), and assuming $\hat{\hat{\lambda}}(\xi)$ is continuous at zero, we have

$$\hat{\hat{\lambda}}(a\xi)=\mathcal{P}(\xi)\hat{\hat{\lambda}}(\xi), \quad \xi\in R^3, \tag{6}$$

$$a^3\mathcal{P}(\xi)=\sum_{k\in Z^3} P_k\cdot\exp\{-ik\xi\}. \tag{7}$$

Let W_j $(j\in Z)$ denote the orthocomplement subspace of Y_j in Y_{j+1} and there exist a^3-1 vector-valued functions $G_s(t)\in L^2(R^3,C^v)$, $s=1,2,\cdots,a^3-1$ such that their translations and dilations form a Riesz basis of W_j, i.e.,

$$W_j=clos_{L^2(R^3,C^v)}(\text{span}\{G_s(a^j t-u):\ s=1,2,\cdots,a^3-1,\ u\in Z^3\}),\ j\in Z. \tag{8}$$

Since $G_s(x)\in W_0\subset Y_1$, $s=1,2,\cdots,a^3-1$ there exist sequences $\{B_k^{(s)}\}_{k\in Z^3}$ such that

$$G_s(t)=\sum_{u\in Z^3} B_u^{(s)} F(at-u),\ \ s=1,2,\cdots,a^3-1. \tag{9}$$

$$a^3\mathcal{B}^{(s)}(\xi)=\sum_{u\in Z^3} B_u^{(s)}\exp\{-iu\xi\}. \tag{10}$$

$$\hat{G}_s(a\xi)=\mathcal{B}^{(s)}(\xi)\hat{h}(\xi),\ \ s=1,2,3,\cdots,a^3-1,\ \ \xi\in R^3. \tag{11}$$

If $\hat{\lambda}(x)\in L^2(R^2,C^v)$ is an orthogonal vector one, then it follows from (3) that

$$\left\langle \hat{\lambda}(\cdot),\hat{\lambda}(\cdot-u)\right\rangle=\delta_{0,u}I_v, \quad u\in Z^3. \tag{12}$$

We say that $G_s(x)\in L^2(R^3,C^v)$, $s\in A=\{1,2,\cdots,a^3-1\}$ are orthogonal vector-valued wavelet functions associated with the vector-valued scaling function $\hat{\lambda}(x)$, if

$$\left\langle \hat{\lambda}(\cdot-k),G_s(\cdot-u)\right\rangle=O, \quad s=1,2,\cdots,a^3-1,\ k,u\in Z^3, \tag{13}$$

and $\{G_s(x-u), s\in A,\ u\in Z^3\}$ is an orthonomal basis of W_0. Thus we have

$$\left\langle G_r(\cdot),\ G_s(\cdot-n)\right\rangle = \delta_{r,s}\delta_{0,n}I_u,\ \ r,s=1,2,\cdots,a^3-1;\ \ n\in Z^3. \tag{14}$$

Lemma 1[6]. Let $\hat{\lambda}(x)\in L^2(R^3,C^v)$. Then $\hat{\lambda}(x)$ is an orthogonal vector-valued function if and only if

$$\sum_{k\in Z^3}\hat{\hat{\lambda}}(\xi+2k\pi)\hat{\hat{\lambda}}(\xi+2k\pi)^* = I_v,\quad \xi\in R^3. \tag{15}$$

Lemma 2. If $\hat{\lambda}(x)\in L^2(R^3,C^v)$, defined by (5), is an orthogonal vector-valued trivariate scaling function, then for $a=2$, $\forall\, u\in Z^3$, we have the following equalities,

$$\sum_{\sigma\in Z^2}P_\sigma(P_{\sigma+4u})^* = 4\delta_{0,u}I_v. \tag{16}$$

$$\sum_{\iota=0}^{3}\mathcal{P}(\xi+\sigma_\iota\pi)\mathcal{P}(\xi+\sigma_\iota\pi)^* = I_v,\ \ \xi\in R^3,\ \ \sigma_\iota\in Z^3. \tag{17}$$

Proof. By substituting (5) into the relation (12), for $\forall\, k\in Z^2$, we obtain that

$$\delta_{0,k}I_v = \left\langle F(\cdot-k),F(\cdot)\right\rangle = \sum_{l\in Z^3}\sum_{u\in Z^3}\int_{R^3}P_l\hat{\lambda}(2x-2k-l)\hat{\lambda}(2x-u)^*(P_u)^*dx$$

$$= \frac{1}{4}\cdot\sum_{l\in Z^3}\sum_{u\in Z^3}P_l\left\langle\hat{\lambda}(\cdot-2k-l),\hat{\lambda}(\cdot-u)\right\rangle(P_u)^* = (1/8)\sum_{u\in Z^3}P_u(P_{u+8k})^*.$$

3 Construction of a Class of Vector-Valued Wavelets

In the following, we begin with considering the existence of a class of compaly supported orthogonal vector-valued wavelet functions for the case of $a=2$.

Theorem 1. Let $\hat{\lambda}(x)\in L^2(R^3,C^v)$ defined by (5), be an orthogonal vector-valued scaling function. Assume $G_s(x)\in L^2(R^3,C^v),\ s=1,2,3.$ Then $G_s(x)$ are orthogonnal vector-valued wavelet functions associated with $\hat{\lambda}(x)$ if and only if

$$\sum_{\iota=0}^{3}\mathcal{P}(\xi+\sigma_\iota\pi)\mathcal{B}^{(s)}(\xi+\sigma_\iota\pi)^* = O,\quad \sigma_\iota\in Z^3, \tag{18}$$

$$\sum_{\iota=0}^{3}\mathcal{B}^{(r)}(\xi+\sigma_\iota\pi)\mathcal{B}^{(s)}(\xi+\sigma_\iota\pi)^* = \delta_{r,s}I_v,\quad \sigma_\iota\in Z^3, \tag{19}$$

where $r,\ s\in 1,2,3,\ \xi\in R^3$, or equivalently,

$$\sum_{l\in Z^3}P_l(B^{(s)}_{l+2u})^* = O,\quad s=1,2,3,\ u\in Z^3; \tag{20}$$

$$\sum_{l\in Z^3} B_l^{(r)}(B_{l+2u}^{(s)})^* = 8\delta_{r,s}\delta_{0,u}I_{v,}\ r,s=1,2,3,\ u\in Z^3. \tag{21}$$

Proof. Firstly, we prove the necessity. By Lemma 1 and (6), (11) , we have

$$O=\sum_{u\in Z^3}\hat{\mathfrak{h}}(2\xi+2u\pi)\hat{G}_s(2\xi+2u\pi)^*=\sum_{t=0}^{3}\mathcal{P}(\xi+\sigma_t\pi)\mathcal{B}^{(s)}(\xi+\sigma_t\pi)^*,\quad \sigma_t\in Z^3.$$

It follows from formula (14) and Lemma 1 that

$$\delta_{r,s}I_v=\sum_{u\in Z^2}\hat{G}_r(2\xi+2u\pi)\hat{G}_s(2\xi+2u\pi)^*=\sum_{t=0}^{3}\mathcal{B}^{(r)}(\xi+\sigma_t\pi)\mathcal{B}^{(s)}(\xi+\sigma_t\pi)^*,\sigma_t\in Z^2.$$

Next, the sufficiency will be proven. From the above calculation, we have

$$\sum_{u\in Z^3}\hat{\mathfrak{h}}(2\xi+u\pi)\hat{G}_s(2\xi+u\pi)^*=\sum_{\sigma=0}^{3}\mathcal{B}^{(r)}(\omega+\sigma\pi/2)\mathcal{B}^{(s)}(\omega+\sigma\pi/2)^*=\delta_{r,s}I_v.$$

$$\left\langle \mathfrak{h}(\cdot),G_s(\cdot-k)\right\rangle=\frac{1}{\pi^3}\int_{[0,\pi/2]^3}\sum_{u\in Z^3}\hat{\mathfrak{h}}(2\omega+2u\pi)\hat{G}_s(2\xi+2u\pi)^*e^{8ik\xi}d\xi=O,$$

$$\left\langle G_r(\cdot),G_s(\cdot-k)\right\rangle=\frac{1}{\pi^2}\int_{[0,\pi/2]^2}\sum_{u\in Z^2}\hat{G}_r(2\omega+2u\pi)\hat{G}_s(2\xi+2u\pi)^*\cdot e^{4ik\xi}d\xi=\delta_{0,k}\delta_{r,s}I_v.$$

Thus, it follows that $F(t)$ and $G_s(t), s=1,2,3$ are mutually orthogonal ones, and $\{G_s(t),\ s=1,2,3\}$ are a family of orthogonal vector-valued functions. This shows the orthogonality of $\{G_s(\cdot-u),\ s=1,\ 2,3\}_{u\in z^2}$. Similar to [6,Proposition 1], we can prove its completeness in W_0.

Theorem 2. Let $\mathfrak{h}(x)\in L^2(R^3,C^v)$ be a 5-coefficient compactly supported orthogonal vector-valued scaling functions satisfying the following refinement equation:

$$\mathfrak{h}(x)=P_0\mathfrak{h}(2x-\mu_0)+P\mathfrak{h}(2x-\mu_1)+\cdots+P_3\mathfrak{h}(2t-\mu_3),\quad \mu_0,\cdots\mu_3\in Z^3.$$

Assume there exists an integer , $0\le\ \le 3$, such that $(8I_v-P(P)^*)^{-1}P(P)^*$ is a positive definite matrix. Define $Q_s(s=1,2,3)$ to be two essentially distinct Hermitian matrice, which are all invertible and satisfy

$$(M_s)^3=[8I_v-P(P)^*]^{-1}P(P)^*. \tag{22}$$

$$\begin{cases} B_j^{(s)}=M_sP_j, & j\ne\mu\ , \\ B_j^{(s)}=-(M_s)^{-1}P_j, & j=\mu\ , \end{cases}\quad s=1,2,3,\ \ j\in\{\mu_0,\mu_1,\mu_2,\mu_3\}. \tag{23}$$

Then $G_s(t)$ $(s=1,2,3)$, defined by (24), are orthogonal wavelets with $\mathfrak{h}(x)$:

$$G_s(x)=B_0^{(s)}\mathfrak{h}(2x)+B_1^{(s)}\mathfrak{h}(2x-1)+\cdots+B_4^{(s)}\mathfrak{h}(2x-4). \tag{24}$$

Example 1. Let $\hat{\lambda}(x) \in L^2(R^3, C^3)$ be 4-coefficient orthogonal vector-valued scaling function satisfy the following equation:

$$\hat{\lambda}(x) = P_0\hat{\lambda}(2x-\mu_0) + P_1\hat{\lambda}(2x-\mu_1) + \cdots + P_3\hat{\lambda}(2x-\mu_3), \quad \mu_0, \cdots \mu_3 \in Z^3$$

where $P_3 = O, \quad P_0(P_3)^* = O, \quad P_0(P_0)^* + P_1(P_1)^* + P_2(P_2)^* + P_3(P_3)^* = 8I_3.$

$$P_0 = \begin{pmatrix} 4\sqrt{2} & 4\sqrt{2} & 0 \\ -4 & \frac{8\sqrt{2}}{3} & 8 \\ 0 & 0 & \frac{16\sqrt{3}}{3} \end{pmatrix}, \; P_1 = \begin{pmatrix} 8 & 0 & 0 \\ 0 & \frac{4\sqrt{2}}{3} & 0 \\ 0 & 0 & \frac{8\sqrt{3}}{3} \end{pmatrix}, P_2 = \begin{pmatrix} 4\sqrt{2} & -4\sqrt{2} & 0 \\ 4 & \frac{8\sqrt{2}}{3} & -8 \\ 0 & 0 & \frac{16\sqrt{3}}{3} \end{pmatrix},$$

Suppose $=1$. By using (22), we can choose

$$M_1 = diag(4\sqrt{2}, 8/\sqrt{53}, 2\sqrt{2}), \quad M_2 = diag(4\sqrt{2}, 8/\sqrt{53}, -2\sqrt{2}).$$

By applying formula (23), we get that $B_3^{(1)} = B_3^{(2)} = O,$

$$B_0^{(1)} = \begin{pmatrix} 4 & 4 & 0 \\ -\frac{4\sqrt{53}}{53} & \frac{8\sqrt{106}}{159} & \frac{8\sqrt{53}}{53} \\ 0 & 0 & \frac{4\sqrt{6}}{3} \end{pmatrix}, \quad B_2^{(2)} = \begin{pmatrix} 4 & -4 & 0 \\ \frac{4\sqrt{53}}{53} & \frac{8\sqrt{106}}{159} & -\frac{8\sqrt{53}}{53} \\ 0 & 0 & -\frac{4\sqrt{6}}{3} \end{pmatrix}$$

$$B_2^{(1)} = \begin{pmatrix} 4 & -4 & 0 \\ -\frac{4\sqrt{53}}{53} & \frac{8\sqrt{106}}{159} & -\frac{8\sqrt{53}}{53} \\ 0 & 0 & \frac{4\sqrt{6}}{3} \end{pmatrix}, \quad B_0^{(2)} = \begin{pmatrix} 4 & 4 & 0 \\ -\frac{4\sqrt{53}}{53} & \frac{8\sqrt{106}}{159} & \frac{8\sqrt{53}}{53} \\ 0 & 0 & -\frac{4\sqrt{6}}{3} \end{pmatrix}.$$

$B_1^{(1)} = -diag(8\sqrt{2}, 4\sqrt{106}/3, 16\sqrt{6}/3), \; B_1^{(2)} = diag(-8\sqrt{2}, -4\sqrt{106}/3, 16\sqrt{6}/3).$

Applying Theorem 2, we obtain that $G_t(t) = B_0^{(t)}\hat{\lambda}(2x-\mu_0) + B_1^{(t)}\hat{\lambda}(2x-\mu_1) + \cdots + B_4^{(t)}\hat{\lambda}(2x-\mu_3), t = 1,2,3$ are orthogonal vector-valued wavelet functions associated with the orthogonal vector-valued scaling function.

4 Conclusion

Electronic engineering is a discipline that utilizes the behavior and effects of electrons for the production of electronic devices, systems, or equipment. In many parts of the

world, electronic engineering is considered at the same level as electrical engineering, so that general programs are called electrical and electronic engineering. A necessary and sufficient condition on the existence of a class of orthogonal vector-valued binary wavelets is presented. An algorithm is proposed.

References

[1] Zhang, N., Wu, X.: Lossless of color masaic images. IEEE Trans. Image Delivery 15(6), 1379–1388 (2006)
[2] Efromovich, S., et al.: Data-Diven and Optimal Denoising of a Signal and Recovery of its Derivation Using Multiwavelets. IEEE Trans. Signal Processing 52(3), 628–635 (2004)
[3] Shen, Z.: Nontensor product wavelet packets in $L_2(R^s)$. SIAM Math. Anal. 26(4), 1061–1074 (1995)
[4] Xia, X.G., Suter, B.W.: Vector-valued wavelets and vector filter banks. IEEE Trans. Signal Processing 44(3), 508–518 (1996)
[5] Chen, Q., Huo, A.: The research of a class of biorthogonal compactly supported vector-valued wavelets. Chaos, Solitons & Fractals 41(2), 951–961 (2009)
[6] Chen, Q., Shi, Z.: Construction and properties of orthogonal matrix-valued wavelets and wavelet packets. Chaos, Solitons & Fractals 37(1), 75–86 (2008)

Photovoltaic Grid Parity in Spain

Rodolfo Dufo-López and José L. Bernal-Agustín

Department of Electrical Engineering, University of Zaragoza,
Calle María de Luna 3, 50018, Spain
{rdufo,jlbernal}@unizar.es

Abstract. Photovoltaic (PV) grid parity can be considered as the point at which PV electricity is generated at the same price as the electricity of the electrical grid is generated (which is a mix of different generation sources) or as the point at which PV electricity becomes cost-competitive with conventional grid-supplied electricity. Depending on which definition of grid parity is used, the year when it is reached may vary. A study of grid parity, considering both definitions, has been carried out with regard to Spain.

Keywords: Photovoltaic, grid parity, Spain.

1 Introduction

The countries that have contributed most to the growth of the photovoltaic (PV) market in the world are Germany, Japan, Spain and USA.

Spain has very abundant solar resources (irradiation over horizontal surface 1100–1900 kWh/m^2/yr), with maximum values in the south and minimum in the far north of the country.

Some countries have applied different PV promotion programs. In Spain, PV generation was highly subsidized. Electricity generated by PV systems of nominal power lower than 100 kW was sold to the electrical grid at 45 c€/kWh before September 2008. Then, subsidies were revised and a tariff of 34 c€/kWh (roof retrofit mount) or 32 c€/kWh (other installation types) was applied to systems of less than 20 kW. The Royal Decree 1578/2008 [1] changed the subsidies policy, reducing up to 2.5% per quarter if the total PV power required in the country is more than the pre-assigned quota. In these moments, the new PV systems have not subsidies. However, in some years, PV electricity will reach "grid parity." In this paper, we study PV grid parity for the case of Spain.

2 Grid Parity

Photovoltaic grid parity can be considered as the point at which PV electricity is generated at the same price as the electricity generated by the electrical grid (which is a mix of different generation sources) or as the point at which PV electricity becomes cost-competitive with conventional grid-supplied electricity [2].

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 235–239.
springerlink.com

In [3], "Grid Parity Year" is defined as the time point in the future when the electricity from grid-connected PV will be able to compete with the grid electricity generated from conventional fuels.

In [4], grid parity price is defined as the total cost to consumers of PV electricity being equal to retail grid electricity price. Taking into account the U.S. average electricity price of 11.6 c$/kWh in 2008, the authors concluded the installed PV system cost should be reduced to $2/W in order to reach grid parity. However, the results are highly dependent on the assumption of a 6% of interest rate.

In [2], the assumptions of reports suggesting that solar grid parity in sunnier areas would be reached by 2012 [5] are rejected, given that this would require a reduction in installed PV system costs from $8.73/W to $2/W in three years, unlikely based on historical experience curves.

3 PV System Cost

The total installed PV system cost includes the modules, the inverter, the support structure, the electrical panel and protections, the cables and structure anchor, as well as the engineering, the mechanical and electrical installation, and VAT (value-added tax). In Table 1, the different costs of a 5.58 kWp (5 kW inverter) are shown, taking into account 0.528 €/W for the inverter (which can be considered a stable cost) and different costs for the PV modules, from 2.37 €/Wp (current average in Europe) to 0.5 €/Wp (average which could be reached in 2030 [6], though the cheapest PV modules could reach this value much earlier). A value of 1 €/Wp could be the average for PV module cost in 2020 [6] or before.

The total specific cost of the PV system varies from 4.6 €/Wp (for PV module cost of 2.37 €/Wp, currently) to 2.39 €/Wp (for PV module cost of 0.5 €/Wp). Taking into account the prices in Table 1 along with the retail grid electricity price and its expected inflation, the grid parity will be studied.

4 Retail Grid Electricity Price

In Spain, the electrical market was liberalized and any consumer may choose his own electricity provider and negotiate the tariff freely. However, there is a regulated tariff for low-power consumers who do not want to negotiate with the manufacturers. Currently, most of the private consumers have chosen the electricity tariff regulated by law. However, most of the companies have chosen a tariff in the liberalized market.

Anyway, for private consumers, the tariff would be similar. Currently (October 2011), the regulated tariff is 17.4 c€/kWh, including taxes.

From 2007 to the end of 2011, the regulated tariff has increased by 56%, with an average of 11.7% annually. In 2011 so far, the increase has been 13%. An additional 6% was planned to be increased in October 2011; however, the Spanish government has delayed it in the end.

Table 1. Cost of the PV system of 5580 Wp (5000 W inverter)

	PV Module Cost (€/Wp)				
	2.37	**2**	**1.5**	**1**	**0.5**
PV module (5580 Wp)	13225 (61%)	11160 (57%)	8370 (50%)	5580 (40%)	2790 (25%)
Support structure	2000 (9%)	2000 (10%)	2000 (12%)	2000 (14%)	2000 (18%)
Inverter (5000 W) (0.528 €/W)	2640 (12%)	2640 (13%)	2640 (16%)	2640 (19%)	2640 (23%)
Electrical panel and protections	800 (4%)	800 (4%)	800 (5%)	800 (6%)	800 (7%)
Other materials (cables, anchor, etc.)	500 (2%)	500 (3%)	500 (3%)	500 (4%)	500 (4%)
Mechanical installation	1620 (7%)	1620 (8%)	1620 (10%)	1620 (11%)	1620 (14%)
Electrical installation and engineering	960 (4%)	960 (5%)	960 (6%)	960 (7%)	960 (8%)
TOTAL PV installed cost (€)	21745	19680	16890	14100	11310
VAT (18%)	3914	3542	3040	2538	2036
TOTAL PV installed cost, including taxes (€)	**25659**	**23222**	**19930**	**16638**	**13346**
Total specific cost (€/Wp)	**4.60**	**4.16**	**3.57**	**2.98**	**2.39**
Total specific cost (€/W)	**5.13**	**4.64**	**3.99**	**3.33**	**2.67**

If we take into account the electricity prices in the electrical market (pool), in September 2011, the final price of the market (price for direct consumers and marketers, which can be considered the generation cost) is 6.8 c€/kWh, while one year before it was 5.1 c€/kWh. Thus there was an increase of 33% last year. However, during the previous year, the increase was only about 10%.

Taking into account that the Spanish energy system is characterized by its high dependence on imports (80% of energy consumption is from imported sources) and that oil, natural gas and coal are set to peak and go into decline within the next decade, we can conclude that the electrical tariff will continue growing.

In next section we will consider values for the average annual electricity inflation (increase in the electricity tariff), g (%), during the lifetime of the system.

5 Cases of Study

Five different scenarios for the PV system installed cost will be studied (Table 1). The system will be installed in Zaragoza (Spain); the irradiation over horizontal surface is 1614 kWh/m^2/yr. The PV modules are fixed south-oriented, and their slope is 35º (optimal), so the irradiation over the PV modules surface is 1834 kWh/m^2/yr. The annual O&M cost considered is 2% of total PV installed cost. The interest rate considered is 4% and the general inflation 2% (to be applied to the O&M annual cost and the replacement of the inverter in about 10 years). The system lifetime is 25 years, as is the PV module lifetime. A performance ratio of 75% will be taken into

account (90% efficiency of the inverter and about 15% losses in PV modules due to temperature, dirt, etc). The regulated electricity tariff is Pr_{2011} = 17.4 c€/kWh in 2011, including taxes, and the average annual electricity price inflation considered will be g = 3%, 5% and 10%. The grid parity will be reached in n years from 2011, when the price of the electricity makes it so that the Net Present Value (*NPV*) of the PV system is equal to zero ($Pr_{NPV=0}$).

$$Pr_{NPV=0} = Pr_{2011} \cdot (1+g/100)^n \quad (1)$$

$$n = \log(Pr_{NPV=0}/Pr_{2011})/\log(1+g/100) \quad (2)$$

Results are shown in Table 2. For each case, the table show 1) the price of PV electricity that would make *NPV* = 0, $Pr_{NPV=0}$; also shown are 2) the price of PV electricity that would make it so amortization occurs in 10 years, and 3) the year in which grid parity is reached (2011+n).

Table 2. Results for the case of average annual electricity inflation g = 3%, 5%, 10%. Grid parity considering the regulated tariff (Pr_{2011} = 17.4 c€/kWh).

		Total PV System Installed Specific Cost (€/Wp)				
		4.60	**4.16**	**3.57**	**2.98**	**2.39**
g = 3%	Price of PV electricity so that *NPV* = 0, $Pr_{NPV=0}$ (c€/kWh)	22.3	19.8	17.2	14.5	11.9
	Price of PV electricity so that amortization is in 10 years ($c€_{2011}$/kWh)	40.5	36.1	30.9	25.7	20.5
	Grid parity in year	2020	2016	< 2011	< 2011	< 2011
g = 5%	Price of PV electricity so that *NPV* = 0, $Pr_{NPV=0}$ (c€/kWh)	17.4	15.4	13.4	11.3	9.3
	Price of PV electricity so that amortization is in 10 years ($c€_{2011}$/kWh)	36.4	32.5	27.8	23.1	18.4
	Grid parity in year	2011	< 2011	< 2011	< 2011	< 2011
g = 10%	Price of PV electricity so that *NPV* = 0, $Pr_{NPV=0}$ (c€/kWh)	8.8	7.8	6.8	5.7	4.7
	Price of PV electricity so that amortization is in 10 years ($c€_{2011}$/kWh)	27.9	24.8	21.2	17.6	14.1
	Grid parity in year	< 2011	< 2011	< 2011	< 2011	< 2011

If we take into account the final price of the market (price for direct consumers and marketers, which can be considered the generation cost), 6.8 c€/kWh in September 2011, instead of the regulated tariff electricity price, the results for the year in which grid parity is reached (2011+n) are quite different (Table 3).

Table 3. Grid parity considering the final price of the market (6.8 c€/kWh)

	Total PV System Installed Specific Cost (€/Wp)				
	4.60	**4.16**	**3.57**	**2.98**	**2.39**
Annual average electricity price inflation g = 3%	2052	2047	2042	2037	2030
Annual average electricity price inflation g = 5%	2031	2028	2025	2022	2018
Annual average electricity price inflation g = 10%	2014	2013	2011	< 2011	< 2011

6 Conclusions

In this paper, a study of photovoltaic grid parity in Spain has been discussed.

If we consider grid parity the point at which PV electricity is generated at the same price as the electricity generated by the electrical grid (which is a mix of different generation sources), grid parity is highly dependent on the PV system cost and on the annual average electricity inflation rate. If we consider grid parity as the point at which PV electricity becomes cost-competitive with conventional grid-supplied electricity, this point has already been reached in most scenarios.

Acknowledgements. This work was supported by the Ministerio de Ciencia e Innovación (Innovation and Science Ministry) of the Spanish Government, under Project ENE2009-14582-C02-01.

References

1. Spanish Ministry of Industry. Royal Decree 1578/2008, of September 26, rewards of the activity of production of electricity by solar photovoltaic installations after the deadline for the continued payment of wages of Royal Decree 661/2007 (2008)
2. Yang, C.J.: Reconsidering solar grid parity. Energy Policy 38, 3270–3273 (2010)
3. Bhandari, R., Stadler, I.: Grid parity analysis of solar photovoltaic systems in Germany using experience curves. Solar Energy 83, 1634–1644 (2009)
4. Stavy, M.A.: Financial worksheet for computing the cost (c/kWh) of solar electricity generated at grid connected photovoltaic PV generating plants. Journal of Solar Energy Engineering 124(3), 319–321 (2002)
5. Hutchinson, A.: Solar panel drops to $1 per watt: is this a milestone or the bottom for silicon-based panels? Popular Mechanics (February 26, 2009)
6. Nowak, S., Aulich, H., Bal, J.L., Dimmler, B., Garnier, A., Jongerden, G.: The European photovoltaic technology platform. In: IEEE Photovoltaic Energy Conversion Conferences, vol. 2, pp. 2485–2489 (2006)

Developing a HTML5-Based Real-Time Monitoring System for Wireless Sensor Networks

Si-Ho Cha

Dept. of Multimedia Science, Chungwoon University
Daehakgil-25, Hongseong-eup, Hongseong-gun, Chungnam, South Korea
shcha@chungwoon.ac.kr

Abstract. A wireless sensor network (WSN) consists of spatially distributed autonomous sensors to monitor physical or environmental conditions, such as temperature, sound, vibration, pressure, humidity, motion or pollutants and to cooperatively pass their data through the network to a main location. Web architecture offers a high degree of interoperability and a low entry barrier to combine the physical world with Web contents. Today WSNs in electronic engineering are used in many industrial and consumer applications and the need for Web-based real-time data from the physical world collected by a large number of sensors and actuators is becoming more and more in electronic engineering. HTML5 standards offer an ideal harmony between performance and functionality. This paper describes a Web-based real-time monitoring system with special circuit for WSNs in electronic engineering using WebSocket, Canvas, Geolocation, Chart, and various other new techniques of the HTML5 standards. The proposed system can monitor and control WSNs in electronic engineering with the Web in real-time manner.

Keywords: Wireless Sensor Network, HTML5, WebSocket, Geolocation, Canvas, Real-Time Monitoring, Electronic Engineering.

1 Introduction

Wireless sensor networks (WSNs) in electronic engineering are used in many industrial and consumer applications, such as industrial process monitoring and control, environment and health monitoring, and so on. A WSN consists of spatially distributed autonomous sensors to monitor physical or environmental conditions, such as temperature, sound, vibration, pressure, humidity, wind speed, motion or pollutants and to cooperatively pass their data through the network to a main location [1]. WSNs in electronic engineering gather data about the real world and the data have been increasingly used in various applications ranging from battlefield surveillance to health care. The development of WSNs in electronic engineering was motivated by military applications such as battlefield surveillance; today such networks are used in many industrial and consumer applications, such as industrial process monitoring and control, environment and health monitoring, and so on [1]. With recent advantages in sensing and wireless communications, it is possible to construct a network consisting of a large number of heterogeneous sensing devices. However, their inherent

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 241–245.
springerlink.com

limitations like power constraints, limited bandwidth, distributed coordination, and ad hoc deployment lead to a number of challenges in the deployment and monitoring of a network [2]. Therefore, the monitoring system of WSNs in electronic engineering is required to control of sensor activity in with real-time easily and simply. However, as described previously, developing a real-time monitoring system for WSNs in electronic engineering is challenging because sensor nodes have limited energy and computational resources.

Future Web applications will increasingly require real-time data from the physical world collected by a myriad of sensors and actuators. Currently, integration of such devices requires customized solutions because of the lack of widely adopted protocols for devices [3]. The advantages of the Web are quick and easy updates, reach anybody and anywhere, available anytime and cross-platform, always up-to-date, zero install, and reduce business costs. The Web of Things [4] is about re-using the Web standards to connect the quickly expanding eco-system of embedded devices built into everyday smart objects. In the Web of Things, embedded devices or computers are connected by fully integrating them to the Web.

Therefore, this paper extends the Web of Things paradigms to support more scalable and interoperable WSN monitoring applications and to support real-time bi-directional messaging. WebSocket [5] enables the HTML5-based monitoring system to monitor WSNs in electronic engineering directly over the Web. The WebSocket is a web technology providing for bi-directional, full-duplex communications channels, over a single Transmission Control Protocol (TCP) socket giving clients and servers a simple way to communicate over a persistent stream without the need for third-party plugins or hacks. Additionally, it serves a purpose for Web applications that require real-time bi-directional communication. The proposed system uses Canvas API, Geolocation API, Chart API, and various other new techniques of the HTML5 standards to represent data from WSNs in electronic engineering easily and efficiently. In the proposed HTML5-based motoring system, sensor nodes can be interacted with and controlled through WebSocket API and their status and functions can be monitored on Canvas through Geolocation and Chart APIs. We can perform the deployment of sensor nodes and the configuration management of WSNs in electronic engineering through mash-up of the Geolocation API and Google Maps API.

The rest of the paper is organized as follows. Section 2 discusses the related work, and Section 3 introduces the development of the proposed HTML5-based real-time monitoring system. Finally, in Section 4, conclusions are made including the future research.

2 Related Work

Many efforts are currently going towards networking smart things from the physical world (e.g. RFID, wireless sensor and actuator networks, embedded devices) on a larger scale. The Internet of Things has mainly focused on establishing connectivity in a variety of challenging and constrained networking environments and the next logical objective is to build on top of network connectivity by focusing on the application layer [6]. The Web of Things is a vision inspired from the Internet of

Things where everyday devices and objects, i.e. objects that contain an embedded device or computer, are connected by fully integrating them to the Web. Examples of smart devices and objects are WSNs, ambient devices, household appliances, RFID tagged objects, etc. Unlike in the many systems that exist for the Internet of Things, the Web of Things is about re-using the Web standards to connect the quickly expanding eco-system of embedded devices built into everyday smart objects. Well-accepted and understood standards and blueprints (such as URI, HTTP, REST, Atom, etc.) are used to access the functionality of the smart objects [7]. The concept of the Sensor Web is a type of sensor network that is especially well suited for environmental monitoring [8]. The phrase the Sensor Web is also associated with a sensing system which heavily utilizes the World Wide Web. OGC's Sensor Web Enablement (SWE) framework defines a suite of web service interfaces and communication protocols abstracting from the heterogeneity of sensor (network) communication [9].

Most of the existing Web-based approaches use HTTP to transport data between devices. HTTP is half duplex, and functions as a request-response protocol in the client-server computing model. In HTTP, a browser acts as a client, while an application running on a computer hosting a Web site functions as a server [10]. Ajax [11] is a group of interrelated web development techniques used on the client-side to create asynchronous web applications. With Ajax, web applications can send data to, and retrieve data from, a server asynchronously without interfering with the display and behavior of the existing page. Data is usually retrieved using the XMLHttpRequest object [12].

WebSocket [5] technology provides similar functionality to regular connections to the web with some additional protocol overhead while multiplexing several WebSocket services over a single TCP port. Before the implementation of WebSocket, such bi-directional communication was only possible using Comet channels; however, a Comet is not trivial to implement reliably, and due to the TCP handshake and HTTP header overhead, it may be inefficient for small messages.

3 Design and Implementation

The core idea of the proposed system is to enable Web-based interactions with sensor nodes by using HTML5 standards for monitoring WSNs. Fig. 1 shows the proposed HTML5-based monitoring system for WSNs. The functionality of sensor nodes can be manipulated using WebSocket on HTTP in a real-time manner.

HTML5-based Web clients are applications that run in a Web browser, communicate to the Web server and use HTML5 (and JavaScript) to process and present data and interact with the user and the Web server.

The Web server of the proposed system uses Canvas API, Geolocation API, RGraph API, and Google Maps API to process and present data and interacts with administrators and the Web server. Canvas API provides scripts with a resolution-dependent bitmap canvas, which can be used for rendering graphs, or other visual images on the fly. The Canvas enables the Web browser to natively manipulate, compose, and layer image, video, and chart data. The Geolocation API [13] defines a high-level interface to location information associated only with the device hosting

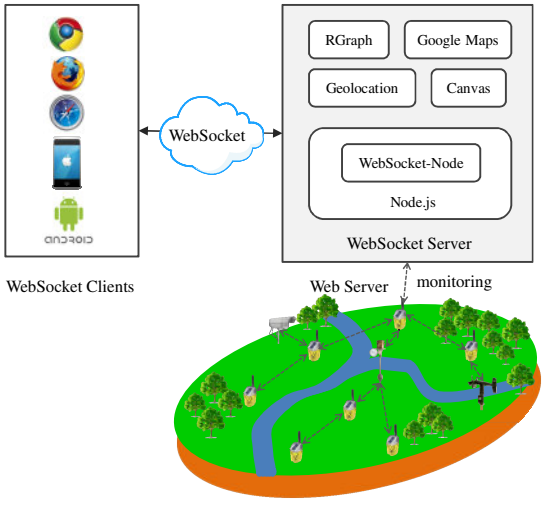


Fig. 1. System Architecture

the implementation, such as latitude and longitude. The API itself is agnostic of the underlying location information sources. Common sources of location information include Global Positioning System (GPS) and location inferred from network signals such as IP address, RFID, WiFi and Bluetooth MAC addresses, and GSM/CDMA cell IDs, as well as user input.

The geographical location of sensor/actuator nodes can be identified on the Google Maps by mash-up the Geolocation API with Google Maps API [14]. By using the Google Maps API, it is possible to embed Google Maps into the Web server, on to which server specific sensor data can be overlaid. RGraph [15] is a HTML5 JavaScript Charts library to represent the data gathered from sensors. RGraph is supports over 20 different types of JavaScript based charts. Using the HTML5 Canvas tag, RGraph creates these charts in the Web browser, meaning quicker pages and less Web server load [15].

Our HTML5-based monitoring system for WSNs in electronic engineering can send real-time commands to actuators (*turn on LEDs*), retrieve sensor data (*get water/wind gauges*), handle events (*trigger alarm*) or change state (*change camera's sampling frequency*) directly through WebSocket API. Notifications are sent when an event occurs from sensor nodes to the WebSocket server and then the WebSocket server sends those notifications the WebSocket clients in real time. The monitoring actions are sent promptly from the WebSocket server to actuators and/or sensor nodes to respond the notifications. And sensor data is collected periodically and sent to a sink to be processed and/or stored.

WebSocket requires its own backend application to communicate with (server side). Therefore, Node.js is used to develop WebSocket server. Node.js [16] is a platform built on Chrome's JavaScript runtime for easily building fast, scalable network applications. Node.js uses an event-driven, non-blocking I/O model that makes it lightweight and efficient, perfect for data-intensive real-time applications

that run across distributed devices. Node.js itself doesn't have support for WebSocket but there are already some plugins that implement WebSocket protocols. In this paper I have used the WebSocket-Node [17] which is a pure JavaScript implementation of the WebSocket protocol versions 8 and 13 for Node.js.

4 Conclusion and Future Study

In this paper, a HTML5-based real-time monitoring system for WSNs in electronic engineering has been proposed. By utilizing the WebSocket, Canvas, Geolocation, Chart, and various other new techniques provided by HTML5, the Web-based monitoring system can be implemented to control and monitor sensor activity in with real-time easily and simply.

Future research activities will be devoted to implement and evaluate the proposed system in real WSN environments.

References

1. Wikipedia: Wireless Sensor Networks, http://en.wikipedia.org/wiki/Wireless_sensor_network
2. Vaidya, D., Peng, J., Yang, L., Rozenblit, J.W.: A Framework for Sensor Management in Wireless and Heterogeneous Sensor Network. In: Proceedings of the 12th IEEE International Conference and Workshops on the Engineering of Computer-Based Systems, ECBS 2005, pp. 4–7 (2005)
3. Trifa, V., Guinard, D., Davidovski, V., Kamilaris, A., Delchev, I.: Web Messaging for Open and Scalable Distributed Sensing Applications. In: Benatallah, B., Casati, F., Kappel, G., Rossi, G. (eds.) ICWE 2010. LNCS, vol. 6189, pp. 129–143. Springer, Heidelberg (2010)
4. Tridium: The Web of Things. White Paper, http://www.tridium.com/galleries/white_papers/WP-SedonaWeb.pdf
5. Hickson, A.: The WebSocket API, http://dev.w3.org/html5/websockets/
6. Guinard, D., Trifa, V., Wilde, E.: Architecting a Mashable Open World Wide Web of Things. Computer and Information Science, vol. 663. Institute for Pervasive Computing, ETH Zurich (2010)
7. Wikipedia: Web of Things, http://en.wikipedia.org/wiki/Web_of_Things
8. Delin, K.: Sensor Webs in the Wild. Wireless Sensor Networks: A Systems Perspective. Artech House (2005)
9. Botts, M., Percivall, G., Reed, C., Davidson, J.: OGC® Sensor Web Enablement: Overview and High Level Architecture. In: Nittel, S., Labrinidis, A., Stefanidis, A. (eds.) GSN 2006. LNCS, vol. 4540, pp. 175–190. Springer, Heidelberg (2008)
10. Swamy, N.R., Mahadevan, G.: Event Driven Architecture using HTML5 Web Sockets for Wireless Sensor Networks. White Papers, Planetary Scientific Research Center (2011)
11. Garrett, J.J.: Ajax: A New Approach to Web Applications (2005), http://AdaptivePath.com
12. Ullman, C.: Beginning Ajax. Wrox Press (2007)
13. Popescu, A.: Geolocation API Specification. W3C Editor's Draft (2012)
14. Google Maps API, https://developers.google.com/maps/
15. RGraph: HTML5 Javascript charts library, http://www.rgraph.net/
16. Node.js, http://nodejs.org/
17. WebSocket-Node, https://github.com/Worlize/WebSocket-Node

Development of Monitoring Systems by Serial Port Based on VB Software and Electronic Communication

Yu Tu[1], Chao Wang[2], and Peng Du[3]

[1,2] Software Institute
[1,2] Henan Polytechnic Institute
[3] Nanyang City thoroughfare Highway Planning Survey and Design
Henan 473009, China
Hygy2011@126.com

Abstract. Two Real-Time monitoring systems were developed with MSComm control in VB6.0 software by accomplishing communication protocols of Fischertechnik-board and Siemens S7-200 PLC. The monitoring systems can fulfill such functions as manual, automatic, stop controlling and reset motion to the manipulator. Moreover, the system can demonstrate an analogue value varying with the time by curve drawing. Animation simulation with Flash playing can synchronize with the physical model motion perfectly.

Keywords: VB, Computer, Monitoring system, Interface Board, Development.

1 Monitoring System Hardware Components

The project monitoring model by Fischer COMBI-PACK PROFI COMPUTING 30330-type structures. Model gripper open limit switch, arm up, down limit switches, limit switches are located in the fuselage right rear gripper arm junction with the left bracket, the body chassis department. Fischertechnik Intelligent Interface board monitoring system interface board by Fischer as the next crew, through the RS232 serial port connected to the serial port on a PC, smart interface board receives model input signals coming from the sensor, the output is driven model of the motor movement.

In the PLC serial port monitoring system, using S7-200 for the next crew, CPU 226CN main unit and EM223 unit receives the sensor signal that electrical and model. The main unit through the serial port connected to the serial port on the PC, to receive command signals from the monitoring system to control the output of the signal driving circuit through the DC motor control model of action. EM235 work unit used to collect the scene of the sensor analog signals.

2 Fischer Intelligent Serial Interface Board Computer Monitoring System

2.1 Robot Motion Control

Reset action process is running robot shown in Figure 1.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 247–251.
springerlink.com

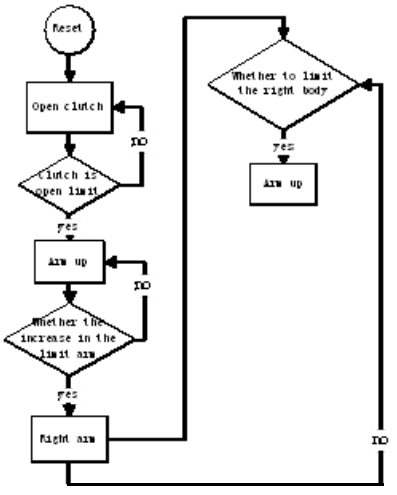


Fig. 1. Robot reset action process

Control, press the button to automatically run, first determine the current location, if not the initial position, is automatically reset after the implementation of automatic operation; If you have a reset, the direct implementation of automatic run: 3 seconds left after the body, arms down; down to lower bit (the location of the material) when the gripper is closed; delay 3 seconds after the arm up; up first after the machine limit The body continues to turn left; delay 3 seconds down the arm; down to lower bit gripper open; gripper arms open up to the time limit; arm right up to the limit when the body; body right to the right Limit, the workflow ends.

2.2 VB Monitoring System

In the VB control interface, to form (Form) by adding a communication control (MSComml) to send and receive the code via the serial port; a timer control (Timer) to periodically send the code to the serial port; two check boxes Control mr (0), mr (1), mr (2), ml (0), ml (1), ml (2) as a manual adjustment of 3 to move the motor. 4 shows four stroke SHAPE switch inputs and state; 1 PICTUREBOX used to draw the analog input, analog dynamically displays the current value; three progress bar to display the operating status; the corresponding number of command buttons. A database control ADODB, used to store analog data. VB-based intelligent interface board Fischertechnik interface control program shown in Figure 2, serial programming using MSComm communication control, communications processing using event-driven approach, the interception of the input signal for the main event in the OnComm, three-way motor control, mainly in Timer Event driven.

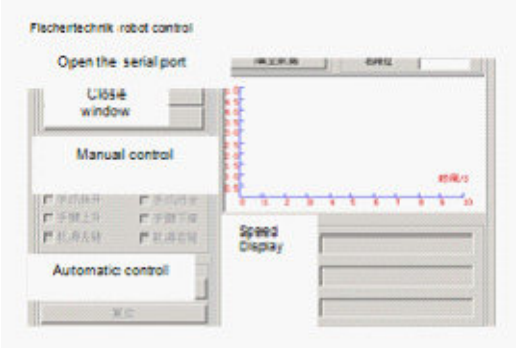


Fig. 2. Fischertechnik interface board monitoring system interface

3 Siemens S7-200PLC Computer Serial Port Monitoring System

3.1 Robot Motion Control

PLC controlled robot I / O wiring principle as shown in Figure 3, the robot automatically moves the process shown in Figure 4. Press Auto Run button, first determine the current location, if not the initial position, it can not exercise, have to run after reset; 4 seconds left after the body, arms down; down to lower bit (the location of materials), the gripper closed grab objects; delay 2 seconds after the arm rose, rising limit after the first body to turn left; delay 3 seconds down the arm; down to lower bit gripper open, put down objects; gripper open to the limit, arms up; arm up to the limit right after the body; body right to the right limit, Workflow ends.

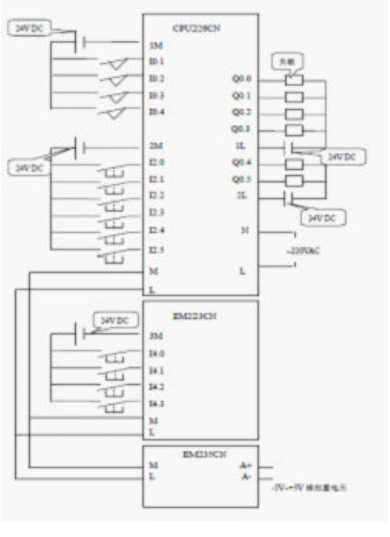


Fig. 3. PLC controlled robot I / O wiring diagram

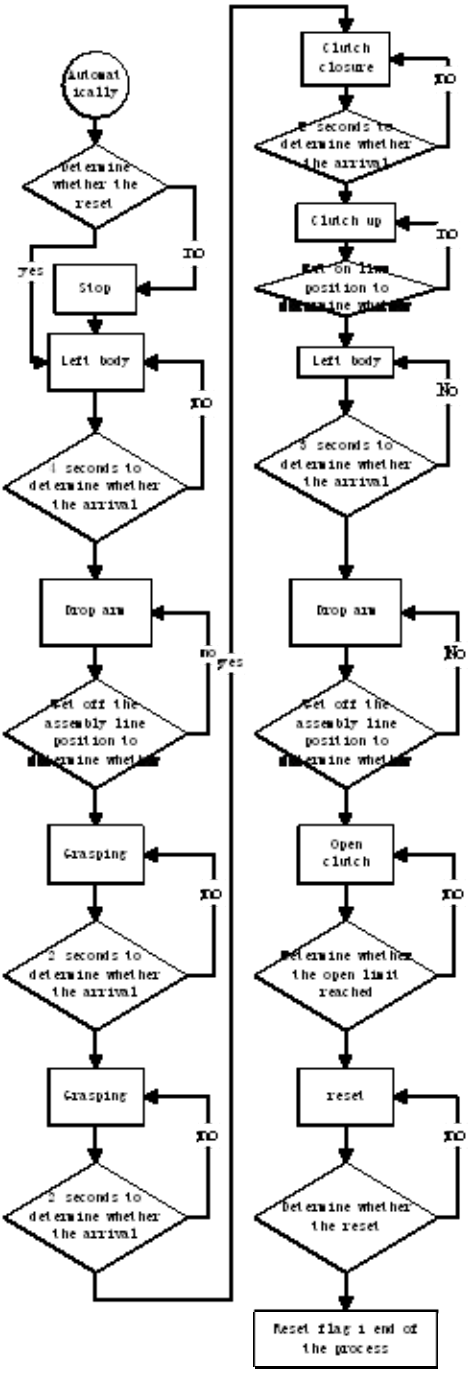


Fig. 4. Robot PLC controlled automatic operation flow chart

3.2 Serial Manipulator Control System PLC

Add a number in the interface command button, respectively, to move the control button 6, a stop button, an auto-run button, a reset button, two monitor control buttons, there are two port selection control button; 3 SHOCKWAVEFLASH control, used to play the corresponding FLASH, real-time display the current robot state; a MSCOMM control; 2 TIMER control, TIMER1 is used to run the PLC internal data read command, TIMER2 to add a variable to record the current frame FLASH Player Number; a PICTUREBOX control, to display the analog curve; there SHAPE control group 2, respectively, displays the current PLC's I / O status; a LABEL, the text to real-time display with the current state of the robot; a BSE controls.

PLC programming, in the hardware control buttons for each input point, an intermediate parallel register, the host computer through the serial port write command to the PLC sends commands to the middle register is ON, so that you can indirectly control the completion of the PLC. System sends a read command, write command, are sent to the serial port to an array of hex data.

Run the actual results show that the two VB-based computer running very successful serial port monitoring system. Have achieved a physical model of the robot hand control, automatic operation, work site analog input dynamic sensor curve drawing, the current state of detection. Among them, Fischer board monitoring system with the progress bar the way to display the model of the current operation progress, S7-200PLC monitoring system is used animation synchronized simulation robot physical action, with a more intuitive way to realize the PLC I / O ports and the robot physical action of the real-time monitoring, detection in the analog interface board control than the Fischer system more stable.

References

1. Morgan, C., Pardoe, D., Smith, N.: Toward a Standard for Remote MicroScope Control Systems. Scanning 20(2), 110–116 (1998)
2. Meng, K.: Development of Robot-monitoring Systems by Serial Port Based on VB Software. Nanyang Polytechnic (July 2009)
3. Shen, D.: Application of VB6.0 to achieve the robot's control Fischertechnik. Electrical Engineering Technology 36(1), 67–68 (2007)
4. Li, X., Zou, X.: Visual Basic Serial communication. Mine Automation (2), 38–39 (2002)
5. Ignatius, F., Chen, Y.: Visual Basic RS232 serial communication and control. Beijing Tsinghua University Press (2002)
6. High, Y.: MSCOMM based on the control environment using VB serial communication. Luliang College 24(2), 31–33 (2008)

Using Binary Decision Diagram for Test Generation of Power Supply Noise in Digital Circuits

Zhongliang Pan and Ling Chen

School of Physics and Telecommunications Engineering, South China Normal University, Guangzhou 510006, China
panz@scnu.edu.cn

Abstract. The scaling of circuit design technology to ultra deep submicron has increased the sensitivity of circuit technology to various noise mechanisms such as power supply noise and leakage current noise, etc. The propagation of noise can happen through various modes like the power supply and the substrate to a sensitive receptor. The power supply noise can reduce the actual voltage level reaching a device, results in the loss of signal integrity. The type of noise may also cause the degradation of switching speed, timing faults, or logic functional failures. In this paper, a new approach is presented for the test generation of power supply noise in digital circuits, the approach is based on the binary decision diagrams. The test vectors of power supply noise are generated by building and operating several binary decision diagrams, such as the binary decision diagrams respectively corresponding to the normal circuits and the circuits whose signal lines have been assigned special logic values. A lot of experimental results show that if there are test vectors for the power supply noise in a circuit, then the test vectors can be generated by the approach proposed in this paper.

Keywords: Digital circuits, power supply noise, test generation, fault detection, binary decision diagrams.

1 Introduction

The trend of VLSI technology scaling is predicted to continue in recent years, the feature sizes continue to shrink while the clock frequencies continue to increase. The circuit density has been increased aggressively as the availability of interconnects for millions of transistors. Because of the higher switching speeds and increased aspect ratios of interconnect, the noise problem in the circuit is becoming to be more and more severe.

The noise in circuit is the any unwanted deviation in the voltages and/or currents at various nodes within a circuit. If a noise acts against a stable logic level on a circuit node, it can transiently destroy logical information carried by the node [1]. For example, the noise can cause an incorrect machine state stored in a latch, therefore it can result in circuit functional failure.

The power supply noise is the difference between the local voltage references of the driver and the receiver. The increased amount of current on power supply lines

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 253–258.
springerlink.com

causes a raise in IR drop on voltage references [2]. This makes the gate more highly sensitive to noise spikes. The ground bounce is the phenomenon of shifting ground and power voltage levels, for example between the IC internal ground and power levels and those at the board. The power supply noise (including IR-drop and ground bounce etc) has become more and more significant in the design and manufacturing of VLSI circuits, it needs to be efficiently taken into consideration.

For the reduction of power supply noise, Tuuna [3] proposed an approach for the simultaneous reduction of both crosstalk noise and power supply noise being caused by on-chip buses, the approach uses the analytical resistance, inductance and capacitance models. Fiori [4] analyzed the propagation of the switching noise in system-on-chip, it was shown that the substrate voltage bounce due to switching noise is able to be reduced when the low impedance power supplies are replaced by high impedance power supplies. Taparia [5] presented a localization method for the power supply noise by using on-chip active inductors.

For the testing about power supply noise, Feng [6] studied the detection of delay defects under the constraints of power supply noise effects, the pseudo-functional testing techniques was used to produce the test vectors. Jing [7] investigated the two types of low cost power supply noise models, and shown their effects in delay testing. Meng [8] studied an automatic test pattern generation scheme that effectively performs compressible low capture power, the scheme can reduce the launch cycle switching activity.

In this paper, a new approach is presented for the test generation of power supply noise in digital circuits, the approach is based on the binary decision diagrams.

2 Test Generation of Power Supply Noise

The scaling of circuit technology to ultra deep submicron has increased the sensitivity of circuit technology to various noise mechanisms such as crosstalk noise, power supply noise, leakage current noise, etc. The switching activity in a VLSI circuit is the main source of noise. The propagation of noise can happen through various modes like the power supply and the substrate to a sensitive receptor.

For example, for VLSI circuit, the switching activity increases as the gate density increases. This may lead to increase current density and voltage drop along the power supply net, and produce simultaneous switching activity within a small time interval, therefore increase the instantaneous current. The power supply noise reduces the actual voltage level reaching a device, results in the loss of signal integrity. The type of noise may also cause the degradation of switching speed, timing faults, or logic functional failures. Therefore, it is necessary to detect the noise in order to insure the reliability of circuits.

2.1 The Model of Power Supply Noise

There are two components in power supply noise: inductive ΔI noise, and power net IR voltage drop. The power supply noise (PSN) is given by the following equation.

$$PSN = L \cdot \frac{dI}{dt} + IR$$

The first component of above equation is the inductive ΔI noise, it depends on the rate of change of the instantaneous current. The second component of above equation is IR voltage drop, which is caused by the instantaneous current through the resistive power and ground network.

Make use of the circuit analytical and simulation such as SPICE simulation, we can obtain that the power supply noise depends on the switching activity inside the circuit, hence it is dependent on the input pattern of circuit. In general, in order to stimulate the maximal effect of power supply noise, it is needed to maximize both the current change rate $\frac{dI}{dt}$ and the peak current drawn.

Therefore, for a given circuit under test, the detection approach for power supply noise is made up of the following two steps: (1) Produce a test pattern pair, i.e. (V_1, V_2), which is applied to the primary inputs of the circuit under test, so that the effect of power supply noise is stimulated and is propagated to the primary outputs of the circuit. (2) Use the set of pattern pair (V_1, V_2) to perform the circuit simulation in order to accurately measure the effect of power supply noise. In the following, a detection approach based on BDD is given for the power supply noise.

2.2 Test Generation by BDD

The binary decision diagrams (BDD) use a directed acyclic graph to represent the logic Boolean functions [9,10]. There are two types of nodes in a BDD, i.e., the leaf nodes also called terminal nodes, and nonterminal nodes. A leaf node is labeled 0 or 1 which is corresponding to the value of the Boolean function for a specific assignment to the input variables. A nonterminal node is labeled with a variable, it has two edges directed toward two children, the two edges are named 0-branch and 1-branch [11]. The 1-branch corresponds to the case where the variable is assigned 1, the 0-branch corresponds to the case where the variable is assigned 0.

For the logic functions $g = x_1 \cdot x_2 \cdot \bar{x}_3$ and $h = x_1 \cdot \bar{x}_3 + x_1 \cdot x_2 + x_2 \cdot \bar{x}_3$, their BDD is shown in the Fig.1(a) and Fig.1(b), respectively. In the Fig.1, the 0-branch is shown by a dashed line, the 1-branch is shown by a solid line. Several BDDs can be joined into a single graph that consists of the BDDs sharing their subgraphs, the graph is called shared binary decision diagram (SBDD). The SBDD is an improved BDD, it has the following property: many functions are represented compactly and simultaneously by the sharing isomorphic subgraphs. An example of shared binary decision diagram is shown in Fig.1(c).

For example, for the BDDs corresponding to the function *g* and *h*, the total number of nodes in this two BDDs is 11. The SBDD of logic functions *g* and *h* is given in Fig.1(c), the number of nodes in the SBDD is 7. Therefore, the SBDD not only can represent the logic functions of *g* and *h* simultaneously, but also the number of nodes in the SBDD is cut down.

In the following, we use binary decision diagram to produce the test pattern pair (V_1, V_2) for power supply noise in digital circuits. The circuit with normal logic functionality is called normal circuit, therefore there are not noise in the normal circuit. The normal circuit whose a signal line has been assigned a value 0 or 1 is called circuit_0, or circuit_1, respectively.

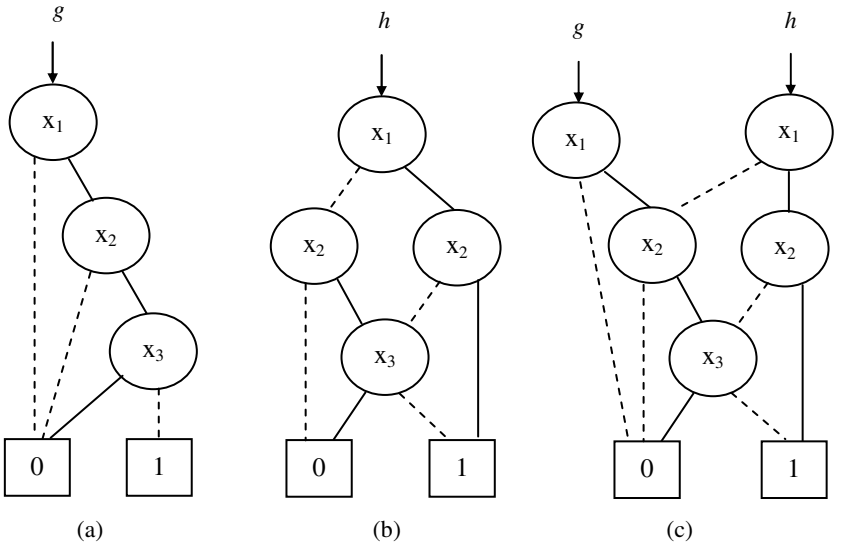


Fig. 1. The SBDD of two functions

The approach in this paper for producing the test pattern pair (V_1, V_2) consist of following four steps:

Step 1, the BDD for the normal circuit is built, the BDD is called as normal BDD.

Step 2, the BDD for the circuit_0 is built, the BDD is called as circuit_0 BDD, which describes the functionality of the circuit whose a signal line has been assigned a value 0.

Step 3, the BDD for the circuit_1 is built, the BDD is called as circuit_1 BDD, which describes the functionality of the circuit whose a signal line has been assigned a value 1.

Step 4, the BDD called as test_0 BDD is built, it is the XOR operation of the normal BDD and circuit_0 BDD. Each input assignment that leads to the leaf node with attribute value 1 is a vector V_1. Similarly, the BDD called as test_1 BDD is built, it is the XOR operation of the normal BDD and circuit_1 BDD. Each input assignment that leads to the leaf node with attribute value 1 is a vector V_2.

2.3 Experimental Results

The approach in this paper for using binary decision diagram for test generation of power supply noise in digital circuits, has been implemented in C++ language, and run them on a personal computer with 3.0GHz and 512MB memory. We have performed a lot of experiments for ISCAS'85 benchmark circuits C432, C499, C1355, C1908, C2670, C5315 and C6288.

For these experiments of the benchmark circuits, we use following method to get the BDD corresponding to a circuit. In general, a circuit is made up of many circuit blocks. A circuit block implements a specific logic function. The logic function of

whole circuit is expressed as a sequence of operations on the Boolean functions realized by these circuit blocks. This needs a fast method to build the BDD of whole circuit by using these BDDs of all circuit blocks.

The operator *ite* is defined by following mode: For given three logic functions f, g, and h, the *ite*(f, g, h) = f·g + $\bar{f}$ ·h. The *ite* operator can be used for realizing all Boolean operations with two variables. For example, $\bar{f}$ = *ite*(f,0,1), f+g = *ite*(f,1,g), f·g = *ite*(f,g,0), f ⊕g = *ite*(f, $\bar{g}$,g), where ⊕ denotes XOR operation.

Suppose F, G, and H is the BDD corresponding to logic functions f, g and h, respectively. The following method can be used for computing the BDD corresponding to *ite*(F, G, H). The Shannon's decomposition of the F is expressed by the following equation:

$$F = w \cdot F_w + \overline{w} \cdot F_{\overline{w}}$$

The $w \in \{x_1, x_2, \cdots, x_n\}$, the F_w denotes $F|_{w=1}$, the $F_{\overline{w}}$ denotes $F|_{w=0}$. The F_w and $F_{\overline{w}}$ are F evaluated at w=1 and w=0 respectively. From the above equation, we can get

$$ite(F, G, H) = ite(v, ite(F_v, G_v, H_v), ite(F_{\overline{v}}, G_{\overline{v}}, H_{\overline{v}}))$$

The $v \in \{x_1, x_2, \cdots, x_n\}$. The terminal cases of the recursion equation are:

$$ite(1, F, G) = ite(0, G, F) = ite(F, 1, 0) = F.$$

The BDD corresponding to a circuit is built by a forward process.

For the test generation of power supply noise, the main task is to produce test pattern pair (V_1, V_2), subsequently, the circuit analytical is performed by using SPICE simulation tools. We randomly select 30 power supply noises from these ISCAS'85 benchmark circuits to produce the test pattern pair. The BDDs corresponding to these ISCAS'85 benchmark circuits are built by using the *ite* operator.

The experimental results show that if there are test pattern pairs for a power supply noise, then the test pattern pairs can be produced by the approach in this paper. The time of producing all test pattern pairs of a power supply noise is less than one minute for all these benchmark circuits. One advantage of the approach in this paper is that all test pattern pairs of a power supply noise can be generated, therefore, the test set with minimal number of test pattern pairs can be obtained by using the approach.

3 Conclusions

The power supply noise may cause the degradation of switching speed, timing faults, or logic functional failures, thus it is necessary to test the noise in order to insure the reliability of circuits. In this paper, a new approach based on the binary decision diagrams is presented for the test generation of power supply noise in digital circuits, the test pattern pair is generated by building and operating several binary decision diagrams. In the future, some work needs to be done for the high effective building method of binary decision diagrams.

Acknowledgments. This work were supported by National Natural Science Foundation of China (No.61072028), Guangdong Province & Chinese Ministry of

Education Cooperation Project of Industry, Education, and Academy (No. 2009B090300339).

References

1. Bohannon, E., Urban, C., Pude, M.: Passive and active reduction techniques for on-chip high-frequency digital power supply noise. IEEE Trans. on VLSI 18(1), 157–161 (2010)
2. Junxia, M., Mohammad, T.: Layout-aware critical path delay test under maximum power supply noise effects. IEEE Trans. on CAD 30(12), 1923–1934 (2011)
3. Tuuna, S., Tenhunen, H.: Skewing-based method for reduction of functional crosstalk and power supply noise caused by on-chip buses. IET Computers and Digital Techniques 6(2), 114–124 (2012)
4. Fiori, F.: On the use of high-impedance power supplies to reduce the substrate switching noise in system-on-chips. Microelectronics Reliability 52(1), 282–288 (2012)
5. Taparia, A., Banerjee, B., Viswanathan, T.R.: Power-supply noise reduction using active inductors in mixed-signal systems. IEEE Trans. on VLSI 19(11), 1960–1968 (2011)
6. Feng, Y., Xiao, L., Qiang, X.: Pseudo-functional testing for small delay defects considering power supply noise effects. In: IEEE International Conference on Computer-Aided Design, pp. 34–39 (2011)
7. Jing, W., Walker, D.M., Xiang, L., Ananta, M.: Modeling power supply noise in delay testing. IEEE Design and Test of Computers 24(3), 226–234 (2007)
8. Meng, W., Jiun, H., Xiaoqing, W., Kohei, M.: Power supply noise reduction for at-speed scan testing in linear-decompression environment. IEEE Trans. on CAD 28(11), 1767–1776 (2009)
9. Keren, O.: Reduction of average path length in binary decision diagrams by spectral methods. IEEE Trans. on Computers 57(4), 520–531 (2008)
10. Lhotak, O., Curial, S., Amaral, J.N.: An optimal encoding to represent a single set in an ROBDD. IEEE Trans. on Computers 59(4), 574–575 (2010)
11. Stefan, B.: Power indices of simple games and vector-weighted majority games by means of binary decision diagrams. European Journal of Operational Research 210(2), 258–272 (2011)

Generation of Test Vectors with Low Power by Co-evolution Algorithm for Digital Circuits

Ling Chen and Zhongliang Pan

School of Physics and Telecommunications Engineering, South China Normal University, Guangzhou 510006, China
panz@scnu.edu.cn

Abstract. The digital circuits should be tested thoroughly in order to find the faults due to manufacturing defects or errors. The generation of test vectors is one of main task of circuit test, the test vectors can be obtained by finding appropriate logic assignments to the circuit primary inputs such that the given faults can be detected. The power consumption during test process may be higher than that power consumption during normal operations of circuits. Therefore it is necessary to design the test vectors with low power. In this paper, a new method is proposed for the generation of test vectors with low power. The method makes use of the co-evolution algorithms, first of all, the coding of an individual corresponding to the circuit primary inputs is given, secondly, the populations consisting of a lot of individuals are built, the better individuals corresponding to the test vectors with low power are obtained by applying the evolutionary operations to the populations. A lot of experimental results show that the test generation method in this paper is able to produce the test vectors with lower power, and can obtain higher fault coverage.

Keywords: Digital circuits, test vectors, low power, faults detection, evolutionary algorithms.

1 Introduction

The power consumption of digital circuits has become a major concern in recent year. In general, a circuit being designed must be simultaneously low power and high performance. The reduction of circuit power dissipation is a design goal even for non-portable circuits, because the excessive power dissipation may result in the increased cooling costs and potential reliability problems [1,2]. On the other hand, the circuits should be tested thoroughly, the purpose of test is to discover any faults caused due to manufacturing defects or errors. Sometimes, the heat dissipated during circuit test is higher than that during normal operations of circuits. This results in that the circuit may be damaged. Therefore, it is necessary to design the test vectors with low power.

In the aspect of the low power of built-in self-test (BIST), Yang [3] used linear feedback shift registers (LFSRs) in built-in self-test to reduce the excessive power dissipation associated with scan-based designs, where two split LFSRs were used to

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 259–264.
springerlink.com

reduce the amount of the switching activity. Abu [4] investigated a modified linear feedback shift register (LFSR), which can reduce the number of transitions at the inputs of the circuit under test by using a bit-swapping approach.

In the aspect of the low power test of system on chip (SoC), Rau [5] aimed at intellectual property cores and SoC, studied the shift-in power and compression ratio in low-cost automatic test equipment environment. Hong [6] studied the scan architecture for both low power testing and test volume compression, a graph-based heuristic was proposed to partition the scan chains into several segments under SoC test environment. Chun [7] proposed a scan flip-flop reordering method for low power test, the method can compress the test data.

In the aspect of the low power test of the digital circuits with scan chain structure, Mishra [8] studied a modified design method of a scan flip-flop, which eliminates the power consumption due to unnecessary switching in the combinational circuit during scan shift. Kothe [9] studied a modified multiple scan chains architecture and an optimization strategy that aim at reduced power consumption during scan test.

For the low power test of the circuits with general structure, Moghaddam [10] studied a method that is not clocking a high proportion of scan chains during both scan shift and test response capture, the method can simultaneously reduce test data volume and test power in at speed delay test utilizing clock gating. Meng [11] investigated a low power test pattern generator, which minimized the peak power consumption during the shift and capture cycles for the test of faults.

In this paper, a new method is presented for the generation of test vectors with low power, the method makes use of the co-evolution algorithms.

2 Power Dissipation in Digital Circuits

The total power dissipation in CMOS circuits can be divided into static dissipation, short circuit dissipation, leakage dissipation and dynamic power dissipation. The static power is associated with maintaining the logic values of internal circuit nodes between the switching events. The short circuit power dissipation caused by short circuit current during switching and power dissipated by leakage currents. The dynamic power dissipation is the dominant component of total power dissipation.

The dynamic power is due to the short circuit current charging and discharging of load capacitance during output switching. The dynamic power is expended usefully, because it is associated with the switching of logic states, which is central to performing logic operations. The dynamic power is proportional to the $C \cdot V_{dd} \cdot \Delta V \cdot \beta \cdot f$, where the C denotes the total capacitance, the V_{dd} denotes the supply voltage, the ΔV denotes the voltage swing, the β denotes the expected number of transitions per clock cycle, and the f denotes the clock frequency.

The dynamic power is in direct proportion to the complexity of the logic, the rate of computation, and the switching activity of nodes in circuits. It can be cut down by scaling the supply voltage, adjusting the speed of circuits, and making use of low voltage swing signaling techniques.

3 Generation of Test Vectors with Low Power

3.1 Power Consumption of Test Vectors

The circuit test is able to detect and verify the correctness of functions for the manufactured circuits. The process of circuit test is as follows. The test vectors are applied to the primary inputs of the circuit under test, the output response of the circuit is compared with the expected output response. The circuit under test is considered good (i.e. there is not faults in the circuit) if the output responses match. Therefore, the test quality depends upon the thoroughness of test vectors. The generation and evaluation of test vectors is one of the main subjects in the circuit test.

The circuits are usually designed for implementing logic functions and carrying out the normal operations. It is needed to consider power consumption during test process, otherwise the higher power consumption during test process may results in that the circuits are damaged. The following two aspects may be used in circuit test: the peak power dissipation and the average power dissipation. The peak power dissipation gives the maximum sustained power in a circuit. The average power dissipation is able to accurately estimate the average switched capacitance at the internal nodes in the circuit. In general, in order to compute the average power dissipation, first of all, the circuit simulation is carried out to determine the current waveforms from the supply voltage for a large number of circuit input vectors. Secondly, the average power dissipation is computed by determining the average current from the supply.

For a circuit, the power consumption P during a test vector is applied to the primary inputs of circuit is given by the equation:

$$P = 0.5 \cdot C \cdot V_{dd}^2 \cdot f \cdot \sum_{i=1}^{M} \sum_{j=1}^{N} S(j) \tag{1}$$

The M is the length of test vector, the N is the number of the gates in the circuit, the $S(j)$ is the switching activity factor of the output of the gate j.

3.2 Generation of Test Vectors

In the following, we use co-evolution algorithm to generate the test vectors with low power. The co-evolution algorithm is a type of search algorithms, which imitates the biological co-evolution that is a series of reciprocal changes in two or more interacting species. There are two main forms of co-evolution algorithm: cooperative co-evolution algorithm and competitive co-evolution algorithm [12,13]. For the cooperative co-evolution algorithm, the individuals of several different species are combined before they are evaluated with respect to the evolution target. For the competitive co-evolution algorithm, the individuals compete with each other individuals in the population, for instance, the fitness of individuals are evaluated through competition with other individuals.

The process for the generation of test vectors with low power is given in the following Algorithm 1. For a circuit with a number of primary inputs, we divide the primary inputs into four parts, each of part is corresponding to a subpopulation. The four subpopulations are named as $B(m,n)$, m=1 to 4. The n represents the generation

of evolutions. Assume the number of individuals in each subpopulation is N_1, N_2, N_3 and N_4, respectively.

Algorithm 1

Step 1. Set parameter n=0; Randomly produce an initial population consisting of (N_1+N_2+N_3+N_4) individuals, and divide the initial population into four subpopulation $B(1,n)$, $B(2,n)$, $B(3,n)$ and $B(4,n)$.

Step 2. For each subpopulation B(1,n) to B(4,n), perform following operations: For each individual in a subpopulation, carry out the fault simulation, and compute the power P of the individual by the equation (1), therefore obtain the fitness of each individual, the fitness is computed by expression 1/(P+0.01). Here, the fitness of an individual (i.e., subpopulation member) is obtained by combining it with the current best individuals of the remaining (temporarily frozen) subpopulations.

Step 3. Perform the genetic operations for the individuals in $B(i,n)$, i=1 to 4, respectively, such as selection, crossover and mutation operations.

Step 4. Carry out fault simulation to compute the fitness of each individuals in $B(i,n)$, i=1 to 4.

Step 5. Randomly remove one individual from the current subpopulation $B(i,n)$, i=1 to 4. Add the best individual in the previous subpopulation to the current one.

Step 6. If the termination conditions are satisfied, then stop the algorithm, otherwise, set n : =(n+1), go to Step 3.

In the Algorithm 1, the coding of an individual in subpopulation is as follows. For instance, the number of primary inputs corresponding the subpopulation B(1,n) is N_1. Each individual Z in the B(1,n) is represented by a binary string ($z_1\, z_2 \cdots z_n$), where z_i=0 or 1 for i=1,2,$\cdots$, N_1.

3.3 Experimental Results

The method in this paper for the generation of test vectors with low power, has been implemented in C++ language, and run it on a personal computer with 3.0GHz and 512MB memory. We have performed a lot of experiments for ISCAS'89 benchmark circuits s349, s382, s420, s641, s820, s838, s1196, s1423, s1488, s9234.

The following parameters are used in the Algorithm 1: the maximal evolution generation is 1200; the roulette wheel selection scheme and three-point crossovers are used; the crossover rate is 0.65; the mutation rate is 0.003. Besides, a population that does not contain any subpopulations is used for the circuits s349, s382 and s1488. The population that contains two subpopulations is used for other benchmark circuits.

The experimental results show that the Algorithm 1 can produce the test sets of each benchmark circuit, and the fault coverage is 100%. Besides, for the test vector generation, we also carry out a lot of experiments by using conventional random test algorithm, this algorithm is setting each primary input to the value 0 or 1 randomly to produce the input vector of circuits. If the input vector can detect a fault in circuit, then it is chosen as a test vector.

We have compared the experimental results of the Algorithm 1 and conventional random test algorithm. The experimental results show that the power consumption reduction that the Algorithm 1 has been obtained is from 23% to 41% when compared

with random test algorithm, the power consumption reduction of benchmark circuits s349, s382, s420, s641, s820, s838, s1196, s1423, s1488, s9234, is 36%, 32%, 23%, 34%, 29%, 31%, 25%, 41%, 35%, 27%, respectively. These experimental results show that the test generation method in this paper is able to produce the test vectors with lower power, and can obtain higher fault coverage.

4 Conclusions

The power consumption during test process for digital circuits may be higher than that the power consumption during circuit normal operations, therefore it is needed to design the test vectors with low power. In this paper, a new method is proposed for the generation of test vectors with low power, the method makes use of the co-evolution algorithms. In the future, some work needs to be done for the design of more accurate models of power consumption for digital circuits.

Acknowledgments. This work were supported by National Natural Science Foundation of China (No.61072028), Guangdong Province & Chinese Ministry of Education Cooperation Project of Industry, Education, and Academy (No.2009B090300339). The corresponding author of this paper is Zhongliang Pan.

References

1. Tuan, D., Zhihui, K., Seng, Y.: Design and sensitivity analysis of a new current-mode sense amplifier for low-power SRAM. IEEE Trans. on Very Large Scale Integration (VLSI) Systems 19(2), 196–204 (2011)
2. Yong, K., Sang, G.P., Ta, J.P.: An ultra low-power CMOS transceiver using various low-power techniques for LR-WPAN applications. IEEE Trans. on Circuits and Systems I: Regular Papers 59(2), 324–336 (2012)
3. Yang, M.H., Kim, Y., Park, Y., Lee, D., Kang, S.: Deterministic built-in self-test using split linear feedback shift register reseeding for low-power testing. IET Computers and Digital Techniques 1(4), 369–376 (2007)
4. Abu, A.S., Quigley, S.F.: Bit-swapping LFSR for low-power BIST. Electronics Letters 44(6), 401–403 (2008)
5. Rau, J.C., Wu, P.H.: Power-aware multi-chains encoding scheme for system-on-a-chip in low-cost environment. IET Computers and Digital Techniques 5(1), 25–35 (2011)
6. Hong, S.K., Sungho, K., Michael, S.H.: A new scan architecture for both low power testing and test volume compression under SoC test environment. Journal of Electronic Testing: Theory and Applications 24(4), 365–378 (2008)
7. Chun, S., Kim, Y., Yang, M.H., Kang, S.: IOC-LP: Hybrid test data compression/ decompression scheme for low power testing. IEE Proceedings: Circuits, Devices and Systems 153(4), 391–398 (2006)
8. Mishra, A., Sinha, N.: Modified scan flip-flop for low power testing. In: 19th IEEE Asian Test Symposium, pp. 367–370 (2010)
9. Kothe, R., Vierhaus, H.T.: A scan controller concept for low power scan tests. Journal of Low Power Electronics 4(3), 420–428 (2008)

10. Moghaddam, E.K., Rajski, J., Reddy, S.M., Janicki, J.: Low test data volume low power at-speed delay tests using clock-gating. In: 20th Asian Test Symposium, pp. 267–272 (2011)
11. Meng, F.W., Kai, S.H., Jiun, L.H.: LPTest: A flexible low-power test pattern generator. Journal of Electronic Testing: Theory and Applications 25(6), 323–335 (2009)
12. Zeng, W., Chow, M.Y.: Optimal tradeoff between performance and security in networked control systems based on coevolutionary algorithms. IEEE Trans. on Industrial Electronics 59(7), 3016–3025 (2012)
13. Cartlidge, J., Boudaoud, D.: Autonomous virulence adaptation improves coevolutionary optimization. IEEE Trans. on Evolutionary Computation 15(2), 215–229 (2011)

Research of Code Construction for OCDMA System

Fugen Su[1,2] and Hongli Jin[3]

[1] Century College, Beijing University of Posts and Telecommunications, China, 102613
[2] State Key Laboratory of Information Photonics and Optical Communications Beijing University of Posts and Telecommunications, China, 100876
[3] Department of Mechanical & Electrical Engineering, Beijing Institute of Economic Management, China, 100102

Abstract. The construction of three optical codes family for SAC OCDMA are presented and analyzed, dynamic cyclic shift (DCS) code, random diagonal (RD) code and zero cross correlation (ZCC) code. RD code is constructed using code level and data level, Dynamic Cyclic Shift (DCS) code is more flexible in selection the code length with variable cross correlation and increasing of the number of users, which is no need to increase the weight value. Zero cross correlation (ZCC) code provides much easier code construction, given any number of users and weights to have minimum code length.

Keywords: Spectral Amplitude Coding (SAC) OCDMA, Dynamic Cyclic Shift (DCS) code, Random Diagonal (RD) code, Zero Cross Correlation (ZCC) code.

1 Introduction

Optical Code Division Multiple Access (OCDMA) technique has been paying more attention recently due to its much advantage. One of the main advantages of OCDMA technique is supporting many users to access the optical network simultaneously and asynchronously by allocating the unique code to each user. However, when the many users involved the OCDMA system, the system performance will be degraded due to the Multi Access Interference (MAI) Spectral Amplitude Coding(SAC) OCDMA scheme has the talent to reduce the MAI by using code with fixed cross correlation Many new codes have been proposed for OCDMA systems in recent years to eliminate the MAI effect.

Optical code is always represented in the form of (N, W, λc), where N is the code length (number of total chips), W is the weight of the code (chips that have a value of 1), and λc is the maximum cross-correlation value between any two code sequences (number of spectral overlapping in the sequence). $\lambda_c = \sum_{i=1}^{N} x_i y_i$ which $X = (x1, x2, \ldots \ldots xN)$ $Y = (y1, y2, \ldots \ldots yN)$. When $\lambda c \leq 1$, it is considered that the code possesses ideal cross correlation properties in OCDMA systems. N, W, λc are three key factor of OCDMA system, increasing the weight will increase the signal power of the users and gives a better signal to noise ratio (SNR). But if the code length is too long, it is considered disadvantageous in implementation since either very wide band sources or very narrow filter bandwidths are required [1][2] .

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 265–270.
springerlink.com

2 Construction of DCS Code [3][4]

2.1 First: The Weight Sequence Si

Si will be set as: Si={S^o,S^1,S^2,…$S^{i.}$, S^{w-1}}, where i={0,1,2,3,…,w-1} $S^o=2^0$, $S^1=2^1$ and $S^i=S^{i-1}+2^{i-}$($S^i \geq 2$), so the position of 1 in weight sequence would be:

$$S^i = \begin{cases} 1, & \text{for } i = 0,1,2,\ldots . w-1 \\ 0, & \text{else} \end{cases} \quad (1)$$

2.2 Second: The Dynamic Sequence T

T = {T_1, T_2,… T_D}, Where D is integer number and the length of the dynamic part). In order to achieve the cyclic property, the values of T are always zero (T=0) and D >7.

2.3 Third

The total combination of the two sequences represents the first code sequence for the first user. that is

$$C_i = \{S^i\} + \{T\} = \{S^0, S^1, \ldots . S^{w-1}, T_1, T_2, \ldots . T_D\}, \quad (2)$$

Based on the cyclic shift method we can get the code sequence for the rest users. The code sequence length is given by

$$N = \sum_{i=1}^{w-1} 2^i + D \quad (3)$$

Based on the cyclic shift the number of the code sequence is equal to the code length

$$(K=N). \quad (4)$$

Table 1. DCS Code Word, W=3, K=14, D=8

K	Code word													
C1	1	1	0	0	0	1	0	0	0	0	0	0	0	0
C2	0	1	1	0	0	0	1	0	0	0	0	0	0	0
C3	0	0	1	1	0	0	0	1	0	0	0	0	0	0
C4	0	0	0	1	1	0	0	0	1	0	0	0	0	0
C5	0	0	0	0	1	1	0	0	0	1	0	0	0	0
C6	0	0	0	0	0	1	1	0	0	0	1	0	0	0
C7	0	0	0	0	0	0	1	1	0	0	0	1	0	0
C8	0	0	0	0	0	0	0	1	1	0	0	0	1	0
C9	0	0	0	0	0	0	0	0	1	1	0	0	0	1
C10	1	0	0	0	0	0	0	0	0	1	1	0	0	0
C11	0	1	0	0	0	0	0	0	0	0	1	1	0	0
C12	0	0	1	0	0	0	0	0	0	0	0	1	1	
C13	0	0	0	1	0	0	0	0	0	0	0	0	1	1
C14	0	0	0	0	1	0	0	0	0	0	0	0	0	1

Examples are presented in Table 1, the code structure is simple, by increasing D can increase the number of users without increasing w.

3 Construction of ZCC Code [5][6]

3.1 First

The KxK Unit Code Matrix. P_k is as flowing:

$$p_k = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}_{k*k}$$

3.2 Second

Similarly the code matrix Ak2 for K users and weight w=2, is given by KxL matrix

$$A_k^2 = |p_k : p_k|_{k*2k} = \begin{vmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{vmatrix}_{k*2k} \quad (5)$$

Noticed that, $N = 2K$ (6)

Similarly the code matrix Akw, for K users and weight w in KxwK matrix

$$A_k^w = |p_k : p_k : \ldots\ldots p_k|_{k*wk} \quad (7)$$

Noticed that, the code length $N = wK$ (8)

Thus, for a given number of users K, and a given weight w, a set of Zero Cross Correlation code of minimum length can be constructed by Equation (3).

This procedure will be explained with the help of an example. Assume that, it is desired to generate a set of minimum length zero iPCC optical codes for 3 users having weight w=2. Let's, K=3 and w=2.

The unit code matrix for this code is $p_3 = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}_{k*k}$

The code matrix for this code is $A_3^2 = |p_3 : p_3| = \begin{vmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{vmatrix}_{k2k*}$

From Example shows, the matrix A is unique and independent of each other, with unspecified number of users.

4 Construction of RD Code

An (N, W,λ) RD code is a family of (0,1) sequence of length N, weight W and λ is the in-phase cross correlation which satisfy the following two properties: 1- Zero cross-correlation will minimized the λ and reduce PIIN (Phase Induced Intensity Noise) 2- No cross correlation in data level. The design of this new code can be performed by dividing the code sequence into two groups which are code level (segment) and data level (segment).[7][8]

4.1 First, Data Segment

let the elements in this group contain only one 1 to keep cross correlation zero at data level ($\lambda = 0$), this property is represented by the matrix (K x K) where K will represent number of user these matrices have binary coefficient and a basic Zero cross code (weight=1) is defined as [Y1] for example three users (K=3), y(K x K) can be expressed as[Y1].Where [Y1] –consists of (K × K) identity matrices. Notice, for above expression the cross correlation between any two rows is always zero.

$$Y_1 = \begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix}$$

4.2 Second, Code Segment

The representation of this matrix can be expressed as follows for W=4

$$Y_2 = \begin{vmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 \vdots & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{vmatrix}$$

Where [Y2] – consists of two parts weight matrix part and basics matrix part [B] and weight part [M], which can be expressed as

$$[B] = \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{vmatrix} \qquad [M] = \begin{vmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{vmatrix}$$

matrix [M] is responsible for increasing number of weights, let i= (W-3) and Mi, where i represents number of Mi matrix on [M],given by

$$[M] = [M1M2\ M3.....Mi\ ] \tag{9}$$

For example if W=5, from Eq.(9) i=2, so that

$$[M] = [M1/M2] = \begin{vmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{vmatrix}$$

Notice that to increase the number of users simultaneously with the increase of code word length we can just repeat each row on both Matrixes [M] and [B], for K user matrix [M] and [B] can be expressed as

$$|M|(j) = \begin{vmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ \vdots & \vdots \\ a_{j1} & a_{j2} \end{vmatrix} \qquad |B|(j) = \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 1 \end{vmatrix}$$

Where j represents the value for K- th user (j=1,2…K), and the value of aj is either zero or one. The weights for code part for both matrix [M], [B] are equal to W-1, so the total combination of code is represented as(K × N) where K=3, N=8, as given by[Z1]=[Y|1Y2]

$$[Z1] = \begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$$

From the above basic matrix Z1, the number of users (K) and the code length (N), is given by (K × N) matrix. Notice that the code weight of each row is equal 4, the relation between N and K for this case (W=4) can be expressed as

$$N=K+5 \tag{10}$$

As a result we can find that for W=5,6, and 7 code word length N can be expressed as K+7, K+9 and K+11 respectively. As a result the general equation describing number of users K, code length N and code weight W is given as

$$N = K+ 2W -3 \tag{11}$$

but the key points of RD code is that the code length (N), RD code offers better performance than other codes in term of code length for same number of users, K, of 30 as shown in Table(2). Short code length limit the addressing flexibility of the codes, while long code length are considered disadvantage in implementation, since either very wide –bandwidth source or very narrow filter bandwidth are required, RD codes exists for practical code length that are neither too short nor too long.

5 Comparison and Conclusion

In optical CDMA systems using Spectral Amplitude Coding, the length of the code is an important parameter. It is desirable to have smaller code length as this will required smaller bandwidth. Moreover, code with smaller length will require less number of filters at the encoder as well decoder. This will reduce the complexity and cost of the system.

Table 2. Comparison of SAC-OCDMA codes for 30 Users

CODE	USERS(K)	WEIGHT(W)	LENGTH(N)	CROSSCORRELATIONλc
RD	30	3	33	DATESEGMENTλc=0, CODESEGMENTλc>=1
DCS	30	2	30	0 or 1
ZCC	30	3	90	0

The larger the code weight can increase the complexity and also lead to high power losses at encoder/decoder, the large cross-correlation value results in intensify the MAI effects, codes with minimum cross correlation shown the numerous advantages including code efficient in term of BER performance, free from suffering of intensity noise even from incoherent source.

References

[1] Aljunid, S.A., Ismail, M., Ramli, A.R., Ali, B.M., Abdullah, M.: A New Family of Optical Code Sequences for Spectral-Amplitude-Coding Optical CDMA Systems. IEEE Photonics Technology Letters 16(10) (October 2004)

[2] Anuar, M.S., Aljunid, S.A., Saad, N.M.: Development of a New Code For Spectral-Amplitude Coding Optical Code. Journal of Computer (2006)

[3] Rashidi, C.B.M., Aljunid, S.A., Ghani, F., Anuar, M.S., AI-Khafaji, H.: New Design of Zero Cross Correlation Codes for Spectral Amplitude Coding in OCDMA systems. IEEE (2011)

[4] Fadhil, H.A., Aljunid, S.A., Badlishah Ahmed, R.: Random Diagonal Codes for Spectral Amplitude Coding Optical CDMA Systems Using Fiber Bragg-Grating. IEEE (2008)

[5] Fadhil, H.A., Aljunid, S.A., Badlishah Ahmad, R.: Improving BER using RD Code for Spectral Amplitude Coding Optical CDMA _etwork. In: 2009 International Conference on Computer Engineering and Technology (2009)

[6] Anuar, M.S., Aljunid, S.A., Arief, A.R., Saad, N.M.: LED Spectrum Slicing for ZCC SAC-OCDMA Coding System. IEEE (2010)

[7] Abda, T.H., Aljunida, S.A., Fadhila, H.A., Ahmad, Aa, R., Saadb, N.M.: Suppression of the Phase Induced Intensity Noise Based on the Dynamic Cyclic Shift Code for SACOCDMA Access Networks. IEEE (2011)

[8] Sahbudin, R.K.Z., Abdullah, M.K., Mokhtar, M.: Performance improvement of hybrid subcarrier multiplexing system using code for OCDMA systems using spectral direct decoding detection technique. J. Optical Fiber Technology 15, 266–273 (2009)

BER Analysis on Mono-tone Jamming in the Field of Electronic Engineering

Youyou Xi and Naiping Cheng

Academy of Equipment
Beijing, China
yyxi0226@126.com

Abstract. The error performance analysis is an important research indicators in the field of electronic technology and electronic engineering.To improve the accuracy and the practicality of the bit error rate performance in the presence of mono-tone jamming on the DS system, a new method and some results more accurate and applicable than the methods presented before is derived here, by assuming the jamming phase and pseudo-random sequence as a uniformly distributed random variable. In both case for mono-tone jamming frequency equal to and unequal to the carrier frequency of the DS system, it makes a theoretical analysis on bit error rate of communication system. The assumptions is fitter with the actual communication system more practical and higher accuracy of error performance in this method. At last, The Matlab simulation base on Monte Carlo simulation model can also indicate theoretical analysis and simulation results are consistent, and the method has accurate bit error rate performance.

Keywords: electronic technology, communication of the DS system, direct sequence spread spectrum system; mono-tone jamming; bit error rate.

1 Introduction

The communication of the DS system used high-rate pseudo-code sequence to spread spectrum of the baseband data in the sender, at the receiving end with the same sequence of pseudo-code to dispreading [1]. DS system has excellent anti-jamming ability and makes it widely used in military communications; Mono-tone jamming effect is obvious on DS system. At present, the research of the mono-tone jamming of DS system is limited at home and abroad. It equivalents to the jamming signal as part of white-noise, then calculate the bit error rate of jamming signal use the formula of white-noise error rate of DS system[2],[3],[4]. However, the synthetic component of mono-tone jamming signal and noise signal may not satisfy the Gaussian distribution. Therefore, this method is not very accurate. Some paper analyses the phase with the jamming signal as the nonrandom signal, it does not match with the actual situation and applies only to a few special cases [5],[6],[7],[8],[9]. Although it analyses the phase with the jamming signal as the random signal, it makes an uncertainty analysis of BER performance of BPSK use vector graphics analysis method, but it doesn't make qualitative analysis, also, it doesn't consider the situation of DSSS. Although it

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 271–278.
springerlink.com

analyses the phase with the jamming signal as random signal, but it doesn't considers the situation of DSSS. There-fore, In order to understand the mono-tone jamming BER of BPSK DS system better, by assuming the jamming phase and pseudo-random sequence as a uniformly distributed random variable, we get the accurate bit error rate formula of the mono-tone jamming DS system and done a simulation.

2 BER Analysis on Mono-tone Jamming of DSSS

Monte Carlo simulation method is a numerical computation method to calculate the approximate solution by statistical testing of random variable, it usually used for the engineering and technical. Monte Carlo simulation model in the situation of mono jamming DS System is shown in Figure1.

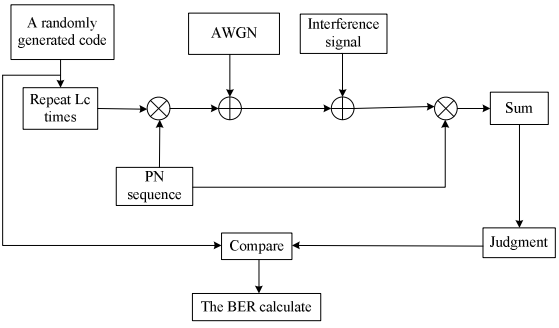


Fig. 1. Monte Carlo simulation model of DSSS

Assuming a period of the spread spectrum code only has a carrier. Therefore, after the spread spectrum, a randomly generated code and BPSK modulation can be expressed as the formula (1).

$$S_i(t) = \begin{cases} S_1(t) = A\cos(2\pi f_c t)P(n) \\ S_{-1}(t) = -A\cos(2\pi f_c t)P(n) \end{cases} \tag{1}$$

In formula (1), $S_1(t)$ is the BPSK modulation spread spectrum signal when send symbol "1", $S_{-1}(t)$ is the BPSK modulation spread spectrum signal when send symbol "0", The carrier frequency of the carrier signal is f_c , code rate is $1/T$, $P(n)$ is the bipolar rectangular pulse spreading sequence, an information code corresponding to N spreading codes and N carrier. If the energy of each symbol period is E_b , then $A = \sqrt{2E_b / T}$.

First, we assume that send the symbol "1", mono-tone jamming signal is $J_t = A_j \cos(2\pi f_j t + \varphi)$, A_j is signal amplitude, f_j is frequency, φ is phase, $n(t)$ is channel noise and it is the additive white Gaussian noise, $N_0 / 2$ is bilateral power spectral density, then, $M(t)$ is the signal after channel transmission.

$$M(t) = S_1(t) + J_t + n(t) \tag{2}$$

Suppose the received signal is synchronized. The signal $M(t)$ can be expressed as $K(t)$ after coherent demodulation.

$$\begin{aligned} K(t) &= \frac{A^2}{2} P(n)(1+\cos(4\pi f_c t)) + An(t)\cos(2\pi f_c t) \\ &+ \frac{AA_j}{2}\cos\left[2\pi(f_c + f_j)t + \varphi\right] + \frac{AA_j}{2}\cos\left[2\pi(f_c - f_j)t - \varphi\right] \end{aligned} \tag{3}$$

After dispreading, the demodulated signal is:

$$\begin{aligned} N(t) &= \frac{A^2}{2} P^2(n)(1+\cos(4\pi f_c t)) + AP(n)n(t)\cos(2\pi f_c t) \\ &+ \frac{AA_j}{2} P(n)\cos\left[2\pi(f_c + f_j)t + \varphi\right] + \frac{AA_j}{2} P(n)\cos\left[2\pi(f_c - f_j)t - \varphi\right] \end{aligned} \tag{4}$$

The sampled values when $t = nT$ is

$$\begin{aligned} R(t) = \int_0^T N(t) &= \frac{A^2}{2} P^2(n)\int_0^T (1+\cos(4\pi f_c t))dt + AP(n)\int_0^T n(t)\cos(2\pi f_c t)dt \\ &+ \frac{AA_j}{2} P(n)\int_0^T \left\{\cos\left[2\pi(f_c + f_j)t + \varphi\right] + \cos\left[2\pi(f_c - f_j)t - \varphi\right]\right\}dt \end{aligned} \tag{5}$$

1) when $f_c - f_j = 0$,because of $f_c + f_j >> 1$, then:

$$R(t) = \frac{A^2 T}{2} + \frac{AA_j T}{2} P(n)\cos\varphi + AP(n)\int_0^T n(t)\cos(2\pi f_c t)dt \tag{6}$$

Let the random variable:

$$X_1 = P(n); X_2 = \cos\varphi; X_3 = A\int_0^T n(t)\cos(2\pi f_c t)dt \tag{7}$$

By the probability theory, we know the probability density of X_2 is:

$$f_{X_2}(x_2) = \frac{1}{\pi\sqrt{1-x_2^2}}, -1 < x_2 < 1 \tag{8}$$

The probability density of X_3 is:

$$f_{X_3}(x_3) = \frac{1}{\sqrt{\pi E_b N_0}} e^{-\frac{x_3^2}{E_b N_0}} \tag{9}$$

After analysis, we can consider $P(n)$ as a Bernoulli trial of the symbol randomly selected 1or-1, it satisfies the 0-1 distribution and have equal probability to selected 1or-1.

Theorem 1[10]**:** Assume η is continuous random variable, its probability density is $f_\eta(y)$; ξ is discrete random variable, its probability density is

$P(\xi = x_k) = p_k, \left(x_k \neq 0, k = 1, 2..., \sum_{k=1}^{\infty} p_k = 1 \right)$. suppose $\varsigma = \xi\eta$,therefore, ς is continuous random variable, its probability density is $f_\varsigma(z) = \sum_{x_k>0} p_k \frac{1}{x_k} f_{\eta|\xi}(\frac{z}{x_k} | x_k) - \sum_{x_k<0} p_k \frac{1}{x_k} f_{\eta|\xi}(\frac{z}{x_k} | x_k)$. In particular, when η and ξ are independent of each other: $f_\varsigma(z) = \sum_{x_k>0} p_k \frac{1}{x_k} f_\eta(\frac{z}{x_k}) - \sum_{x_k<0} p_k \frac{1}{x_k} f_\eta(\frac{z}{x_k})$.

Apparently, X_1 and X_3 、 X_1 and X_2 are independent of each other. According to theorem 1, the probability density of $Z = X_1 X_3$ is:

$$f_Z(z) = \frac{1}{\sqrt{\pi E_b N_0}} e^{-\frac{z^2}{E_b N_0}} \tag{10}$$

The probability density of $X_4 = X_1 X_2$ is:

$$f_{X_4}(x_4) = \frac{1}{\pi\sqrt{1 - x_4^2}}, -1 < x_4 < 1 \tag{11}$$

Let $Y = \frac{A^2}{2} T + \frac{AA_j}{2} T \cdot X_4$, the probability density of Y is:

$$f_Y(y) = \frac{2}{\pi\sqrt{(AA_jT)^2 - \left(2y - A^2T\right)^2}}, \frac{1}{2}\left(A^2T - AA_jT\right) \le y \le \frac{1}{2}\left(A^2T + AA_jT\right) \tag{12}$$

Therefore, the bit error rate $P(e | s_1)$ of one spread code when send $S_1(t)$ is:

$$P(e | s_1) = P(Y + Z < 0) = \iint_{x+y<0} f_Y(y) f_z(z) dydz \tag{13}$$

Similarly, the bit error rate $P(e | s_{-1})$ of one spread code when send $S_{-1}(t)$ is:

$$P(e | s_{-1}) = P(Y + Z > 0) = 1 - \iint_{x+y<0} f'_Y(y) f_z(z) dydz \tag{14}$$

and:

$$f'_Y(y) = \frac{2}{\pi\sqrt{(AA_jT)^2 - \left(2y + A^2T\right)^2}}, \frac{1}{2}\left(-A^2T - AA_jT\right) \le y \le \frac{1}{2}\left(-A^2T + AA_jT\right) \tag{15}$$

Because it is equiprobable when send symbol "1"or"0", P_e is the bit error rate of a spread spectrum code when the frequency of jamming signals equal to carrier frequency.

$$P_e = \frac{1}{2} P(e | s_{-1}) + \frac{1}{2} P(e | s_1) \tag{16}$$

In the judgment, it usually considers the information symbol error when more than half of the symbol error in received spread spectrum data. Therefore, the information code error rate is shown in formula (17).

$$\sum_{k=N/2+1}^{N}\binom{N}{k}P_e^k\left(1-P_e\right)^{N-k} \tag{17}$$

2) when $f_c - f_j \neq 0$,let $\Delta w = 2\pi(f_c - f_j)$, similar to formula(17).Formula(5)can rewrite as:

$$R(t) = AP(n)\int_0^T n(t)\cos(2\pi f_c t)dt + \frac{A^2T}{2} + P(n)\frac{AA_j}{2}\times\left(\frac{\sin\Delta wT}{\Delta w}\cos\varphi + \frac{\sin\varphi}{\Delta w} - \frac{\cos\Delta wT}{\Delta w}\sin\varphi\right) \tag{18}$$

Let $\frac{\sin\Delta wT}{\Delta w}\cos\varphi + \frac{\sin\varphi}{\Delta w} - \frac{\cos\Delta wT}{\Delta w}\sin\varphi = a\cos\varphi + b\sin\varphi$, by the probability theory, we know the probability density of $a\cos\varphi + b\sin\varphi$ is:

$$\frac{1}{\pi\sqrt{\left(a^2+b^2\right)-y^2}}\ ,-\sqrt{a^2+b^2} < y < \sqrt{a^2+b^2} \tag{19}$$

The probability density of $Y = \frac{A^2T}{2} + \frac{AA_j}{2}P(n)\left(a\cos\varphi + b\sin\varphi\right)$ is:

$$f^{\Gamma}{}_Y(y) = \frac{2}{\pi\sqrt{(AA_j)^2\left(a^2+b^2\right)-(2y-A^2T)^2}}, \frac{A^2T}{2} - \frac{AA_j}{2}\sqrt{a^2+b^2} \le y \le \frac{A^2T}{2} + \frac{AA_j}{2}\sqrt{a^2+b^2} \tag{20}$$

Similarly, the bit error rate of Y when send symbol "-1" is:

$$f^{\Gamma}{}_Y(y) = \frac{2}{\pi\sqrt{(AA_j)^2\left(a^2+b^2\right)-(2y+A^2T)^2}}, -\frac{A^2T}{2} - \frac{AA_j}{2}\sqrt{a^2+b^2} \le y \le -\frac{A^2T}{2} + \frac{AA_j}{2}\sqrt{a^2+b^2} \tag{21}$$

The value of $f_Z(z)$ is equal to the situation of $f_c - f_j = 0$, $P_e{}'$ is the bit error rate of a spread spectrum code when the frequency of jamming signals unequal to carrier frequency.

$$P_e{}' = \frac{1}{2}P'(e \mid s_{-1}) + \frac{1}{2}P'(e \mid s_1) \tag{22}$$

When $f_c - f_j \neq 0$, the BER when send a message code is:

$$\sum_{k=N/2+1}^{N}\binom{N}{k}P'^{\,k}_{\ e}\left(1-P'_{\ e}\right)^{N-k} \tag{23}$$

The formula (17), (23) is very complex, but they accurately reflect the BER of the mono-tone jamming in BPSK DS system. It can be seen the BER is related to the carrier frequency of DS system, the frequency of jamming signal, signal to jamming ratio, signal to noise ratio, pseudo-random code, etc. However, due to formula (17),

(23) have non-integrable integral with forms of $\int_{-\infty}^{+\infty} e^{-x^2} dx$, when there are human disturbance, noise jamming effect is very small, so in order to depth analysis of the impact factor of the BER in mono-tone jamming BPSK DS system. We can ignore the impact of channel noise.

3) When $f_c - f_j = 0$ and there is no channel noise, P_{ej} is the bit error rate of a spread spectrum code when the frequency of jamming signals equal to carrier frequency.

$$P_{ej} = \begin{cases} 0\ ,\ A > A_j \\ \dfrac{1}{2} - \dfrac{1}{\pi} \arcsin \dfrac{A}{A_j}\ ,\ A < A_j \end{cases} \tag{24}$$

4) When $f_c - f_j \neq 0$ and there is no channel noise, P'_{ej} is the bit error rate of a spread spectrum code when the frequency of jamming signals unequal to carrier frequency.

$$P'_{ej} = \begin{cases} 0\ ,\ A > A_j\left(\sqrt{a^2+b^2}/T\right) \\ \dfrac{1}{2} - \dfrac{1}{\pi} \arcsin \dfrac{AT}{A_j\sqrt{a^2+b^2}}\ ,\ A < A_j\left(\sqrt{a^2+b^2}/T\right) \end{cases} \tag{25}$$

Similar to the formula (17),(23),we can calculate the BER of the information code from formula (24),(25),from the formula (24),(25),it can be seen transmission error rate is 0 when the energy of jamming signal lower to a certain degree than transmission signal, also, the BER is increased with the enhancement of the interfering signals. Due to $a^2 + b^2 < T$, The BER is lager when $f_c - f_j = 0$.

3 Simulations

Assume that the bandwidth of information code is 40 KHz; The sampling frequency is 40MHz; The bandwidth of the PN code to 400KHz, so each information code contains 1000 sampling points, each PN code containing 100 sampling points; The PN code generated randomly; Modulation is adopted for BPSK; The carrier signal is a cosine signal, phase is 0; The carrier frequency is 400 kHz; The mono-tone jamming signal frequency deviation from the carrier frequency is 500KHz; It generated a total of one million information symbol randomly.

The theoretical curve of the BER of mono-tone jamming DS system is shown in Figure 2, the simulation curve of the BER of mono-tone jamming DS system is shown in Figure 3, Simulation results show that the BER is lager when the frequency of jamming signals equal to carrier frequency. When carrier amplitude is larger than the jamming amplitude, the system error rate is 0.

The BER curve when the frequency of jamming signals equal to carrier frequency is shown in Figure 4, the BER curve when the frequency of jamming signals unequal to carrier frequency is shown in Figure 5, the simulation results show that the results of theoretical analysis and actual simulation results are the same.

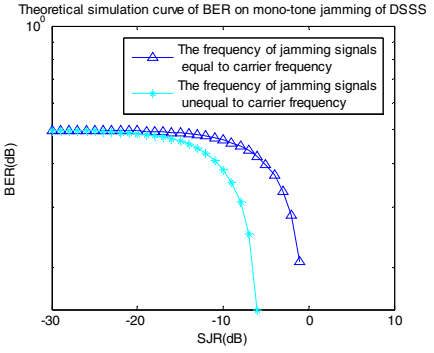


Fig. 2. Theoretical simulation curve of BER on mono-tone jamming of DSSS

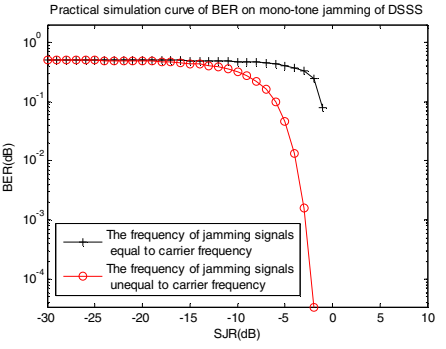


Fig. 3. Practical simulation curve of BER on mono-tone jamming of DSSS

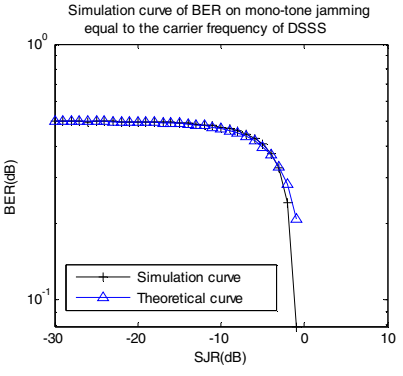


Fig. 4. Simulation curve of BER on mono-tone jamming equal to the carrier frequency of DSSS

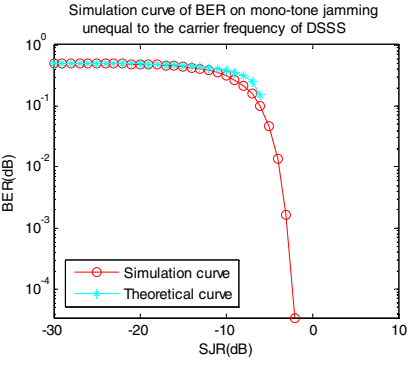


Fig. 5. Simulation curve of BER on mono-tone jamming unequal to the carrier frequency of DSSS

4 Conclusions

Mono-signal as a common continuous wave signal, it widely used to interfere with the direct sequence spread spectrum communication system. The initial conditions are set ideal in some of the existing analytical methods, the results of the analysis is inaccurate and does not have practical value. The other methods only analyses one aspect of the problem. This paper introduces the phase of the jamming signal and the PN code sequence as a random variable, inducted a comprehensive BER analysis based on Monte Carlo simulation method in BPSK DS system. The method got new analysis methods and results, improved the accuracy and practicability of error performance in the mono-signal jamming DS system. The Matlab simulation base on Monte Carlo simulation model indicates the method has accurate bit error rate performance. It provides a new method for accurate analysis of the BER performance in the field of electronic technology and electronic engineering.

References

1. Poisei, R.A.: Modern Communications Jamming Principles and Techniques. Publishing House of Electronics Industry (2005)
2. Huo, X., Liu, J., Jiang, Z., et al.: Monte Carlo Simulation and BER Analysis of DS Spread Spectrum System. Modern Electronic Technique (7), 36–40 (2008)
3. Liu, J.: Research on the method of single-tone jammingof DS/BPSK spread spectrum system and implementation of simulation. Modem Electronics Technique (01), 110–112 (2008)
4. Gui, H.: Research on the best jamming to the BPSK coherent reception system. Radio Communications Technology 27(6), 33–34 (2001)
5. Fang, L., Kuang, J., Lu, X.: Union tracking and detection performance of BPSK receiver in the presence of tone jamming. Television Technology (03), 33–36 (2001)
6. Li, H., Pei, B., Huang, Y.: Performance of the direct sequence spread spectrums system with single-tone jamming. In: Proceedings 2010 IEEE International Conference on Information Theory and Information Security, ICITIS 2010, pp. 458–461 (2010)
7. Liu, Z., Gao, H., Song, Y.: Analysis of single tone interfering effect on BPSK system. Radio Engineering 38(8), 18–21 (2008)
8. Zheng, L., Su, Y., Piao, T., et al.: Analysis and Its Application of Performance Degradation in BPSK Communication Systems under Mono-frequency Jamming. Journal of Astronautics 32(1), 136–141 (2011)
9. Xu, J., Gu, C., Yi, K.: BER Analysis of DS Spread Spectrum System under Mono-frequency Jamming. Space Electronic Technology 32(1), 136–141 (2011)
10. Wang, Z.: The Sum, Difference, Product and Quotient Distributions of Inhomogeneous Variables. Studies In College Mathematics 13(4), 97–99 (2010)

Three-Stage Node Selection Algorithm Based on Network Communication in Electronic Engineering*

YuFei Qin and RongFa Zhou

The network center of Guilin University of Electronic Technology, Guilin Guangxi 541004

Abstract. Communication is an important field researched by electronic engineering. In order to solve the problem that network communication robustness, synchronization delay, quality of streaming service and network load can't be kept all in one, the paper proposed a new method which improved network topology aggregation algorithm in P2P streaming node selection mechanism. The simulation results show that the proposed method improved the performance about the pressure of overlay-network and quality of streaming service.

Keywords: electronic engineering, peer to peer streaming media, peer selection, topology aggregation.

1 Introduction

Network communication is an important field researched by electronic engineering. P2P technology is a hot research in network communication field. The core idea of the P2P technology is that it use the client ordinary node to cache video resources, and through sharing the video resources among client ordinary nodes to reduce server load and the relative delay [1].

In the P2P streaming system, the node selection strategy is one of the key technologies. The node selection strategy [2] [4] point that when a new node join into P2P system, it need to use specific algorithm to select the right node as its transmission sources from other nodes that has the required content set, and establish connection and transmit data, this algorithm is the node selection strategy. In P2P streaming system, the nodes have dynamics, that means every node in the system has uncertainty in time, space, activity cycles of appearance, and in the performance of link among nodes, the uncertainty make that the system can't adopt a static method to deal with the dynamic problem[2]. Therefore, the node selection strategy in the hybrid P2P streaming media system can't use the traditional node selection methods which rely on fixed servers or router, so the node selection strategy faces more challenging.

This paper improve the present node selection algorithm--network topology aggregation algorithm in allusion to P2P streaming media system based on hybrid P2P network topology structure, which enhance the system robustness and balance network load on the premise to guarantee that the network resources have high utilization rate.

---

* YuFei Qin (1987-), male, GuangXi GuiLin, assistant engineer, the main research direction is the network information security.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 279–285.
springerlink.com

2 The Hybrid P2P Network Topology Structure

As shown in figure 1, hybrid P2P network topology structure is mainly consisted of streaming media server, super node layer and ordinary node layer.

(1) Streaming Media Server (Media Server, MS): the distribution source of Media file. (2) The Super Node (Super Peer, SP): SP has higher performance processor, and has more network bandwidth and more load capacity than ordinary node. SP has two main functions: 1) maintain the list of information of their child nodes (ordinary nodes), to store the index of sharing resources of their child nodes; 2) accept the resource request of their child nodes. (3) Ordinary node (Ordinary Peer, OP): request node.

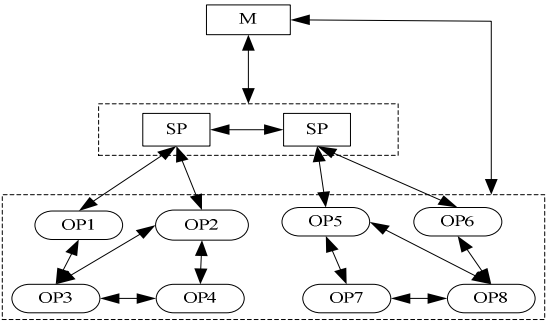


Fig. 1. The hybrid P2P network topology structure

3 Three-Stage Node Selection Algorithm Based on Hybrid P2P Streaming Media

In order to balance the service quality of the flow and keep the robustness of the system as well as to balance the network load and reduce the flow through the backbone network and improve the utilization of network resources, this paper analyzes the above typical node selection algorithms and proposes a new node selection algorithms based on improved network topology aggregation algorithm.

3.1 Algorithms Thought

The algorithm is divided into three processes: ISP priority, bandwidth-first and minimum delay.

Process1: ISP priority. In the time limited, in order to reduce the flow through the backbone network, select nodes in the same ISP and channel as the candidate nodes, and then select a certain number of candidate nodes and record in the super nodes by means of the network topology aggregation algorithm.

Process2: bandwidth-first. Measure the bandwidth of the candidate nodes and arrange them from largest to smallest, then select the appropriate number according to need.

Process 3: minimum delay. Measure the delay of the candidate nodes selected in process 2, then select the node with the minimum delay.

Through the above three processes, the network topology aggregation algorithm, the combination of bandwidth optimization and minimum delay algorithm are combined to ensure the quality of streaming service, also reduces the coverage of network pressure and improve the transmission rate and network utilization.

3.2 Algorithm Description

The flow chart for this algorithm as shown in figure 2.

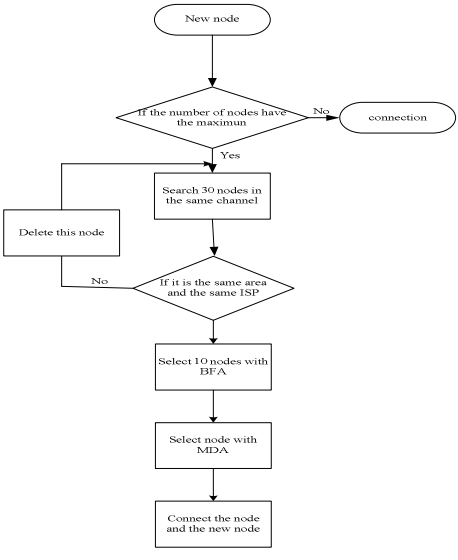


Fig. 2. The flow chart

Step1: When the new node requests to send resources to the super node, it provides its own regional and ISP. According to the connection situation of the super node, If the limit is not reached, it establishes a connection directly with the new node. Otherwise, it searches for a suitable candidate node according to the requested resource.

Step2: Select the candidate nodes according to the region and ISP provided by the new nodes. If the candidate node is not in the same area an ISP with the new node, then delete it. Otherwise, make a record in the super node.

Step3: Repeat the Step2 and Step3 and select 30 candidate nodes.

Step4: Measure the bandwidth of the 30 candidate nodes and arrange them from big to small, then select the first 10 nodes.

Step5: Measure the delay of the 10 nodes generated in the Step4 and select the one with minimum delay as the source node, and then connect it with new node.

4 Experimental Results and Performance Analysis

4.1 Simulation Experiment Environment

Experiment was simulated to P2P streaming system in the PeerSim simulator. It simulated dynamic characteristics of P2P system by random increasing 20% nodes and decreasing 20% nodes. The total number of nodes in overlay network changed be in the range of 0 to 1000, at the same time, it would divide overlay network into 16 ISP service areas. Tested random selection algorithm, bandwidth-first algorithm, the topological aggregation algorithm and the new algorithm, evaluated the respective influence on performance of their system.

4.2 The Analysis of Experimental Results

Four algorithms in any case all could guarantee that overlay network was connected under the dynamic environment, the results of performance test were shown in table 1 for their respective algorithms.

Explain: network topology aggregation algorithm(NTAA), Random selection algorithm(RSA), Bandwidth-first algorithm(BFA), the proposed algorithm(TPA).

Table 1. The results of performance test

algorithm	Connectivity	Radius of overlay network	The quality of streaming service (%)	The pressure of overlay network (%)
RSA	Yes	3	44	74
BFA	Yes	4	75	89
NTAA	Yes	16	40	29
TPA	Yes	4	69	31

(1)The pressure of overlay network

The pressure of overlay network is an important standard of measuring the node selection strategy, it is the ratio of the flow through the backbone network when data distributed to the sum of the flow that be needed by all nodes when they downloaded resources[3]. The experiment simulated pressure changed be in the range of 0 to 1000 nodes, and got four curves of the pressure change in overlay network by respectively

using the four nodes selection algorithms as shown in figure 3, could be seen from the graph, the pressure in overlay network by using this algorithm was only a little higher than that by using the pure topology aggregation algorithm. The utilization of network had been greatly improve.

(2)The quality of streaming service

The quality of streaming service users got is the most important index to measure the streaming system performance[2]. In the streaming media system, the quality of streaming service is the most important parameter to evaluate the system performance[7]. The real arrival rate of nodes can reflect the quality of streaming service in streaming media system. The four curves of the quality of streaming service changed by respectively using the four nodes selection algorithms when the number of nodes changed be in the range of 0 to 1000 as shown in figure 4.

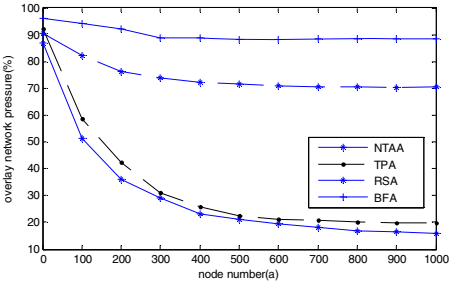


Fig. 3. The relationship between the pressure of overlay network and the number of nodes

From the graph, we can got the conclusion that the bandwidth-first algorithm had the highest quality of streaming service, the quality of streaming service of this algorithm introduced by this paper was close to the quality of streaming service of the bandwidth priority algorithm, and not greatly reduced the quality of streaming service under the environment of topology clustering .

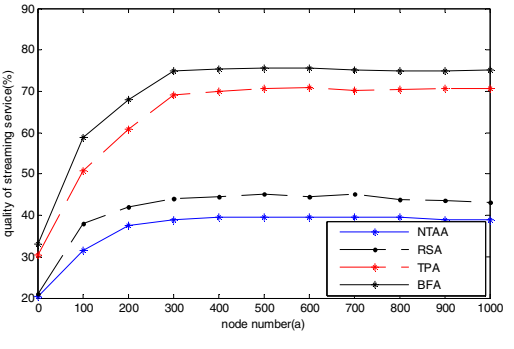


Fig. 4. The relationship between the quality of streaming service and the number of nodes

(3) Synchronous Delay

The synchronicity of the participation node and the data source node is measured by synchronous delay which is evaluated by the radius of the overlay network. The radius of the overlay network was shown in the Fig 1 and we could draw a conclusion that the order of the synchronous delay from small to large was random selection algorithm, bandwidth-first algorithm, the algorithm proposed in this paper and network topology aggregation algorithm. As the proposed algorithm was identified with random selection algorithm on the synchronous delay, the retardance was better.

(4) Robustness

The robustness of the node selection algorithm is generally evaluated by the connectivity of the overlay network and we should avoid the islet appears in the network. The connectivity and the robustness are better in dynamic. We meassured data with numbers of node which is from one to a thousand. From Fig 1 we could draw that the algorithm proposed in this paper was keep 100% connectivity of the network in any case.

5 Conclusion

Communication technology is a key technology in electronic engineering field. The algorithm of this article is a good algorithm for network communication. Firstly, it considers the network topology position of the node, makes a group of nodes which is with new nodes in the same area and the same ISP as the candidate nodes, enhances the network topology clustering of the node. Secondly, a certain number of candidate nodes whose bandwidth complies with the requirements are selected via priority choice mechanism of the bandwidth to this group of candidate nodes. Lastly, the delay time of these candidate nodes is tested, and a minimum delay node is selected as a source node. The simulation results show that the algorithm gets a good balance among the flow service quality, pressures of the overlay network, service delay and robustness of network, is a pretty good selection algorithm of the node.

Acknowledgments. This work was supported by Projects of National Natural Science Foundation (No:61163058).

References

1. Ren, L.-B., Huang, G.-M.: Research on cache mechanism for VOD based on P2P network. Journal of Guilin University of Electronic Technology 31(5), 350–354 (2011)
2. Zhen, J., Zhang, S.: Study of Peer Selection Strategy for Mesh Based P2P Streaming. Microelectronics & Computer 24(12), 95–99 (2007)
3. Sun, M.-S., Zhang, X.-Y., Yao, L.: Double-ended Cooperation Peer Selection Strategy for P2P Media Streaming. Computer Engineering 35(18), 105–107 (2009)
4. Ge, Y., Liang, J.: Improved High-Priority strategy for P2P Streaming Peer Selection. Computer System & Application (11), 45–46 (2009)

5. Zhen, J., Zhang, S., Qi, J., Lei, Y.-C.: Research and simulation for peer selection strategy on P2P streaming. Computer Engineering and Design 28(22), 5396–5399 (2007)
6. Wu, C., Li, B.: Optimal peer selection for minimum-delay peer- to- peer streaming with rateless codes. In: International Multimedia Conference Proceedings of the ACM Workshop on Advances in Peer- to- Peer Multimediastreaming, Hilton, Singapore, pp. 69–78 (2005)
7. Sun, W.-J., Zhou, J.-Q.: Study of node selection strategy in P2P media streaming based on service quality. Microcomputer & Application 11, 80–84 (2010)

Hybrid Quantum Evolutionary Algorithm and Its Application in Multiuser Detection of Electronic Communication System

Qiang Song[1,2] and Xialing Liu[3]

[1] Wuhan University of Technology, Wuhan, 430073, China
[2] Anyang Institute of Technology, Anyang, 455000, China
[3] Anyang Normal University, Anyang, 455000, China
llx@aynu.edu.cn

Abstract. This paper combines quantum evolutionary algorithm (QEA) with particle swarm optimization (PSO) together, proposes two kinds of hybrid quantum evolutionary algorithm. One is particle swarm embedded quantum evolutionary algorithm (PSEQEA); another one is quantum binary particle swarm optimization (QBPSO). The experiment results of multiuser detection problem show that both of the methods not only have simpler algorithm structure, but also perform better than conventional QEA and BPSO in terms of ability of global search optimum. In the third generation mobile communication system, multiuser detection technology is one of the most effective ways to solve multi-access interference in electronic communication system; these two algorithms can simplify and accelerate the detection process.

Keywords: quantum evolutionary algorithm, particle swarm optimization, hybrid, evolutionary algorithm, electronic communication system.

1 Introduction

Quantum computing with its powerful computing power has become one of the leading edge of the current era of science. It shows the attractive prospect of quantum information processing, which inspires researchers from quantum computing mechanism to study some traditional algorithm again. The combination of quantum computing and evolutionary computation is starting from the late 1990s, the researchers in literature [1-4] had made beneficial attempt. But through the experiment and analysis, it can be found that QEA ability to solve the complex optimization problem is not so strong, and easy to fall into local optimum. By the way, QEA adopts look-up table to update quantum gate, on the one hand, which needs to set more parameters, on the other hand which is more complicated for realization.

In order to make the quantum evolutionary algorithm solve practical optimization problem more effectively, exploring new quantum evolutionary algorithm is very necessary. The particle swarm optimization (PSO) algorithm is a kind of evolutionary computation method based on swarm intelligence [5]. The operation speed of particle swarm algorithm is very fast, easy to implement, and the performance of algorithm

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 287–293.
springerlink.com

can be comparable to genetic algorithm, in recent years which has been widely used in all areas of engineering and science research. As with other optimization algorithm, particle swarm algorithm also has its own weakness: in later evolution its speed will be clearly slowing, and easy to fall into local optimum.

Because each optimization algorithm has its own advantages and disadvantages, combining the different optimization algorithm with each other, can obviously improve the performance of the algorithm [6]. In order to improve the performance of the QEA and BPSO, this paper puts forward two plans of QEA and PSO combining with each other: embedded particle swarm quantum evolutionary algorithm (PSEQEA) and quantum binary particle swarm optimization algorithm (QBPSO).

2 Quantum Evolutionary Algorithm and Particle Swarm Optimization Algorithm

2.1 Quantum Evolutionary Algorithm (QEA)

Quantum evolutionary algorithm is a kind of probability search algorithm which combines quantum mechanism and basic evolutionary algorithm. Its essential characteristics are making full use of the superposition and coherence of quantum state. QEA adopts a new coding party quantum bit code; the concrete form can be described as:

$$q = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_m \\ \beta_1 & \beta_2 & \cdots & \beta_m \end{pmatrix} \tag{1}$$

Inside the expression, $(\alpha_i, \beta_i)^T (i = 1, \cdots, m)$ represents a quantum bit, and fulfills the expression $|\alpha_i|^2 + |\beta_i|^2 = 1$, m represents the number of quantum bit; q is called a quantum chromosome in QEA. QEA describes the problem with quantum chromosomes.

In the coding method of quantum bits, a chromosome can represent arbitrary linear superposition state, so QEA are more likely to keep the diversity of population than other traditional evolutionary algorithm. And while $|\alpha_i|^2$ or $|\beta_i|^2$ gradually is close to 1 or 0, quantum chromosomes gradually converge to a single state, the diversity of population also decreases, and the algorithm convergences. So the coding method of quantum bits is adopted, meanwhile QEA has exploration and development ability. The general procedure of the quantum evolutionary algorithm is shown below [2]:

(1) initialize population Q (t), t = 0; (2) from Q (t) to generate P (t); (3) evaluate P(t) fitness, and preserve the optimal solution; (4) the pausing condition judgment: if it meets the outage conditions, to output the best current individual, and algorithm is over, otherwise algorithm continues; (5) update Q(t), t = t+1, return (2).

Therein Q (t) represents the No. t generation of population; P (t) represents the set of No. t generation of binary solution. While initializing population, generally all the quantum bits of quantum chromosomes are initialized to $1/\sqrt{2}$, and this means that all possible superposition states appear in the same probability. In (2) step, Q (t) generates P (t) through the random observation. The concrete operating process is as

follows: randomly generating a random number between [0, 1], if it is greater than $|\alpha_i|^2$, the bit value of corresponding binary solution is 1, otherwise its value is 0. In the (5) step, QEA adopts the quantum revolving door u (t) to update Q (t), and mathematical expression can be described as:

$$\begin{pmatrix} \alpha_i' \\ \beta_i' \end{pmatrix} = U(\theta) \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} \tag{2}$$

Therein, θ is the rotation angle of quantum revolving door, its concrete numerical value can be obtained through the look-up table way [2].

2.2 Particle Swarm Optimization (PSO) Algorithm

The particle swarm optimization algorithm is a kind of global optimization algorithm based on swarm intelligence. The individual of particle swarm represents one possible solution of the problem; each particle has two features of position and speed. The objective function values corresponding to the coordinates of particles are used for fitness algorithm through the fitness to measure the merits of the particles. Algorithm firstly in the solution space (assuming D dimensions), randomly generates a group with N particles, among which the position vector and speed vector of No. i particle can be respectively represented as $X_i = (x_{i1}, x_{i2}, \cdots, x_{iD})$ and $V_i = (v_{i1}, v_{i2}, \cdots, v_{iD})$, then through multiple iterations to find the optimal solution. In each iteration, particles renew themselves through tracking two extreme values: one is the optimal solution found by particle itself, which is the individual extremum $P_i = (p_{i1}, p_{i2}, \cdots, p_{iD})$, and the other is the optimal solution found by whole particle swarm, which is $P_g = (p_{g1}, p_{g2}, \cdots, p_{iD})$, and is called global extremum. Particle swarm algorithm is used to solve the real number coding problems, in order to solve discrete optimization problem in engineering, the literature [7] puts forward the binary code particle swarm algorithm (BPSO). The update process of BPSO particles can be expressed as:

$$v_{id} = v_{id} + c_1 r_1 (p_{id} - x_{id}) + c_2 r_2 (p_{gd} - x_{id}) \tag{3}$$

$$sigmoid(v_{id}) = 1/[1 + \exp(-v_{id})] \tag{4}$$

$$if\ (r < sigmoid(v_{id}))\ then\ \ x_{id} = 1;\ else\ \ x_{id} = 0 \tag{5}$$

Among them c_1 and c_2 are called learning factors; r_1 and r_2 are the random numbers between [0, 1]; threshold function $sigmoid(v_{id})$ transforms speed vector into the value between [0, 1]; r is the random number between [0, 1].

3 Hybrid Quantum Evolutionary Algorithms

BPSO and QEA as the emerging intelligent optimization algorithms in recent years, their advantages have been introduced in front of this paper. And there also exist some problems to be solved: in the evolutionary process of BPSO, the diversity of

particles will gradually disappear. In later evolution period algorithm convergence speed will significantly change slowly and easy to fall into local optimal; in the course of evolution QEA, only the best individual information is used, and there is no use of other immature individual that might carry global optimal solution information, which influences evolution speed, and makes evolution easy to fall into local optimum.

In order to overcome the disadvantages of QEA and BPSO, this paper combines PSO and QEA, complements each other's advantages, proposes two kinds of hybrid quantum evolutionary algorithm to improve the performance of BPSO and QEA. Below are detailed introduction.

3.1 Particle Swarm Embedded Quantum Evolutionary Algorithm

The first kind of hybrid quantum evolutionary algorithm is called particle swarm embedded quantum evolutionary algorithm (PSEQEA). Its main idea is the simplified particle swarm evolution equation embedded in the evolution operation of QEA, in PSEQEA, by quantum gate to update Q (t), the size and orientation of rotation angle of the quantum gate is obtained no longer using look-up table way, but using the evolution equation of PSO to do automatic adjustment. And the evolution equation of PSO is simplified, redefining the meaning of expression; its mathematical expression can be expressed as:

$$\theta = c_1(p_{id} - x_{id}) + c_2(p_{gd} - x_{id}) \tag{6}$$

Where the definitions of p_{id}, p_{gd}, c_1 and c_2 are as same as the definitions in PSO.

The procedure of updating Q (t) can be described as:

$$\begin{cases} \theta = c_1(p_{id} - x_{id}) + c_2(p_{gd} - x_{id}) \\ \begin{pmatrix} \alpha_{id}^{'} \\ \beta_{id}^{'} \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \alpha_{id} \\ \beta_{id} \end{pmatrix} \end{cases} \tag{7}$$

Using (6) expression to automatically adjust the size and direction of quantum door rotation angle, there are two main advantages: on the one hand, which simplifies the structure of QEA algorithm, reduces the number of parameters; On the other hand, PSO's evolution equation has memory characteristics, not only using the information of swarm optimal state, and taking advantage of the individual itself local optimum information to adjust the angle more reasonable θ, make PSEQEA than QEA have the capability of jumping out of local optimum.

The basic structure of the PSEQEA is different with QEA only in this step of updating Q (t), so it is no longer repeating the algorithm process.

3.2 Quantum Binary Particle Swarm Algorithm

The second kind of hybrid quantum evolutionary algorithm is called quantum binary particle swarm optimization (QBPSO). The main idea is to lead quantum bit code into BPSO, which can be more natural and simpler to keep population diversity.

According to the basic principles of QEA, while the algorithm guides the evolution of quantum chromosome, it is necessary to ensure $|\alpha_{id}|^2 + |\beta_{id}|^2 = 1$. Apparently the evolution equation in BPSO can not meet this requirement, it must be modified.

The relationship between $|\alpha_{id}|$ *and* $|\beta_{id}|$ can be expressed in right-angle coordinate diagram, which is shown in Figure 1. The θ_{id} is the deflection angle of No. d quantum bit of No. i quantum chromosome.

$$|\alpha_{id}| = \cos\theta_{id}, |\beta_{id}| = \sin\theta_{id} \tag{8}$$

Thus the changes between $|\alpha_{id}|$ and $|\beta_{id}|$ can be expressed in $|\theta_{id}|$ change, that is the evolution which guides quantum chromosome, it can be expressed as the evolutionary of guide $|\theta_{id}|$. Combining the characteristics of BPSO evolution equation, the vector $\theta_i(\theta_{i1}, \theta_{i2}, \cdots, \theta_{iD})$ is used to replace the speed vector, so (3) can be rewritten as:

$$\theta_{id} = \theta_{id} + c_1(p_{id} - x_{id}) + c_2(p_{gd} - x_{id}) \tag{9}$$

In order to generate a binary solution, QBPSO uses random observation to replace the converting method of sigmoid () function. The specific operation is as follows: If $r > |\alpha_{id}|^2$, then $p_{id} = 1$, otherwise $p_{id} = 0$, where r is a random number between [0, 1].

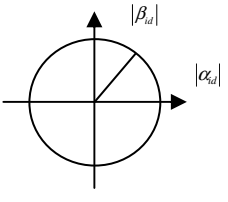


Fig. 1. The relationship between α and β

In addition, in order to improve the search efficiency of QBPSO, avoiding duplication of search, θ_{id} is limited in the first quadrant of coordinate graphs to do change. At the same time, in order to make the algorithm maintain a certain population diversity in late evolutionary period, the upper limit of id is less than $\pi/2$, and the lower limit is greater than θ_{id}, that is, θ_{id} is between $[\theta_{\min}, \theta_{\max}]$, inside $\theta_{\min}, \theta_{\max} \in [0, \pi/2]$.

Therefore PSO is a kind of evolutionary computing method; here QBPSO is regarded as a hybrid quantum evolutionary algorithm. The description of algorithm process is as same as QEA, the difference is the updating method of Q (t).

4 Experimental Simulation and Analysis

In order to study the optimization performance of PSEQEA and QBPSO, this paper implements them into multiuser detection problem of simulation experiment. In order to compare, QEA and BPSO also solve the above mentioned problem. In order to

ensure the objectivity of the experimental results, the experiment results are obtained in average after 20 time's independent experiment. The third generation mobile communication system will use the code division multiple access (CDMA) access method, and multiple access interference (MAI) is the main jamming of CDMA mobile communication system. Its existing seriously reflects the capacity and performance of CDMA communication system, and multiuser detection technology (MUD) is one of the most effective ways to solve multi-access interference. In 1986 Verdu proposed the optimal multiuser detector (OMD), but with the growth of users its calculation load appears exponential growth and becomes a NP problem, which limits its practical application. Suppose the subscribers of a synchronous CDMA communication system are K, OMD may be considered as the solution to make function's value maximal in expression (10) in 2^K solutions:

$$J(b) = 2b^T Ay - b^T Hb \tag{10}$$

Where b and A are the information bit vector and amplitude vector send by users, y is the output signal vector matching filter, H is the related functions for normalization. The optimal multiuser detector because the computational complexity is too high to use, but the expression (10) is used by multiuser detector as objective function. In this paper the globe maximum of expression (10) is also regarded as the two algorithm's calculating target.

Now aiming to the synchronous CDMA system with 10 users under the strict control of power, spreading sequence adopts 31 Gold sequence, the maximum normalized relations with each other is 9, information channel is Gaussian channel. The population scale of four kinds of algorithm is 20; pausing condition takes the maximum evolution algebra value 100.

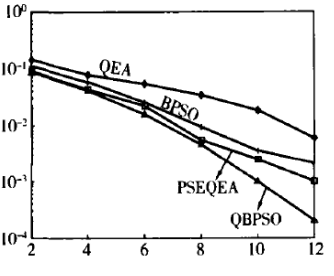


Fig. 2. The comparison of bit error rate (BER)

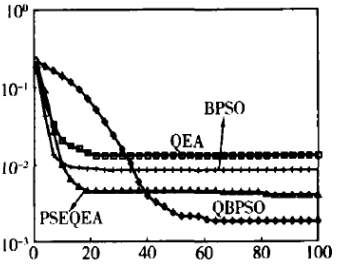


Fig. 3. The comparison of convergence speed

Figure 2 shows in different SNR (db), there are four kinds of average bit error rate of MUD. Figure 3 shows 4 kinds of MUD convergence curve in the same signal-to-noise ratio 8 db. From figure 2 and figure 3, PSEQEA bit error rate is significantly lower than QEA, and convergence speed is faster; About QBPSO, although in the early evolution its convergence speed is slower than BPSO, but the resistance capability of error bit has obvious advantages than BPSO, and it's running speed is faster than BPSO (for QBPSO, the time for each 10 bit is detected as 0. 1452s, but for BPSO it is 0.3895s). For MUD requiring to quickly find the optimal solution, obviously QBPSO has more practical value than BPSO.

5 Conclusions

This paper proposes two kinds of hybrid quantum evolutionary algorithm combining PSO and QEA with each other: embedded particle swarm quantum evolutionary algorithm (PSEQEA) and quantum binary particle swarm optimization algorithm (QBPSO). The main idea of PSEQEA is using simplified evolution equation to lead the evolution of quantum chromosomes, to simplify the structure of QEA algorithm and improve the ability of jumping out of local extremum; The main idea of QBPSO is to lead the concept of quantum chromosomes in QEA into BPSO, which improves the global search ability of BPSO algorithm and computing speed. Through the experiment of multiuser detection problem, the result shows that the new algorithm is not only simpler in operation, but also the global search ability has significantly improved.

References

1. Narayanan, A., Moore, M.: Quantum-inspired Genetic Algorithms. In: Proceedings of IEEE International Conference on Evolutionary Computation, Nagoya, pp. 61–66 (1996)
2. Han, K.-H., Kim, J.-H.: Genetic Quantum Algorithm and its Application to Combinatorial Optimization Problem. In: Proceedings of the 2000 Congress on Evolutionary Computation, Piscataway, vol. (2), pp. 1354–1360 (2000)
3. Han, K.-H., Kim, J.-H.: Quantum-inspired Evolutionary Algorithm for a Class of Combinatorial Optimization. IEEE Trans. Evolutionary Computation 6(6), 580–593 (2000)
4. Yang, S., Liu, F.: Quantum Evolutionary Strategy. Electrical Journal 29(12), 1873–1877 (2001)
5. Kennedy, J., Eberhart, R.C.: Particle Swarm Optimization. In: Proceedings of the International Conference on Evolutionary Computation, Perth Western Australia (1995)
6. Van den Bergh, F., Engelbrecht, A.P.: A Cooperative Approach to Particle Swarm Optimization. IEEE Tran. Evolutionary Computation 8(3), 225–239 (2004)
7. Kennedy, J., Eberhart, R.C.: A Discrete Binary Version of the Particles Swarm Algorithm. In: International Conference on Systems Man and Cybernetics, Orlando (1997)
8. Hacioglu, E.K.: Multiuser Detection Using a Genetic Algorithm in CDMA Communications Systems. IEEE Trans. Communication 48(8), 1374–1383 (2000)

Research of Implement and Comparison on Quick Sort Based on Mobile Agent Communication

Hualiang Wu

Information and Network Administrative Center, Jiangxi University of Technology, Nanchang China, 330098
wywchinazj_2013@126.com

Abstract. Nowadays, using the cluster system for high performance computing becomes a hot spot, this is due to inexpensive high performance PC and workstation and high-speed network technology to provide a foundation of hardware. MPI+PVM programming mainly rely on such a message passing library, transplantation of poor. Java based mobile agent technology for distributed computing provides a new computing model, with good portability, scalability, flexibility. Along with the computer network and the Internet expands increasingly, people on the net need more and more information, the information renewal speed for the amazing, how fast, effective collection, management of these information is the problem that everybody cares. As we know, the current situation, network hardware updates can not meet demand, network bandwidth as information transmission and electronic communication, exchange bottleneck. The appearance of mobile agent technology in a certain extent alleviated the problem of Mobile agent communication, which is autonomous, collaborative and mobility, and has distributed information retrieval, data mining, mobile computing, electronic commerce, system management and other fields have a wide application prospect. Due to dynamic network, every time the result is not the same, but when the difference is 0.0 m s only, M PI + C method of comparison can be seen, Agent based fast sorting algorithm is quite satisfactory. Because the limitations of the mobile platform itself of Java, plus expenses, make than MPI + C, The realization of quick sort for a long time some. This paper make a brief introduction of the characteristics of mobile agent firstly, and then analyzed Aglets, a specific platform of Mobile agent. Finally, it presented a case about a comparison between MPI and C based on quick sort, its result was satisfactory.

Keywords: Mobile agent communication, Quick Sort, Aglet, electronic commerce.

1 Introduction

Mobile Agent[1-5] is a special class of Agent, besides the advantages of intelligent Agent 's basic characteristics: reactive, autonomy, object oriented and aimed at environment, also has the mobility, i.e. it is available on the web from one host to independently moved to another host, on behalf of the user to complete the assigned

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 295–300.
springerlink.com

task. Mobile Agent in heterogeneous network environment of soft, hardware of mobile. Mobile Agent calculation model can effectively reduce the distributed computing network load, improve the communication efficiency, support the disconnected operation, support asynchronous interactive teaching, can dynamically adapt to the network environment, has the advantages of safety and tolerance. Based on the traditional client / server technology compared to the distribution, mobile agent in many respects to the distribution technology for the realization of a better solution, comparatively speaking, it has the advantages: first, whether local or Interne T, there are network connection problems of instability, the resulting network communication abnormal task cannot be completed, normal. And mobile agent is transmitted to the remote host, so as to avoid the above problems. Second, because the network hardware limitations, now the network transmission speed is not very fast, so on the network for information transmission, switching delay is large. Mobile agent technology can integrate information query task are transferred to a server implementation, finally the query results returned, thereby eliminating the need for intermediate results of transmission. In third, mobile agent technology will be a task distribution to different hosts, finally will be the implementation of the results to the original host, thus reducing the workload of each agent, improve work efficiency. From the three points of analysis we can draw a conclusion: the advantages of mobile agents through the calculation of transfer, interactive localization reduces the use of network bandwidth, high speed. Mobile Agent is originated in the distributed system based on distributed technology, is a result of the gradual development. In telnet with the rapid development of mobile Agent technology has brought new opportunities, in the face of the application areas of Internet, mobile Agent technology increasingly shows its strong technological advantages and good development prospect. The word "Agent" will move the Agent and have decades of history of intelligent Agent, Agent together, so that the mobile Agent technology has become an important research field of Agent branch.

2 Aglet of Mobile Agent Platform

Aglets [6-8] by IBM in Tokyo Research Laboratory, is the most complete commercialization of mobile agent system. Now has an aglet1.2 version of the test, we can in the official website to download to IBM, it provides more complete instructions for use. The Aglets system is developed by java language, each aglet agent is a Java object, it can from the network in a mobile host to another. That is to say, a host computer running aglet can suddenly stopped running, distributed to a remote host, continue with perform. When Aglet moves, its program code will carry with it's all the object state transfer. When a host is not sure with the aglet, a built-in security mechanism can make the system safe. In Aglets, has an interface, for each aglet agent provides generic methods and the other agent communication. All agents are operating in the agent environment. The simple diagram (Figure 1) are as follows: by the master Agent of Master is responsible for accepting the task, and then decompose the task then assigned to different from Agent of Slave to perform. Here, allowing the main Agent to delegate tasks to from Agent of Slave, from Agent of Slave moved to the specified destination, to complete the assigned task returns the

result. Aglet development and runtime environment provides a Java class library, support for the aglet to create and run. Create an aglet, use the aglet, can override the method to customize the behavior of aglet. Heavy OnCreation( ) can initialize a aglet. Aglet has a run( ) method, it shows that aglet main thread entrance point. Whenever a aglet to a new aglet engine, run( ) method will be called.

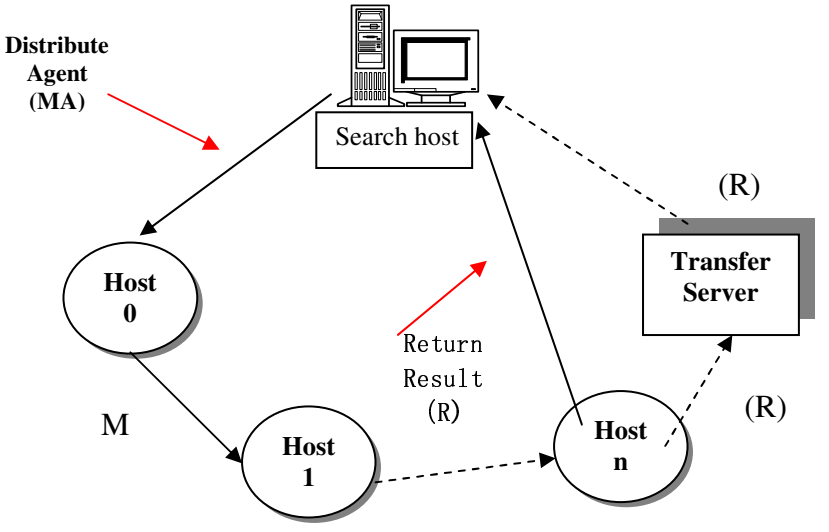


Fig. 1. Mobile Agent

In an aglet lives during the main event occurred before, would call a “callback” method, allows the aglet to the prepared (or refuse to participate in. This is an aglet to know it will be serialized. In the main part of OnDispatch(), aglet must decide whether to go. If aglet decided not to go, will throw an exception, otherwise, must be completed for any unfinished affairs as well as the serialized state ready. When it is from OnDispatch() returns, its state will be serialized, it's all thread will be terminated. Class file and the serial state will be transmitted to new hosts, where aglet will rise. The aglet class defines five callback method, they can be customized by overloading the aglet behavior:

OnCloning()--in the cloning operation prior to being called ;OnDispatch() -- in the dispatch operation prior to being called; OnRevering()-- in the extraction operation prior to being called; OnDeactivating ( ) -- in the activity before being called; OnDisposing() -- in the clear operation (with real life is not the same, if a aglet want to die, it can throw an exception) before being called. The treatment process for each method, the aglet class has a corresponding method to trigger action: Clone ( ), dispatch ( ), retract ( ), deactivate ( ) and dispose ( ). This method is called for a period of time; the aglet host calls the callback method. Whenever a aglet in a host computer is started, the host will provide aglet calls an initialization method. When initialized when the method returns, the host calls the run aglet. According to promote new life events, aglet host will decide to call the following four initialization method:

OnCreation() -- in a aglet first gain life; OnClone() - at the completion of a clone operation; OnArrival() -- in the dispatch or extraction after calling; OnActivation()-- in activated after the transfer.

3 Method

The basic serial quick sort algorithm, in order to make the problem has practical significance, setting a row of data to be stored in a node computer, and asks the sorted data is still in the node computer. According to the quick sort method of partition, algorithm adopts master-slave model, first by the main process in the host to be sorted array A is divided into two blocks (because there are two Agent host), and then put the two pieces of data are sent to two from the process, each from the process in its host of receiving data serial sort, and generates a two block is assigned to another host, so recursively. The final sorted all the data back to the main process and put it in the array A in the original position, complete the sorting process. The two hosts the Agent communication diagram as shown in figure 2.

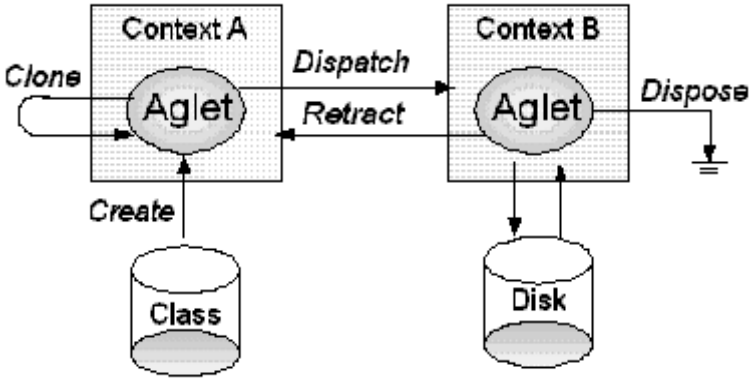


Fig. 2. Mobile agent Communication

The basic idea of the algorithm is as follows:

The first step: The elements of u-1 bit switch to sequence the front bit of A_{00}, which formed sequence of A_{10}; the elements of u-1 bit switch to the 0 bit of the exchange to the back, which formed a sequence of A_{11}, so the A is divided into two sub sequence A_{10}, A_{11}.

The second step: Parallel processing of A_{10}, A_{11}: The elements of u-2 bit to elements of 1 switching to the front of the formation of sequence, which formed $A_{2,2j}$; and the remaining elements of the exchange to the back, which formed $A_{2,2j+1}$ sequence. In this way, it will be divided into four sub sequence of A_{20}, A_{21}, A_{22}, A_{23} …….

The Step third: $A_{u,j}(j=0,1,\cdots,2^{(u-1)}-1)$ parallel processing: The zero bits for the 1 elements of the exchange to the front of the formation of sequence of $A_{u-1,j}(j=0,1,\cdots,2^{(u-1)}-1)$, which formed $A_{u,2j}$, the remaining elements of the exchange to the back, forming a sequence $A_{u,2j+1}$. This A is divided into sub sequence: $A_{u,0}$, $A_{u,1}$,..., $A_{u,2\text{^}u-1}$. In these sequences, some may contain many elements, but with a subset of all elements of a sequence are equal.

4 Result

The experimental environment is to use a fast Ethernet connection of 2 Dell machine structure based on Agent and MPI platform. In contrast, in the same environment, using MPI based parallel quick sort calculation program were experimentally [8], two kinds of methods for run time is shown in Table 1

Table 1. Agent-based and MPI-Based Compared

	N=2000	N=20000	N=50000
C+MPI	0.611	6.064	15.490
Mobile Agent	0.752	6.896	17.075
Serial Quick Sort	1.600	15.000	31.000

(Units:ms)

Due to dynamic network, every time the result is not the same, but difference is 0 ms only, so the values in Table 1 are taken with a 40 average value. As can be seen from the table, and the MPI+C method of comparison can be seen, Agent based fast sorting algorithm is quite satisfactory. Java itself as constraints, coupled with the mobile platform itself overhead, so than using MPI+C achieves rapid sequencing for a long time some.

5 Discuss

The mobile agent technology is a new distributed computing model, now by more and more company, University, Research Institute of wide concern [26]. In this paper, the quick sort algorithm is realized based on the Agent algorithm, and MPI+C algorithm were compared, although it is more than a little overhead, but always make people satisfied. If the mobile agent platform and Java virtual machine further optimization, believe that this gap will become smaller and smaller, of course, this is only in terms of performance and do some work, the next step I will do more work from the aspects of safety. Because the Agent itself may be influenced by a number of malicious code, if they are not protected, then the result is not correct.

Acknowledgment. Mrs. Yan helps me for the languages. There are other colleagues, hereon, to thank. This work is supported by IT projects of Jiangxi Office of Education [No. GJJ10273] and Nature Project of Jiangxi University [No.XYKJ1004] .The authors are grateful for the anonymous reviewers who made constructive comments.

References

1. The Voyage SOA platform, PP 2-5, Recusion Software, Inc. (September 9, 2005)
2. http://www.javaworld.com/javaworld/jw-04-1997/jw-04-agents.html
3. Nwana, H.S.: Nwana, Software Agents: An Overview. Knowledge Engineering Review 11(3), 205–244 (1996)
4. Nwana, H.S., Ndumu, D.T.: A Perspective on Software Agents Research. The Knowledge Engineering Review 14(2), 1–18 (1999)
5. Bradshaw, J.M.: An Introduction to Software Agents. In: Bradshaw, J.M. (ed.) Software Agents. MIT Press, Cambridge (1997)
6. Marc, S., Steve, O., Steven, H.L., et al.: MPI: The Complete Reference. MIT Press, Cambridge (1996)
7. Lange, D.B., Oshima, M.: Programming and Deploying Java mobile agents with Aglets. Addison-Wesley, Reading (1998)
8. Wong, D., et al.: Java-based Mobile Agents. CACM 42(3), 92–102 (1999)

Backward P-Reasoning and Fault Status Identification-Discovery in Electronic System

Chengxian Fan[1] and Jiahe Liu[2]

[1] School of Electrical Engineering, Shandong University, Jinan, Shandong, 250061 China
[2] School of Electrical Engineering and Automation, Shandong Polytechnic University, Jinan, 250353 China
fcxmar@sdu.edu.cn, hitljh@sina.com

Abstract. Using P-reasoning(P=Packet), backward P-reasoning is presented. Backward P-reasoning is consisted of internal backward P-reasoning and outer backward P-reasoning. Backward P-reasoning has dynamic characterisitic, while the fault status in electronic system also has the dynamic characteristic. The identification-discovery of elctronic fault status is to find the location and property of fault by using the information collected, and the fault is the attribute set α, so the fault status discovery-identification in electronic system is similar to backward P-reasoning. Under a certain condition, backward P-reasoning can be restored to the general reasoning. The paper gives the structure of backward P-reasoning and some theory results. Using the study results, the application of backwark P-reasoning in fault status identification- discovery of electronic system is proposed.

Keywords: backward P-reasoning, fault status, identification-discovery, elctronic system.

1 Introduction

The opposite problem to [1] is presented in the paper: 1°. if information $X_k^{\bar{F}}$, $X_{k+1}^{\bar{F}}$ satisfy $X_{k+1}^{\bar{F}} \Rightarrow X_k^{\bar{F}}$, what will the attribute set α_k^F of $X_k^{\bar{F}}$ become? or if $X_{k+1}^{\bar{F}} \Rightarrow X_k^{\bar{F}}$, then $\alpha_k^F \Rightarrow ?$; 2°. if information X_k^F, X_{k+1}^F satisfy $X_{k+1}^{\bar{F}} \Rightarrow X_k^{\bar{F}}$, what will the attribute set $\alpha_k^{\bar{F}}$ of X_k^F become?; 3°. if information $X_k^{\bar{F}}, X_{k+1}^{\bar{F}}$, X_k^F, X_{k+1}^F satisfy $(X_{k+1}^{\bar{F}}, X_k^F) \Rightarrow (X_k^{\bar{F}}, X_{k+1}^F)$, then $(\alpha_k^F, ?) \Rightarrow (?, \alpha_k^{\bar{F}})$. Obviously, the opposite problem has also "sequence characteristic". For example, in 1°, $X_{k+1}^{\bar{F}} \Rightarrow X_k^{\bar{F}}$ appears before $\alpha_k^F \Rightarrow ?$. Because the existence of 1°～3°, backward P-reasoning is presented, which tells people: if the information element becomes, how will the attribute becomes? For example, in the identification of electronic system, it can known the location and property of fault using the information collected by backward P-reasoning, i.e. the attribute set α.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 301–307.
springerlink.com

Backward P-reasoning and P-reasoning are two kind of different reasoning models, which acquire the application[2-4] in different area of dynamic information system.

Using P-reasoning[1] and the characteristics of P-sets[5-10], section **2** gives:

2 Backward P-Reasoning and Its Structure

Literature [1] gives:

$$\text{if} \quad X_n^{\bar{F}} \subseteq X_{n-1}^{\bar{F}} \subseteq \cdots \subseteq X_2^{\bar{F}} \subseteq X_1^{\bar{F}} \tag{1}$$

$$\text{then} \quad \alpha_1^F \subseteq \alpha_2^F \subseteq \cdots \subseteq \alpha_{n-1}^F \subseteq \alpha_n^F \tag{2}$$

$$\text{if} \quad X_1^F \subseteq X_2^F \subseteq \cdots \subseteq X_{n-1}^F \subseteq X_n^F \tag{3}$$

$$\text{then} \quad \alpha_n^{\bar{F}} \subseteq \alpha_{n-1}^{\bar{F}} \subseteq \cdots \subseteq \alpha_2^{\bar{F}} \subseteq \alpha_1^{\bar{F}} \tag{4}$$

where, α_k^F is the attribute set of $X_k^{\bar{F}}$, and $\alpha_k^{\bar{F}}$ is that of X_k^F, $k=1,2,\cdots,n$.

Obviously, if (1) is defined as condition, (2) is the conclusion of (1). If (3) is defined as condition, (4) is the conclusion of (3). In other words, the attribute change of internal P-set can be obtained by change of internal P-set, the attribute change of outer P-set obtained by change of outer P-set. Thus it can be seen that a new reasoning form is hidden in (1)-(4), which is called backward backward P-reasoning.

Definition 1. if $\alpha_k^F, \alpha_{k+1}^F$ are the attribute sets of $X_k^{\bar{F}}, X_{k+1}^{\bar{F}}$, (5) is called internal backward P-reasoning generated by internal P-set, briefly called internal backward P-reasoning, moreover:

$$\text{if} \quad X_{k+1}^{\bar{F}} \Rightarrow X_k^{\bar{F}}, \text{ then} \quad \alpha_k^F \Rightarrow \alpha_{k+1}^F \tag{5}$$

$X_{k+1}^{\bar{F}} \Rightarrow X_k^{\bar{F}}$ is called internal backward P-reasoning condition, and $\alpha_k^F \Rightarrow \alpha_{k+1}^F$ is called internal backward P-reasoning conclusion.

Definition 2. if $\alpha_k^{\bar{F}}, \alpha_{k+1}^{\bar{F}}$ are the attribute sets of X_k^F, X_{k+1}^F, (6) is called outer backward P-reasoning generated by outer P-set, briefly called outer backward P-reasoning, moreover:

$$\text{if} \quad X_k^F \Rightarrow X_{k+1}^F, \text{ then} \quad \alpha_{k+1}^{\bar{F}} \Rightarrow \alpha_k^{\bar{F}} \tag{6}$$

$X_k^F \Rightarrow X_{k+1}^F$ is called outer backward P-reasoning condition, and $\alpha_{k+1}^{\bar{F}} \Rightarrow \alpha_k^{\bar{F}}$ called outer backward P-reasoning conclusion.

Definition 3. (7) is called Backward P-reasoning generated by internal backward P-reasoning and outer backward P-reasoning, briefly called backward P-reasoning, moreover:

$$\text{if} \quad \left(X_{k+1}^{\bar{F}}, X_k^F\right) \Rightarrow \left(X_k^{\bar{F}}, X_{k+1}^F\right), \quad \text{then} \quad \left(\alpha_k^F, \alpha_{k+1}^{\bar{F}}\right) \Rightarrow \left(\alpha_{k+1}^F, \alpha_k^{\bar{F}}\right) \tag{7}$$

$\left(X_{k+1}^{\bar{F}}, X_k^F\right) \Rightarrow \left(X_k^{\bar{F}}, X_{k+1}^F\right)$ is called backward P-reasoning condition, and $\left(\alpha_k^F, \alpha_{k+1}^{\bar{F}}\right) \Rightarrow \left(\alpha_{k+1}^F, \alpha_k^{\bar{F}}\right)$ called backward P-reasoning conclusion.

where, $\left(X_{k+1}^{\bar{F}}, X_k^F\right) \Rightarrow \left(X_k^{\bar{F}}, X_{k+1}^F\right)$ denotes $X_{k+1}^{\bar{F}} \Rightarrow X_k^{\bar{F}}$, $X_k^F \Rightarrow X_{k+1}^F$; $\left(\alpha_k^F, \alpha_{k+1}^{\bar{F}}\right) \Rightarrow \left(\alpha_{k+1}^F, \alpha_k^{\bar{F}}\right)$ denotes $\alpha_k^F \Rightarrow \alpha_{k+1}^F$, $\alpha_{k+1}^{\bar{F}} \Rightarrow \alpha_k^{\bar{F}}$.

By definitions 1-3, it can be got:

Theorem 3 (attribute supplemented theorem of internal backward P-reasoning) if there exists set $\nabla X_{k+1}^{\bar{F}}$, $\nabla X_{k+1}^{\bar{F}} \subset X_k^{\bar{F}}$, then attribute $\nabla \alpha_{k+1}^F$ is supplemented into the attribute set α_k^F of $X_k^{\bar{F}}$ and satisfies:

$$\text{if } \left(X_k^{\bar{F}} - \nabla X_{k+1}^{\bar{F}}\right) \Rightarrow X_k^{\bar{F}}, \text{ then } \alpha_k^F \Rightarrow \left(\alpha_k^F \cup \nabla \alpha_{k+1}^F\right) \tag{8}$$

Theorem 4 (attribute deleted theorem of outer backward P-reasoning) if there exists set $\Delta X_{k+1}^F \neq \phi$, then attribute $\Delta \alpha_k^{\bar{F}}$ is deleted from the attribute set $\alpha_k^{\bar{F}}$ of X_k^F and satisfies:

$$\text{if } X_k^F \Rightarrow \left(X_k^F \cup \Delta X_{k+1}^F\right), \text{ then } \left(\alpha_k^{\bar{F}} - \Delta \alpha_{k+1}^{\bar{F}}\right) \Rightarrow \alpha_k^{\bar{F}} \tag{9}$$

Theorem 5 (attribute supplemented-deleted theorem of backward P-reasoning) if there exists sets $\nabla X_{k+1}^{\bar{F}}$, ΔX_{k+1}^F, then attribute $\nabla \alpha_{k+1}^F$ is supplemented into α_k^F, $\Delta \alpha_k^{\bar{F}}$ is deleted from $\alpha_k^{\bar{F}}$ and satisfies:

$$\begin{aligned} &\text{if } \left((X_k^{\bar{F}} - \nabla X_{k+1}^{\bar{F}}), X_k^F\right) \Rightarrow \left(X_k^{\bar{F}}, (X_k^F \cup \Delta X_{k+1}^F)\right), \\ &\text{then } \left(\alpha_k^F, (\alpha_k^{\bar{F}} - \Delta \alpha_{k+1}^{\bar{F}})\right) \Rightarrow \left((\alpha_k^F \cup \nabla \alpha_{k+1}^F), \alpha_k^{\bar{F}}\right) \end{aligned} \tag{10}$$

Where, $\left((X_k^{\bar{F}} - \nabla X_{k+1}^{\bar{F}}), X_k^F\right) \Rightarrow \left(X_k^{\bar{F}}, (X_k^F \cup \Delta X_{k+1}^F)\right)$ denotes $(X_k^{\bar{F}} - \nabla X_{k+1}^{\bar{F}}) \Rightarrow X_k^{\bar{F}}$, $X_k^F \Rightarrow (X_k^F \cup \Delta X_{k+1}^F)$; $\left(\alpha_k^F, (\alpha_k^{\bar{F}} - \Delta \alpha_{k+1}^{\bar{F}})\right) \Rightarrow \left((\alpha_k^F \cup \nabla \alpha_{k+1}^F), \alpha_k^{\bar{F}}\right)$ denotes $\alpha_k^F \Rightarrow (\alpha_k^F \cup \nabla \alpha_{k+1}^F)$, $(\alpha_k^{\bar{F}} - \Delta \alpha_{k+1}^{\bar{F}}) \Rightarrow \alpha_k^{\bar{F}}$.

Theorem 6 (the first relation theorem between backward P-reasoning and general reasoning) if $F = \bar{F} = \phi$, then backward P-reasoning and general reasoning satisfy:

$$\left\{\text{if } (X_k^{\bar{F}} \Rightarrow X, X \Rightarrow X_k^F), \quad \text{then } (\alpha \Rightarrow \alpha_k^F, \alpha_k^{\bar{F}} \Rightarrow \alpha)\right\}_{\bar{F}=F=\phi} = \left\{\text{if } X^*, \text{then } \alpha^*\right\} \tag{11}$$

Theorem 7 (the second relation theorem between backward P-reasoning and general reasoning) if $F = \bar{F} = \phi$, then backward P-reasoning and general reasoning satisfy:

$$\left\{\text{if } (X_k^{\bar{F}} \Rightarrow X, X \Rightarrow X_k^F), \quad \text{then } (\alpha \Rightarrow \alpha_k^F, \alpha_k^{\bar{F}} \Rightarrow \alpha)\right\}_{\bar{F}=F=\phi} = \left\{\text{if } X^o, \text{then } \alpha^o\right\} \tag{12}$$

Using the results in section **2**, **3** gives:

3 Attribute Residual Discovery Generated by Backward P-Reasoning

Definition 4. Given internal backward P-reasoning

$$\text{if } X_{k+1}^{\bar{F}} \Rightarrow X_k^{\bar{F}}, \quad \text{then} \quad \alpha_k^F \Rightarrow \alpha_{k+1}^F \tag{13}$$

$\nabla\alpha_k^F$ is called internal backward P-attribute residual of α_k^F generated by internal backward P-reasoning, briefly called internal-attribute residual, if

$$\nabla\alpha_k^F = \alpha_{k+1}^F - \alpha_k^F \tag{14}$$

Definition 5. Given outer backward P-reasoning

$$\text{if } X_k^F \Rightarrow X_{k+1}^F, \quad \text{then} \quad \alpha_{k+1}^{\bar{F}} \Rightarrow \alpha_k^{\bar{F}} \tag{15}$$

$\Delta\alpha_{k+1}^{\bar{F}}$ is called outer backward P-attribute residual of $\alpha_{k+1}^{\bar{F}}$ generated by outer backward P-reasoning, briefly called outer-attribute residual, if

$$\Delta\alpha_{k+1}^{\bar{F}} = \alpha_k^{\bar{F}} - \alpha_{k+1}^{\bar{F}} \tag{16}$$

Definition 6 The set pair $\left(\nabla\alpha_k^F, \Delta\alpha_{k+1}^{\bar{F}}\right)$ consisted of $\nabla\alpha_k^F$ and $\Delta\alpha_{k+1}^{\bar{F}}$ is called attribute residual of backward P-reasoning, briefly called backward P-reasoning attribute residual. $\nabla\alpha_k^F$ is internal-attribute residual, and $\Delta\alpha_{k+1}^{\bar{F}}$ is outer-attribute residual.

$$\left(\nabla\alpha_k^F, \Delta\alpha_{k+1}^{\bar{F}}\right) \tag{17}$$

By definitions 4-6, it can be got:

Theorem 8 (the discovery theorem of internal-attribute residual) if $X_{k+1}^{\bar{F}} \Rightarrow X_k^{\bar{F}}$ is internal backward P-reasoning condition, then the attribute set α_{k+1}^F and α_k^F satisfy:

$$\alpha_k^F \Rightarrow \left(\alpha_k^F \cup \nabla\alpha_k^F\right) \tag{18}$$

$\nabla\alpha_k^F$ is internal-attribute residual of α_k^F, and $\nabla\alpha_k^F$ is discovered out α_k^F.

Theorem 9 (the discovery theorem of outer-attribute residual) if $X_{k+1}^{\bar{F}} \Rightarrow X_k^{\bar{F}}$ is outer backward P-reasoning condition, then the attribute set $\alpha_k^{\bar{F}}$ and $\alpha_{k+1}^{\bar{F}}$ satisfy:

$$\alpha_{k+1}^{\bar{F}} \Rightarrow \left(\alpha_{k+1}^{\bar{F}} \cup \Delta\alpha_{k+1}^{\bar{F}}\right) \tag{19}$$

$\Delta\alpha_{k+1}^{\bar{F}}$ is outer-attribute residual of $\alpha_{k+1}^{\bar{F}}$, and $\alpha_{k+1}^{\bar{F}}$ is discovered in $\alpha_{k+1}^{\bar{F}}$.

Theorem 10. (the discovery theorem of backward P-reasoning attribute residual) if $X_k^{\bar{F}} \Rightarrow X$, $X \Rightarrow X_k^F$ are backward P-reasoning condition, then there must exist non-empty attribute set pair $\left(\nabla\alpha_k^F, \Delta\alpha_k^{\bar{F}}\right)$ to satisfy:

$$\left(\alpha_k^F, \alpha_k^{\bar{F}}\right) \Rightarrow \left((\alpha_k^F \cup \nabla\alpha_k^F), (\alpha_k^{\bar{F}} \cup \Delta\alpha_k^{\bar{F}})\right) \tag{20}$$

Using the results of sections **2** and **3**, **4** gives:

4 The Fault Status Identification-Discovery Application of Backward P-Reasoning in Electronic System

Assumption: In the output of the elctronic system, if the output voltages or currents deviate the normal output, the deviation value is regarded as the empietement for the normal output, which is denoted in the set as a new output.

That the attribute change in α can be known by the element change in X is the important characteristic of backward P-reasoning. It can be known that some attributes are delted from α through supplementing some elements into X, and it is the important characteristic of outer backward P-reasoning, which complies with the chararcteristic of fault iedentification in elctronic system. Using the characteristic, the paper gives an application example about the fault status identification in elctronic system. The data in the example come from the output voltages and currents of a certain key laboratory in a certain university. The normal output data are denoted with data set $X = \{x_1, x_2, x_3, x_4, x_5\}$, which is shown in the table 1. The status of lines or elements are denoted with attribute set $\alpha = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$.

Table 1. The normal output data X

	1	2	3	4	5
X	1.23	1.34	1.56	1.68	1.65

If the output x_2, x_3 have the abanormal output, then set X becomes $X_1^F = \{x_1, x_2, x_3, x_4, x_5, x_2', x_3'\}$, $X \subseteq X_1^F$, which is shown in table 2.

Table 2. Abanormal output data X_1^F

	1	2	3	4	5	6	7
X_1^F	1.23	2.56	3.28	1.68	1.65	1.22	1.72

According to backwark P-reasoning (6) "if $X_k^F \Rightarrow X_{k+1}^F$, then $\alpha_{k+1}^{\bar{F}} \Rightarrow \alpha_k^{\bar{F}}$", it can be got that attribute set α becomes $\alpha_1^{\bar{F}} = \{\alpha_2, \alpha_4\}$, $\alpha_1^{\bar{F}} \subseteq \alpha$.

If the output x_5 also has the abanormal output, then set X becomes $X_2^F = \{x_1, x_2, x_3, x_4, x_5, x_2', x_3', x_5'\}$, $X_2^F \subseteq X_2^F$, which is shown in table 3.

Table 3. Abanormal output data X_2^F

	1	2	3	4	5	6	7	8
X_2^F	1.23	2.56	3.28	0.01	1.65	1.22	1.72	-1.64

According to backwark P-reasoning (6) "if $X_k^F \Rightarrow X_{k+1}^F$, then $\alpha_{k+1}^{\bar{F}} \Rightarrow \alpha_k^{\bar{F}}$", it can be got that attribute set α becomes $\alpha_2^{\bar{F}} = \{\alpha_2\}$, $\alpha_2^{\bar{F}} \subseteq \alpha_1^{\bar{F}}$. So the fault point is determined and treated to guarentee the system normal operation.

5 Conclusion

Using the change of X, the change of α can be found by back P-reasoning. Obviously, back P-reasoning has intelligent characteristic, i.e., if $X_n^{\bar{F}} \subseteq X_{n-1}^{\bar{F}} \subseteq \cdots \subseteq X_2^{\bar{F}} \subseteq X_1^{\bar{F}}$, it can be got $\alpha_1^F \subseteq \alpha_2^F \subseteq \cdots \subseteq \alpha_{n-1}^F \subseteq \alpha_n^F$, or, if $X_1^F \subseteq X_2^F \subseteq \cdots \subseteq X_{n-1}^F \subseteq X_n^F$, it can be got $\alpha_n^{\bar{F}} \subseteq \alpha_{n-1}^{\bar{F}} \subseteq \cdots \subseteq \alpha_2^{\bar{F}} \subseteq \alpha_1^{\bar{F}}$. The reasoning process of backward P-reasoning is similar to the fault status in electronic system, so the fault type and location in electronic system can be found system by back P-reasoning. Backward P- reasoning is a novel mathematics model and method of studying the fault status identification-discovery in electronic system.

References

[1] Shi, K.: P-reasoning and P-reasoning discovery-identification of information. Computer Science 38(7), 1–9 (2011) (in Chinese)
[2] Lin, H., Fan, C., Shi, K.: Backward P-reasoning and attribution residual discovery-application. Computer Science 38(10), 189–198 (2011) (in Chinese)
[3] Lin, H., Fan, C.: The dual form of P-reasoning and identification of unknown attribute. International Journal of Digital Content Technology and its Applications 6(1), 121–131 (2012)
[4] Fan, C., Lin, H.: P-sets and the reasoning-identification of disaster information. International Journal of Convergence Information Technology 7(1), 337–345 (2012)
[5] Shi, K.: P-sets. Journal of Shandong University (Natural Science) 43(11), 77–84 (2008) (in Chinese)
[6] Shi, K.: P-sets and its Applied Characteristics. Computer Science 37(8), 1–9 (2010) (in Chinese)
[7] Shi, K.: P-sets and its applications. An International Journal Advances in Systems Science and Applications 9(2), 209–219 (2009)

[8] Lin, H., Li, Y.: P-sets and its P-separation theorems. An International Journal Advances in Systems Science and Applications 10(2), 209–215 (2010)
[9] Wang, Y., Geng, H., Shi, K.: The mining of dynamic information based on P-Sets and its application. An International Journal Advances in Systems Science and Applications 10(2), 234–240 (2010)
[10] Lin, H., Li, Y.: P-sets and its P-separation theorems. An International Journal Advances in Systems Science and Applications 10(2), 209–215 (2010)

Optimizing the Number and Location of Switching and Protective Devices in Power Distribution Networks

Ignacio J. Ramírez Rosado[1] and Enrique Zorzano Alba[2]

[1] Department of Electrical Engineering, University of Zaragoza, Spain
[2] Department of Electrical Engineering, University of La Rioja, Spain

Abstract. For the power distribution systems planner it is essential to know the best number and in which power network sections, the control and protective devices must be placed. This paper presents a method for determining the optimal number and the optimal locations of sectionalizing switches and protective devices, in order to achieve a reliability improvement of the distribution network systems and to reduce the investment cost. Thus, a multi-objective nonlinear optimization model was used for determining all the non-dominated planning solutions (Pareto's solutions) obtained from the optimization of two objective functions: a) objective function of reliability represented by ASIDI, ASIFI and ENS indexes; and b) objective function of economic investment costs. The constraints are considered taking into account technical and economic limitations imposed by the electric company (i.e. coordination of protective devices and sections prohibited for some devices).

Keywords: Power systems modeling, multiobjective optimization, reliability, power distribution systems, optimal planning.

1 Introduction

Power distribution networks are an essential part of the electrical system. Major power losses occur in such distribution networks, as well as the highest failure indexes that directly affect the power quality supplied to consumers. Because of this, electrical companies must support good reliability indexes or otherwise they will incur penalties [1]. These indexes are regulated by Governmental Legislation [2]; thus the companies, the owners of the networks, must carry out optimal planning in order to reach the established reliability indexes.

In short or medium-term time-horizon planning of power distribution networks, appropriate investment decisions should be taken in order to achieve significant improvements in network reliability with the installation of new switching and protective devices. The decisions of the planner will be developed in a multiobjective framework, where, simultaneously, the technical and economic aspects have to be evaluated, in order to obtain the best planning solutions through multiobjective optimization [3]. Some models seek to reduce losses of energy and investment costs looking for the optimal locations of sectionalizing devices [4]. Other models present a

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 309–316.
springerlink.com

formulation for sectionalizing device placement taking into consideration outage, maintenance and investments costs. A solution methodology based on the optimization technique of simulated annealing is proposed to determine the number and the locations of the sectionalizing switches [5].

Through optimization methods, based on genetic algorithms, where the principal objective function is to minimize the investment cost, the optimal locations [6] and optimal number of sectionalizing switches have been determined [7]. The multiobjective ant colony optimization has been applied in others problems to minimize the total cost while simultaneously minimizing several distribution network reliability indexes [8]. In order to determine the solutions for this problem, other mathematical models have used mixed integer nonlinear programming MINLP [9] establishing a codification for each branch of the feeder that is a possible candidate: this is represented by a value, depending on which device has been allocated.

This paper presents a multiobjective optimization method to obtain the best reliability indexes with minimal investment costs in distribution networks. We have used two objective functions to be minimized simultaneously: a) an objective function representing a reliability index (i.e. ASIDI, ASIFI and ENS) and/or other indexes under the Spanish Legislation [9]; b) an objective function representing economical investment cost for the installation of new switching and protective devices.

This work presents an enumerative methodology to achieve the multiobjective optimal planning solutions. From this set of solutions, the non-dominated ones were selected. A solution is non-dominated if there is no other solution that can improve at least an objective function without worsening any other objective one [1]. These planning solutions represent ideal locations for building new switching and protective devices in order to simultaneously improve the reliability indexes and investment costs. Initially we must take into account that the planning of the main feeder network expansion has ended. The next step is the installation of new switching and protection devices to minimize the reliability indexes. The devices cannot be placed in any section of the network, since there are technical constraints and restrictions imposed by the electrical company.

The control and protective devices, which have been considered in this work, are:

Recloser (REC): this element has opening capacity and closes in short circuit. RECs can be installed mainly in feeder heads and important derivations of them. It can be operated by remote control.

Disconnect switch (SERC): this device does not have capacity of opening in short circuit. It can be operated manually or by remote control, although in this paper sectionalizing switch by remote control only was considered. It is mainly used as the main feeder sectionalizing sections, although it can be placed anywhere on the network.

Disconnector (SEM): it allows the opening of loads. It is usually located in the starting of derivations and feeding to consumers. It is manually operated.

Sectionalizer (SLZ): its operation, properly coordinated with the reclosing cycles of switch line head or recloser, allows the elimination of faulty section before completing the cycles, avoiding the definitive interruption of the service in the sections without fault. This has importance in the elimination of the temporary failures, which represent almost 90% of the failures in the distribution network.

Fuses (FUS): this element will be installed in the starting of the derivation of supply line to the consumer.

2 Methodology

A multiobjective optimization problem considers a p-dimensional vector of objective functions **f**(x) in the feasible region **X**.

$$\mathbf{f}(x) = [f_1(x), f_2(x), ..., f_k(x), ..., f_p(x)] . \quad (1)$$

$$\mathbf{X} = \{x:x \in \mathbf{R}^n, h_i(x) \leq 0, x_j \geq 0\} \text{ for all i and j} . \quad (2)$$

$$i = 1, 2, ..., I ; \qquad j = 1, 2, ..., J .$$

where $f_k(x)$ are the objective functions; $h_i(x)$ are the mathematical constraints; and **R** is the set of real numbers.

Multiobjective optimization aims to determine the set non-dominated solutions in the region **X**. This optimization can be described as:

$$\text{max-dominated } \mathbf{f}(x) . \quad (3)$$

Subject to

$$x \in \mathbf{X} . \quad (4)$$

If **P** is the set of non-dominated solutions (Pareto's solutions), then:

$$\mathbf{P} = \{x:x \in \mathbf{X}, \text{ there is no other } x' \in \mathbf{X} \text{ such that}$$

$$f_k(x') > f_k(x), \text{ for some } k \in \{1, 2, ..., p\}$$

$$\text{and } f_q(x') \geq f_q(x), \text{ for all } q \neq k\} . \quad (5)$$

This paper describes a non-linear multiobjective optimization model to obtain the best number and optimal locations of switching and protective devices in power distributions networks. Thus, two objective functions were considered, f_1 and f_2, where f_1 is a reliability index (ENS, ASIDI or ASIFI) and f_2 are the economic investment costs, corresponding to the installation of new devices in the network.

$$ENS = \sum_{i=1}^{k} (PI_i \cdot H_i) \quad \textit{(kWh/year)} . \quad (6)$$

$$\text{ASIDI} = \frac{\sum_{i=1}^{k} (\text{PI}_i \cdot \text{H}_i)}{\sum \text{PI}} \quad \textit{(hours/year)} \cdot \tag{7}$$

$$\text{ASIFI} = \frac{\sum_{i=1}^{k} \text{PI}_i}{\sum \text{PI}} \quad \textit{(interruptions/year)} \cdot \tag{8}$$

where:

ENS (Energy Non Supplied) is defined as the non supplied energy due to system outages during a year,

ASIDI (Average System Interruption Duration Index) is defined as the installed power equivalent interruption time at medium voltage ($1\text{kV} \leq \text{V} \leq 36$ kV),

ASIFI (Average System Interruption Frequency Index) is the installed power equivalent number of interruptions at medium voltage,

k, represents the total number of interruptions during the period considered,

H_i, is the supply interruption time affecting the power PI_i (in hours),

PI_i, is the installed power (in kVA), in the distributor's MV/LV transformers stations plus the power (in kVA), contracted at MV affected by interruption i, associated to H_i ,

$\sum PI$, is the total sum of the installed power (in kVA) in the distributor's MV/LV transformers stations plus the power contracted at MV (in kVA).

Therefore, two objective functions were minimized simultaneously, to achieve the best planning solutions of minimum reliability index value (maximum network reliability) with minimum economic investment costs.

The reliability indexes are directly related to supply interruption time H_i. Such unavailability is affected by types of switching and protection devices, and by their installation locations. Therefore, the methodology considers the used time to isolate the fault, the time to locate it, the access and repair time, and finally the time to leave the power network in the original non-faulted state.

3 Case Study

The proposed methodology was applied to several power distribution networks. Fig. 1 shows a portion of a real-life power distribution system, used for our case study. This network, with $\Sigma\text{PI} = 12520$ kVA, is the reference case, with the following reliability indexes: ASIDI = 13.04 hours/year; ASIFI = 4.26 interruptions/year; and ENS = 163200 kWh/year. In this case, the maximum limit of investment is 110000 Euros.

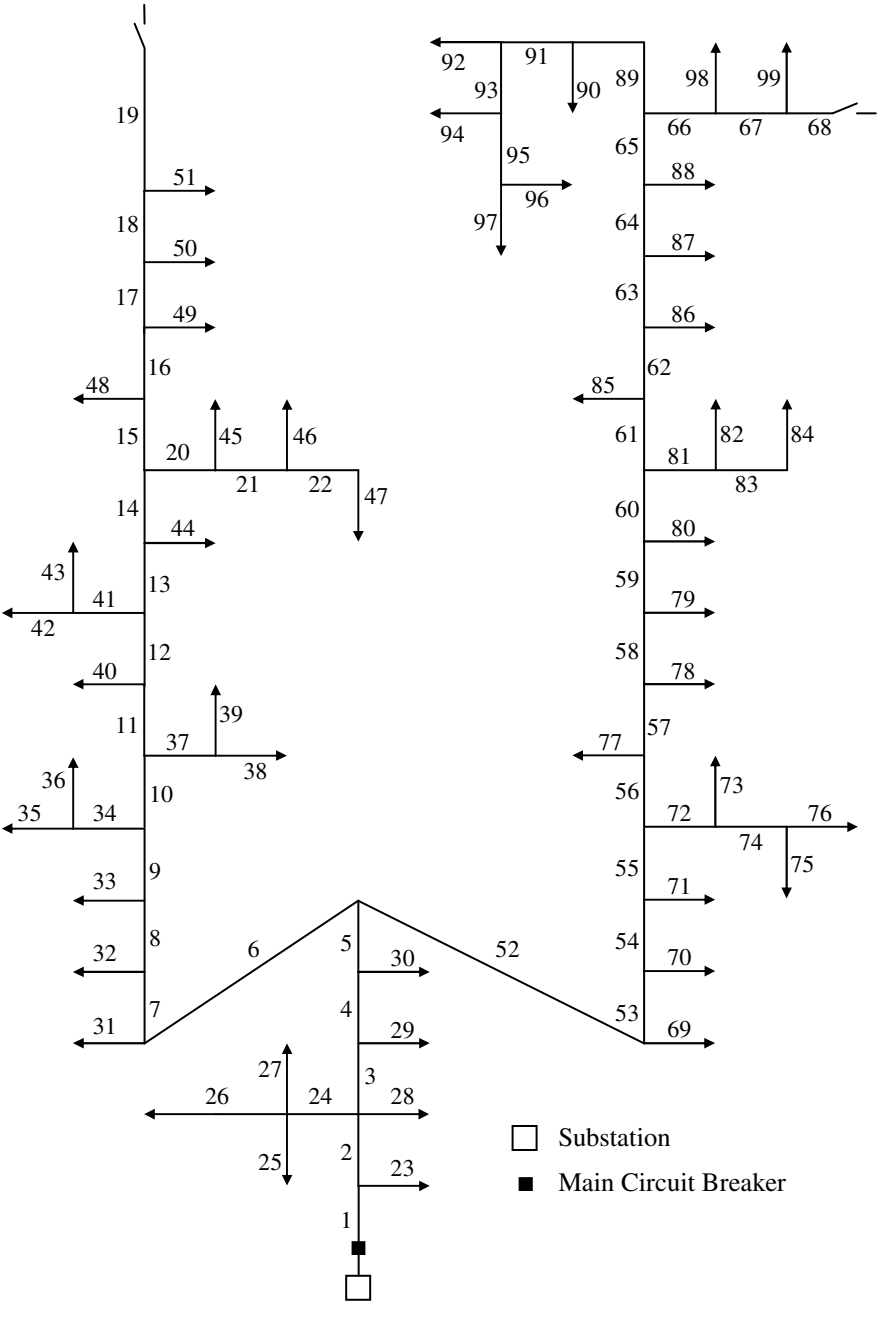


Fig. 1. Real-life power distribution network used as reference case

3.1 Determining the Number of Devices

These tests allow us to determine the number of power devices of each type that has the best ratio in the improvement of the reliability indexes versus investment cost. Firstly, five tests, one for each type of power apparatus, were carried out. For each test, a number of elements were considered, so that it did not exceed the investment limit.

Fig. 2 shows the results obtained in each test. In the first series (curve REC) reclosers (from one to five ones) were installed. The best improvement rate of reliability values appeared in the first part of the curve: for the same increase in investment costs, a substantial improvement in the reliability value was obtained. For example, installing three REC in sections 54, 7 and 65 on network (investment of 62400 Euros), the ASIDI index value was reduced to 4.90 hours/year, i.e. 37.58% of the ASIDI index value of the reference case. This supposes an improvement in ASIDI index of 62.42%. With four REC, the investment cost increased to 83200 Euros and the ASIDI index value decreased to 4.57 hours/year (35.05% of the ASIDI value of the reference case, an improvement of 2.53% with respect to the previous case).

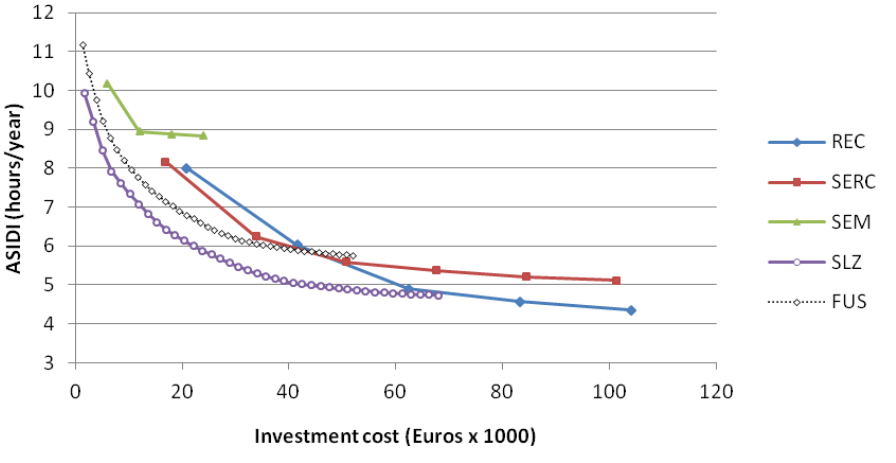


Fig. 2. Test for every type of device. ASIDI versus investment cost

In this network, improvements in the reliability indexes below 3% were not considered in order to determine the number of devices. Therefore, three elements were considered as the maximum number of elements, since this allowed a good practical representation of the costs versus ASIDI.

All this has helped us to establish a criterion for the installation of new devices in the distribution network. Then, using the same procedure, SERC and SEM were considered, up to a maximum of three SERC and two SEM. Fig. 2 shows no more than four elements of SEM series, since increasing the number of devices, the ASIDI index worsens. Finally, sectionalizers and fuses are included. The number of these devices will be obtained from the non-dominated solutions of the next test.

3.2 Determining the Optimal Allocation of Devices

Using the methodology described above, an enumerative procedure was used to identify ideal device installation locations. All feasible power distribution locations were tested, evaluating every device installation in all the possible locations and determining the corresponding reliability indexes.

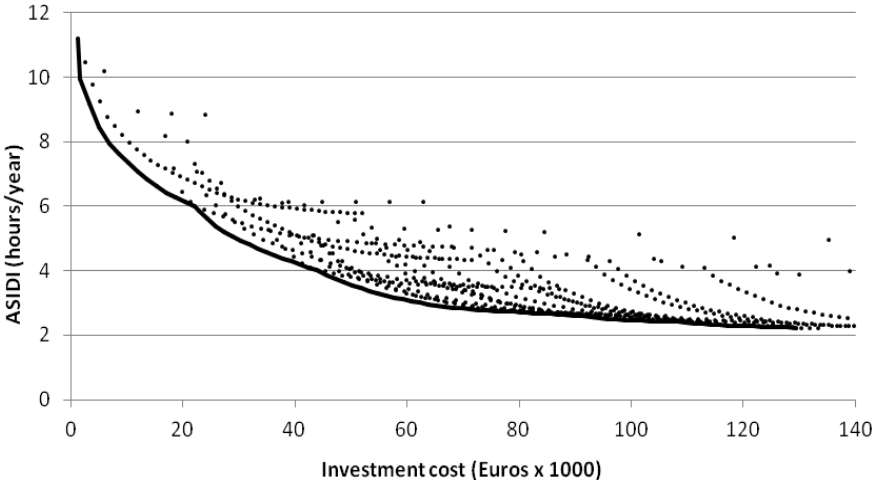


Fig. 3. All the feasible planning solutions obtained for every combination of devices

With the numbers that were obtained in the previous section, the maximum number of devices of each type (for this test) was established. Thus, the combinations with the following elements were performed: 0 to 3 REC, 0 to 3 SERC, 0 to 2 SEM, 0 to 30 SLZ and 0 to 30 FUS. We have chosen a large number of sectionalizers and fuses, in order to achieve a larger number of solutions. Fig. 3 shows all the feasible planning solutions (ASIDI versus investment cost) obtained for each combination of elements. The optimal number and locations of these devices were obtained and, therefore, also the curve of non-dominated solutions (solid line curve in Fig. 3). Table 1 shows some solutions for different investment costs.

Table 1. Optimal numbers and locations of power devices for different investment costs

Number/Type device	Locations	ASIDI (h/y)	ASIFI (int/y)	Investment (Euros)
1SERC, 6SLZ	54, 89, 20, 81, 37, 24, 72	5.38	2.16	25900
2SERC, 11SLZ	54, 7, 89, 81, 20, 24, 72,78, 28, 23, 93, 88, 37	3.55	1.81	50300
2REC, 1SERC, 28SLZ	54, 7 , 65, 81, 20, 24, 72, 93, 78, 28, 23, 88, 69, 37, 29, 41, 85, 86, 95, 77, 89, 34, 99, 94, 48, 32, 40, 70, 92, 90, 31	2.46	0.68	100500

4 Conclusions

This paper presents a methodology for the multiobjective optimization of economic costs and reliability of power distribution systems in order to determine optimal decisions on new investments that achieve the best reliability values (ASIDI, ASIFI and ENS reliability indexes), with the lowest economic investment cost. In the case of the new switching and protection devices, the proposed methodology determined the best locations and the ideal number of reclosers, disconnect switch, disconnector, sectionalizers and fuses to be installed in the distribution system, using a specific procedure: first, reclosers (with the most short switching and restoration times) were installed; followed by disconnect switch, disconnector; and afterwards the sectionalizers and fuses (long switching and restoration times) were installed. The ideal locations of the selected devices were determined using an enumerative procedure, testing the installation of the devices in every distribution system location, and identifying the locations with the best reliability index.

Acknowledgments. The authors would like to thank the "Ministerio de Ciencia e Innovación" for supporting this research under the Project ENE2009-14582-C02-02.

References

1. Billinton, R., Allan, R.: Reliability Evaluation of Power Systems. Pitman Books, England (1984)
2. Spanish Ministry of Economy, Royal Decree 1995/2000, of December 1, for regulation of the activities of transmission, distribution, trading and supply of electrical energy
3. Zeleny, M.: Multiple Criteria Decision Making. McGraw-Hill, New York (1982)
4. Levitin, G., Mazal-Tov, S., Elmakis, D.: Optimal Sectionalizer Allocation in Electric Distribution Systems by Genetic Algorithm. Electric Power Systems Research 31, 97–102 (1994)
5. Billinton, R., Jonnavithula, S.: Optimal Switching Device Placement in Radial Distribution Systems. IEEE Transactions on Power Delivery 11, 1646–1651 (1996)
6. Yang, W., Liu, J., Yu, J., Dong, H., Song, M.: Optimal allocation of switches in distribution networks. In: World Congress on Intelligent Control and Automation, WCICA, pp. 5155–5159 (2004)
7. Liu, B., Kun, S., Zou, J., Duan, X., Zheng, X.: Optimal feeder switches location scheme for high reliability and least costs in distribution system. In: World Congress on Intelligent Control and Automation, WCICA, pp. 7419–7423 (2006)
8. Tippachon, W., Rerkpreedapong, D.: Multiobjective optimal placement of switches and protective devices in electric power distribution systems using ant colony optimization. Electric Power Systems Research 79, 1171–1178 (2009)
9. Wesz Da Silva, L.G., Fernandes Pereira, R.A., Abbad, J.R., Sanches Mantovani, J.R.: Optimized Allocation of Control and Protective Devices in Electric Distribution Systems. Electric Power Components and Systems 38, 1–21 (2010)

A Novel Binary Integrated Circuit Images Encryption Algorithm Based on Chaotic Sequences

Jianjun Wu[1] and Li Tu[2]

[1] Center of Network, Hunan City University,Yiyang, Hunan, 413000, China
[2] School of Computer Science and Engineering, Hunan City University, Yiyang, Hunan, 413000, China
Tujinhe163168@163.com, tulip1907@163.com

Abstract. In this paper, through the introduction and analysis of integrated circuit diagram encryption technology, we have introduced chaos theory and systems models, and how to generate a chaotic sequence. We proposed an improved integrated circuit diagram encryption algorithm based on chaotic sequences, and binary integrated circuit diagram is encrypted and unencrypted using matlab tools by the encryption algorithm based on chaotic sequences. The sensitivity analysis and relevance analysis of the algorithm shows that, this algorithm has high security, and it is fast and easy to implement.

Keywords: Chaotic sequence, Binary integrated circuit diagram, Encryption algorithm.

1 Introduction

The integrated circuit diagram encryption technology originated originated in the early classic cryptography. Cryptography is a method of how to convert readable text (plaintext) to unreadable text (Ciphertext), and how to convert unreadable text (Ciphertext) to readable text (plaintext) through reversible transformation.

Chaotic system has the following characteristics suitable as a password system:

1). Traversal characteristics. In a limited area, the points on the chaotic orbit can be arbitrarily close. This makes it very difficult to predict the initial conditions (plaintext).

2). Mixed characteristics. The anomaly of the chaotic orbit and local expansion, compression, folding make the output of the chaotic system is similar to random noise;

3). Exponential divergence. Two points arbitrarily close in the phase plane will exponentially diverge.

In summary, the chaotic system is a natural system of password. The anomaly of chaos in the two-dimensional phase plane makes chaotic system to be more suitable for the encryption of integrated circuit diagram and data.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 317–322.
springerlink.com

2 Algorithm Principle

2.1 Models of Chaotic Systems

A one-dimensional discrete-time nonlinear dynamical system is defined as follows:

$$X_{k+1} = \in \tau(X_k) \tag{1}$$

And $x_k \in$V, k = 0, 1, 2, … is called the state. Formula τ: V→V is a mapping, it mapped to the current state x_k to the next state X_{k+1}. Started from an initial value x_0, repeated application of parameter τ, we will get a sequence { x_k ;k= 0, 1, 2,…}. This sequence is called the trajectory of the discrete-time dynamical system.

A class of very simple while widely studied dynamical system is the logistic map, it is defined as follows:

$$X_{k+1} = \mu x_k (1 - x_k) \tag{2}$$

And the inequality 0≤μ≤4 is called the branching parameters, $x_k \in$(0,1).When 3.5699456…<μ≤4,the logistic mapping is working in a chaotic state. That is, the sequence $\{x_k; k = 1,2,3...\}$ generated under the initial conditions x0 by the action of logistic map, is a non-cycle, emanative sequence, and it is very sensitive to its initial value. The logistic map can be defined on the interval (-1,1) as follows:

$$X_{k+1} = 1 - \lambda x_k^2 \qquad \lambda \in [0,2] \tag{3}$$

2.2 Algorithm Design

The encryption algorithm as follows:

(1) First converted the binary integrated circuit diagram I to be encrypted into a one-dimensional vector V, and acted the map f on the vector V, got $\tilde{V}$.

(2) Selected the initial value (key) x_0 ,u and the length of the sequence, calculated the chaotic sequence using Logistic map, arranged it in vector p' sequentially, using p' Calculated the matrix P.

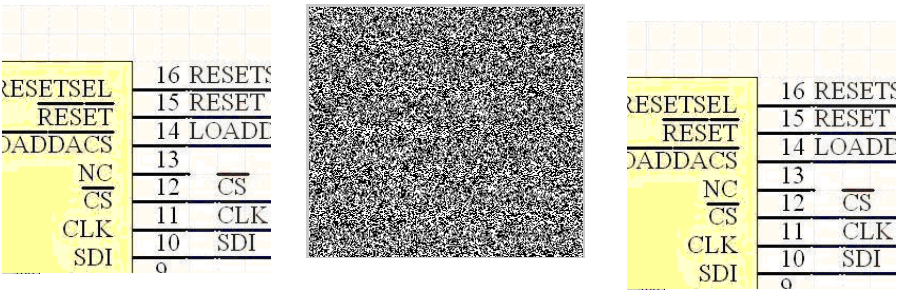
(3) Calculated $\tilde{D} = P \otimes \tilde{V}$, In the of role f^{-1} ,by calculating the vector a $\tilde{D}$ D is converted into D, then arranged it in a new binary integrated circuit diagram $\tilde{I}$

according to the ranks of the relationship of the original binary integrated circuit diagram, it is the encrypted binary integrated circuit diagram.

The decryption algorithm is to input the same key and do the encryption algorithm again.

3 Encrypted Results

After an original integrated circuit diagram was encrypted, the integrated circuit diagrams becomes disorganized and unrecognizable, the encrypted integrated circuit diagram by the correct key to decrypt can be restored to the original integrated circuit diagram. The grayscale integrated circuit diagrams encrypted by the chaotic algorithm. The simulation results are shown as below:



① Original image（256*256） ② Encrypted image ③ Decrypted image

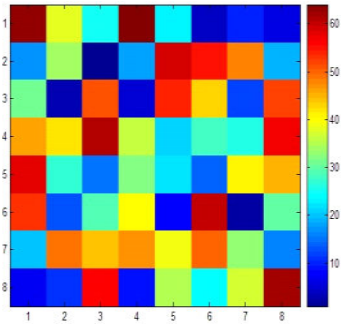
Fig. 1. Effect of encryption

4 Sensitivity Analysis and Relevance Analysis of the Algorithm

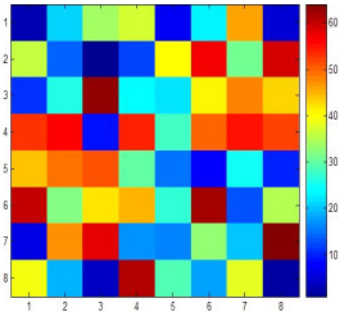
4.1 Sensitivity Analysis of the Algorithm

The sensitivity of the key is the degree of changing of the cipher text when the initial key changes a little. We use the chaos algorithm on matrix transforming to detect the sensitivity of the algorithm.

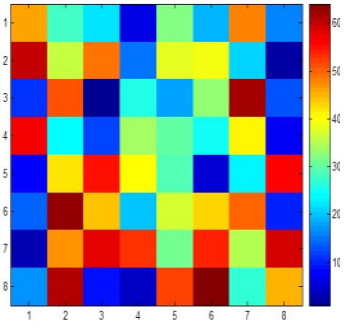
It can be seen from Figure 2, different values of parameter x_0 and parameter r have a great impact on the effect of integrated circuit diagram encryption. When we slightly changed the values of parameter x_0 and parameter r, we can get different encrypted integrated circuit diagrams.



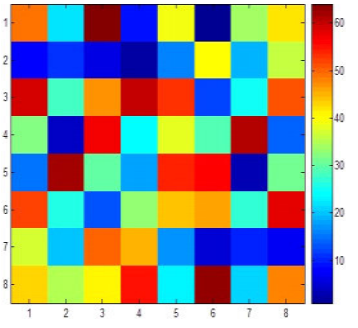
①x0=0.4，r=3.8，m=0



①x0=0.5，r=3.8，m=0



③x0=0.4，r=3.9 m=0

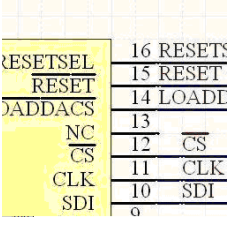


④x0=0.5，r=3.8，m=100

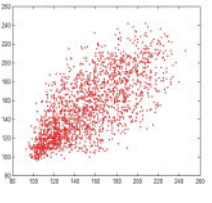
Fig. 2. Sensitivity analysis

4.2 Relevance Analysis of the Algorithm

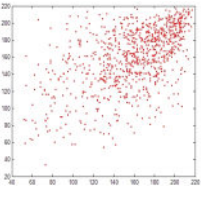
A good integrated circuit diagram encryption algorithm should destroy the correlation between the pixels, and make it as small as possible. The correlation integrated circuit diagrams are shown as follows:



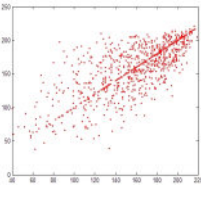
Original image A (256*256)



Level correlation of the pixels



Diagonal correlation of the pixels



Vertical correlation of the pixels

Fig. 3. Relevance analysis of of the pixels in original integrated circuit diagram A

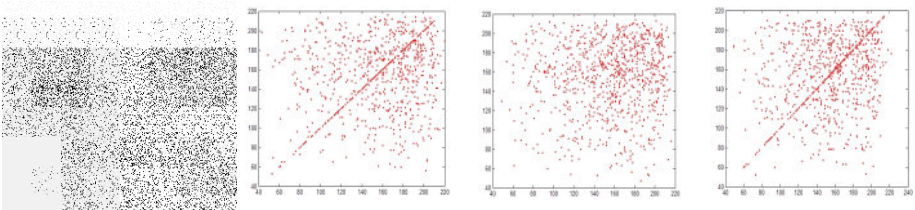


Fig. 4. Relevance analysis of the pixels in encrypted integrated circuit diagram A

The algorithm not only can be used in the encryption of integrated circuit diagram, it can be used for the encryption of ordinary images, the correlation of ordinary images are shown as follows:

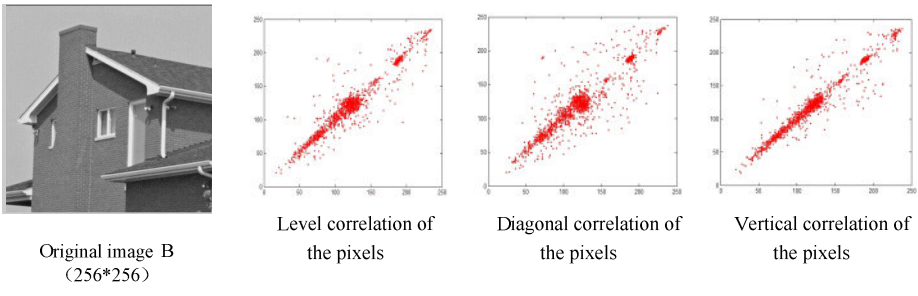


Fig. 5. Relevance analysis of of the pixels in original integrated circuit diagram B

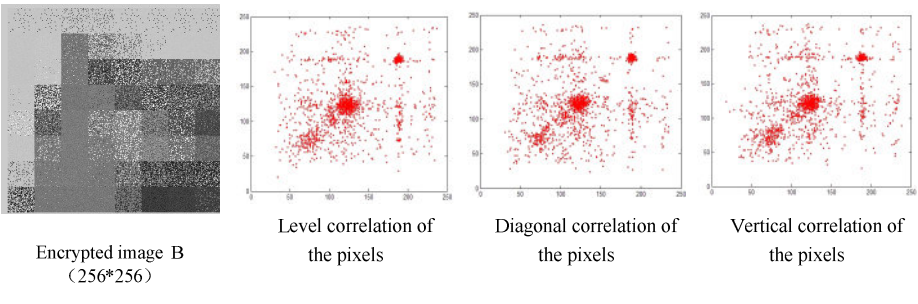


Fig. 6. Relevance analysis of the pixels in encrypted integrated circuit diagram B

It can be seen from the results, there is a high correlation between the adjacent pixels in an original integrated circuit diagram, while there is zero correlation between the adjacent pixels in a encrypted integrated circuit diagram. It shows that the encryption effective of this algorithm is good, and the algorithm has a good security.

5 Conclusion

In this paper, chaos theory is introduced into the integrated circuit diagram encryption algorithm, we propose an integrated circuit diagram encryption algorithm based on chaotic sequences, by using the characteristics of chaotic binary sequences and through the combination of the integrated circuit diagram characteristics. It achieves the encryption goal of "A password is used only once". The algorithm can hide the integrated circuit diagram information safely, it has high security. For the generated chaotic sequence, it had better not to choose the initial partial sequence, this will strengthen the effect of encryption.

References

1. Valle, M.E.: A class of sparsely connected autoassociative morphological memories for large color integrated circuit diagrams. IEEE Trans. Neural Networks 20(6), 1045–1050 (2009)
2. Lu, H.: Global exponential stability analysis of Cohen-Grossberg neural networks. IEEE Trans. Circuits Syst. 52(8), 476–479 (2005)
3. Huang, W.Z., Huang, Y.: Chaos of a new class of Hopfield neural networks. Appl. Math. C 206, 1–11 (2008)
4. Huang, Y., Yang, X.S.: Hyperchaos and bifurcation in a new class of four-dimensional Hopfield neural networks. Neurocomputing 69, 1787–1795 (2006)

A Fuzzy Expectation-Maximization Algorithm of Electronic Remote-Sensing Image

Wenying Ge[1,*] and Pengwei Li[2]

[1] School of Computer and Information Engineering
Anyang Normal University, Anyang, Henan, 455002
[2] Teaching Department of Computer
Anyang Normal University, Anyang, Henan, 455002
xunzhaoxinyi@126.com

Abstract. Among the statistical classification methods of Electronic Remote-Sensing images, the Expectation-Maximization(EM) algorithm is one of the most frequently used methods to estimate the parameters of the mixture models. But it is very sensitive to classification noises, which would probably slow down the running speed and reduce the accuracy of its results. In order to solve the problem, a fuzzy strategy of EM algorithm was proposed to inhibit the negative influence in which the posterior probability is weighted with membership so that the noise and normal samples can make different contributions in the iteration. The classification experiments on the Electronic Remote-Sensing images shows that the proposed algorithm can get better classification result.

Keywords: Expectation-Maximization, mixture model, Electronic Remote-Sensing images, fuzzy strategy, classification.

1 Introduction

The sufficient samples are essential to the correct classification in statistical classification methods [1]. Whereas, some little subjects in Electronic Remote-Sensing images are always mapped into very few pixels, which leads to the misclassifications and decreases the accuracy of classification. So, they are treated as the classification noises.

EM (Expectation-Maximization) algorithm is a general method to estimate parameters of the model employed in statistical classification. It's an iterative procedure starting with an initial assumption and involving two steps: Expectation and Maximization. E-step calculates the distribution expectation of complete samples with an initial assumption while M-step finds the assumption that make the expectation maximum. If the noises are involved in the iteration as the normal samples, it will inevitably bring the negative effects on the accuracy and speed of parameter estimation. To mitigate this defect, this paper introduced the fuzzy theory into EM algorithm. In our proposed method, training samples is represented by a

* Supported by NSFC No. 41001251.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 323–329.
springerlink.com

fuzzy set and the posterior probability is weighted with membership so that the samples make different contributions in the iteration by which depress the influence of the noises. Experiments show that this restriction on posterior probability can promote the performance of EM algorithm and accelerate its convergence.

2 Finite Mixture Model and EM Algorithm

2.1 Finite Mixture Model

The probability density of data x which obeys mixture model distribution in d dimensional space can be expressed as,

$$p(x \mid \theta) = \sum_{k=1}^{K} \pi_k \, p(x \mid w_k, \theta_k) \tag{1}$$

in which, π_k ($\pi_k \in (0,1)$, $\sum_{k=1}^{K} \pi_k = 1$) is the mixture parameter, $\theta = [\pi_1, \pi_2, \cdots, \pi_K; \theta_1, \theta_2, \cdots, \theta_K]^T$ denotes the parameter vector and $p(x \mid w_k, \theta_k)$ represents the density of components. Distribution of each component in a Guassian mixture model is a Guassian function with the component parameter $\theta_k = (m_k, \Sigma_k)$, mean vector m_k and the variance matrix Σ_k.

2.2 EM Algorithm

The incomplete data, unmarked samples, described as $X = \{x_1, x_2, \cdots, x_N\}$ is usually used to represent the observation data whereas the complete data is composed of the observation data and the vector set of hidden mark-value $z = \{z_1, z_2, \cdots, z_N\}$ [2], where z_i is the vector indicated the class of sample x_i. For the class k, $z_i = e_k$ in which e_k is a unit vector whose k th component is 1 and the others are 0.

The idea of EM algorithm is[2~6] to obtain the complete data set $\{x, z\}$ through estimating the hidden data z based on the initial parameter $\theta^{(0)}$ and current data x. And then, estimate the parameter θ by maximize the logarithm likelihood function $\log[P(x, z \mid \theta)]$. Instead of maximizing $\log[P(x, z \mid \theta)]$, EM algorithm employed its expectation $Q = E[\log P(x, z \mid \theta)]$ for that $\log[P(x, z \mid \theta)]$ is about θ and z simultaneously.

The steps of EM algorithm can described as,

(1) Initialization. Give the initial parameter $\theta^{(0)}$.
(2) E-step. Calculate the conditional expectation,

$$\begin{aligned} Q(\theta \mid \theta^{(t)}) &= E[\log P(x, z \mid \theta) \mid x; \theta^{(t)}] \\ &= \sum_{i=1}^{N} \sum_{k=1}^{K} [\log \pi_k p_k(x_i \mid \theta_k)] P(k \mid x_i; \theta^{(t)}) \end{aligned} \tag{2}$$

where, $P(k \mid x_i; \theta^{(t)})$ is the posterior probability expressed as,

$$P(k \mid x_i; \theta^{(t)}) = P(z_i = e_k \mid x_i; \theta^{(t)}) = \frac{\pi_k p_k(x_i \mid \theta_k^{(t)})}{\sum_{k=1}^{K} \pi_k p_k(x_i \mid \theta_k^{(t)})} \tag{3}$$

in which, K is the number of the components.

(3) M-step. Obtain the estimated value of θ through maximizing $Q(\theta \mid \theta^{(t)})$.

(4) $t = t + 1$. Repeat step (2), (3) till convergence.

It's proved theoretically that each iteration step of EM algorithm can guarantee the convergence of the object function on local optimum in the general condition[2~5].

3 Fuzzy Strategy of EM Algorithm and Parameter Estimate

3.1 Fuzzy Strategy of EM Algorithm

Seen from Eq.(2) and Eq.(3), there are no any restriction on the training samples in the calculation of object function $Q(\theta \mid \theta^{(t)})$. That means the contributions of the noise samples to object function are same as the normal ones. By analyzing the common of EM algorithm and fuzzy C-means clustering, DatRran and Michael Wagner[6] had pointed out that the EM algorithm is essentially a fuzzy clustering one in which the posterior probability can be look as membership. But, the exact fact is that this fuzzy cluster is not enough to restrain noise. For that, A noise sample should not belong to any component, but if the membership is defined as probability, with the constraint $\sum_{k=1}^{K} P(k \mid x_i) = 1$, there must be a component whose posterior probability is predominate so that the noise is considered as sample of this component. The misclassification will affect the estimation of this component. So, it's necessary to make further restriction on the fuzziness of EM algorithm.

Let the samples set as domain ($U = X$) and define K fuzzy sets in this domain correspond to the K classification respectively. For any $x_i \in U$, the membership of fuzzy set k, denoted as μ_{ik}, $\mu_{ik} \in [0,1]$, $k = 1, \cdots, K$, doesn't need the constraint $\sum_{k=1}^{K} \mu_{ik} = 1$. For a noise sample, its membership μ_{ik} is close to 0 because the probability of its belonging to a class is very small. So, restrain $P(k \mid x_i)$ with μ_{ik} can inhibit the influence of noises on parameter estimation. Based on the analysis, EM algorithm can modified as,

$$Q(\theta \mid \theta^{(t)}) = \sum_{k=1}^{K}\sum_{i=1}^{N}[\log \pi_k p_k(x_i \mid \theta_k)](\mu_{ik})^d P(k \mid x_i; \theta^{(t)}) \tag{4}$$

in which, d is the restraining parameter. With the constraint, the effect on object function from noises will be weakened so that parameters can be estimate more accurately and the convergence can be accelerated.

3.2 Parameter Estimation of Fuzzy EM Algorithm

Assuming that samples from any component obey the normal distribution, the membership of sample x_i to the k th class can be defined as,

$$\mu_{ik} = \exp(-(x_i - m_k)^T \Sigma_k (x_i - m_k)) \tag{5}$$

With this definition, the membership of samples around the mean value is close to 1 while the others will descent greatly, which is in accord with the distribution of membership. Assuming that each component is a Guassian function, the parameter vector will be $\theta = [\pi_1, \pi_2, \cdots, \pi_K; m_1, m_2, \cdots, m_K; \Sigma_1, \Sigma_2, \cdots, \Sigma_K]^T$. According the fuzzy EM strategy described above, by finding the extreme of object function with Lagrange, the iterative rules are obtained,

$$\pi_k^{(t+1)} = \frac{\sum_{i=1}^{N}(\mu_{ik}^{(t)})^d P(k \mid x_i; m_k^{(t)}, \Sigma_k^{(t)})}{\sum_{k=1}^{K}\sum_{i=1}^{N}(\mu_{ik}^{(t)})^d P(k \mid x_i; m_k^{(t)}, \Sigma_k^{(t)})} \tag{6}$$

$$m_k^{(t+1)} = \frac{\sum_{i=1}^{N} x_i (\mu_{ik}^{(t)})^d P(k \mid x_i; m_k^{(t)}, \Sigma_k^{(t)})}{\sum_{i=1}^{N}(\mu_{ik}^{(t)})^d P(k \mid x_i; m_k^{(t)}, \Sigma_k^{(t)})} \tag{7}$$

$$\Sigma_k^{(t+1)} = \frac{\sum_{i=1}^{N}(\mu_{ik}^{(t)})^d P(k \mid x_i; m_k^{(t)}, \Sigma_k^{(t)})(x_i - m_k^{(t)})(x_i - m_k^{(t)})^T}{\sum_{i=1}^{N}(\mu_{ik}^{(t)})^d P(k \mid x_i; m_k^{(t)}, \Sigma_k^{(t)})} \tag{8}$$

in which, t is the iteration number.

For the finite Guassian mixture model, the iteration step can be described as follows,

(1) Initialization: give the initial estimation of θ denoted as $\theta^{(0)}$, and set the restraining parameter d;

(2) E-step: calculate the posterior probability $P(k \mid x_i; m_k^{(t)}, \Sigma_k^{(t)})$ and membership $\mu_{ik}^{(t)}$ of each sample to each class based on Eq. (4) and Eq.(5);

(3) M-step: update the parameters according Eq.(6)~(8);

(4) $t = t + 1$. Repeat step (2) and (3) till convergence.

The fuzzy EM algorithm imposes the constraint on the membership using posterior probability and adjusts the constraint strength by d. In this method, the membership of a noise sample to any class is close to 0 and its effect on classification will be reduced or even eliminated in the EM iteration.

After the parameters estimation using this fuzzy strategy, the classification can performed according the follow equation,

$$k' = \arg\max_{k=1,2,\cdots,K} \{(\mu_i^k)^d P(k \mid x_i)\} \tag{9}$$

4 Experiments and Analysis

To verify the inhibitory effect of fuzzy EM algorithm on the noises, a 256×256 Electronic Remote-Sensing images with Lands at TM3,4,5 bands shown in Fig.1 is classified by both of our method and the standard EM algorithm. Brightness of three bands are extracted as the feature vector. The image is classified manually into water (river and paddy), bare land (road, bridge and unused area) and dry land. 512 samples are selected from each class respectively to calculate the initial parameters[7~8] and 3000 random unmarked pixels are employed as training set. In our method the restraining parameter $d = 1$. Each class is tested by 512 marked samples and the error matrix is showed in Tab.1. In the Fig.2, (a) is the classification using initial parameter, (b) is the result of standard EM algorithm and (c) shows the classification of the proposed fuzzy EM algorithm.

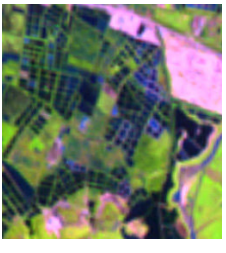


Fig. 1. TM3, 4, 5 image

Conclude from the Fig.2, the proposed fuzzy strategy improved the performance of EM algorithm especially in the marking area. Moreover, the iteration number of standard EM is 15 whereas only 3 iterations in fuzzy EM, which show that our proposed method can accelerate the convergence.

Seen from the error matrix shown in Tab. 1, 77(31+46) errors in water samples by standard EM algorithm vs. 35(26+9) errors by fuzzy EM, our method has better performance than the EM for the inhabitation on noises.

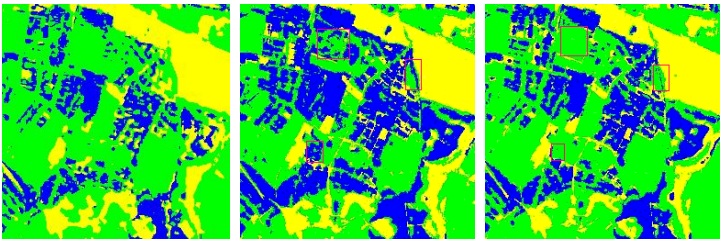


Fig. 2. Classification result (blue represent water, green is dry land and yellow denote bare land) (a) initial classification (b) EM algorithm (c) fuzzy EM

Table 1. Error matrix of standard EM and fuzzy EM

	Standard EM			Fuzzy EM		
	water	dry lands	bare lands	water	dry lands	bare lands
water	465	31	46	477	26	9
dry lands	0	512	0	0	512	0
bare lands	0	0	512	0	3	509

5 Conclusion

EM algorithm, a general method to calculate the maximum likelihood estimation from incomplete data set, does not distinguish the noises from normal samples, which brings the negative effect on classification and convergence speed. And the defect is more obvious in Electronic Remote-Sensing images for existence of subject noise besides imaging noise. In order to finite the influence of noise, fuzzy theory is introduced into the EM algorithm. In the fuzzy strategy, the training samples, expressed by a fuzzy set, have different membership to different class. And then the posterior probability is weighted with these memberships to inhibit the effect of noise in training. Experiments show that the constraint can accelerate convergence and improve the classification accuracy.

Fuzzy EM is yet sensitive to initial values as standard EM algorithm. The membership and posterior probability is calculated according to the former iteration results, which make the initial values are very essential. So, the initial value selection will be paid more attention in our future research.

References

[1] Hsieh, P., Landgrebe, D.A.: Classification of high dimensional data. School of Electrical and Computer Engineering, Purdue University, West Lafayette (1998)

[2] Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum-likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society, Series B (Methodological) 39(1), 1–38 (1977)

[3] Neal, R.M., Hinton, G.E.: A View of the EM Algorithm that Justifies Incremental, Sparse, and Other Variants, pp. 355–368. Kluwer Academic Publishers, Boston (1998)
[4] Redner, R.A., Walker, H.F.: Mixture Densities Maximum Likelihood and the EM Algorithm. SIAM Review 26(2), 195–239 (1984)
[5] Yu, L.-S., Zhang, T.-W.: Filtering EM Algorithm for Image Segmentation. Chinese Journal of Computers 29(6), 928–935 (2006)
[6] Tran, D., Wagner, M.: Fuzzy Expectation-Maximization Algorithm for Speech and Speaker Recognition. In: 18th International Conference of the North American Fuzzy Information Processing Society, NAFIPS, pp. 421–425 (1999)
[7] Yang, Z.-H., Zhou, Y.-Y., Zhu, Y., et al.: Improved EM algorithm for Hyperspectral classification. Journal of Zhengzhou Institute of Surveying and Mapping 23(1), 36–40 (2006)
[8] Luo, J.-C., Wang, Q.-M., et al.: The EM-based Maximum Likelihood Classifier for Remotely Sensed Data. Acta Geodaetica et Cartographica Sinica 31(3), 234–239 (2002)

Comparative Study of Water-Body Information Extraction Methods Based on Electronic Sensing Image

Xiumei Li, Xianbin Liu, Lina Liu, and Kun Xue

Tianjin Key Lab of Marine Resource & Chemistry, Tianjin University of Science & Technology, Tianjin 300457, China
lixiumei@tust.edu.cn

Abstract. In order to extract accurately wetland information of Tianjin Binhai New Area from electronic sensing images, the paper summarized and analyzed several water extraction methods, including Gray-level Threshold Method, Spectrum-photometric Method, Band Ratio Method and Modified Normalized Difference Water Index (MNDWI). Water information of the TM data was extracted using ENVI software by these methods, respectively. Finally, the above-mentioned results were analyzed and compared. The results show that Gray-level Threshold Method is the easiest and the least accurate, it is not appropriate for the extraction work of small water bodies. Spectrum-photometric Method is more accurate and gets clear water boundary. MNDWI is the most accurate, suitable for all kinds of water, but the boundary definition is not as good as the formers. Therefore, no one is accurate completely, but if we combine the method reasonably according to the characteristics of data, we can achieve better results.

Keywords: electronic sensing image, information extraction, water-body information, TM.

1 Introduction

In order to promote the construction of Tianjin Binhai New Area, Tianjin is actively building a good investment environment. As a coastal city, it is particularly important to deal the marine environment with real-time monitoring and governance. The development of remote sensing technology provides technical support for it. However, how to use the information obtained from electronic sensing images effectively is a long-term issue concerned by the people. For this reason, the author sums up the experience of predecessors, makes detailed comparative analysis of several ways that extract water-body information from electronic sensing image, trying to find one or more appropriate way to prepare for the survey of the coast of Tianjin and its surrounding environment.

With seven bands, TM data can Comprehensive reflect all kinds of objects Spectral characteristic, and it is one of the most widely used remote sensing data. Thus this study uses the Landsat TM data of Tianjin in April, 2007, the scope of the study area is Bohai Sea and its surrounding area.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 331–336.
springerlink.com

2 Water-Body Extraction Methods

2.1 Gray-Level Threshold Method (GTM)

Gray-level Threshold Method is a commonly used and also the simplest method of image segmentation. It is a method that divides the gray-scale into many intervals, and chooses the threshold to determine of the region or border of the image[8].

The method usually extracts water body using the single Infrared band, the fifth band, of TM data. This band lies at the absorption band of water body. Water body appears dark tones in the image, which has obvious demarcation with the land border. This study sets proper threshold to distinguish water body from other surface objects using this method, reaches the purpose of extracting water body. After verification, it is found that gray level 17 is the optimal threshold, while 0 is set as the minimum threshold. The result of the data is as follows.

In Fig.1, the white part is large-area water, and it is obvious that the gray parts are water bodies, such as reservoir, salt pan, paddy field and river. Different types of water bodies present different colors, this is because the fifth band of Landsat TM is near-infrared band, whose wavelength is 1.55-1.75 micron, and it is sensitive to reflect water content, the absorption and reflect characteristics of water with different material is slightly different. Due to the effect from other surface objects, which have the similar spectral characteristics with the near-infrared band, part of the extracted water body is not real water body. For example, in the mountain area, reflectivity of the shadow of the mountain in the near-infrared is particular low, resulting in dark tone on the image, and confusion of water body and shadow makes it difficult to extract water body. In the result figure of GTM, there are a lot of noise points in the large-area water.

2.2 Spectrum-Photometric Method (SPM)

Chenghu zhou[3] and other domestic scholars believe that the water has unique characteristic relationship of Spectrum-photometric in the TM image. By analyzing the characteristics of the spectral curve about water bodies and the background, we can identify the changes between them, and then extract water body using logical expression.

According to the spectral characteristic of water body[5], the sum of second band and third band can increase difference in gray value between water body and shadow, gray value of water body is larger than that of shadow. The forth band adding to the fifth band can also increase this difference, while gray value of shadow is larger. TM2-TM5 can distinguish water body and shadow from other surface objects, while using (TM2+ TM3)-(TM4+TM5)>0 means to enhance the difference between water body and shadow[5]. The image is shown in Fig.2.

In the picture, the darkest part is large-area water the lighter parts are water bodies, such as reservoir, salt pan, paddy field, rivers. In this method, the reason that different types of water bodies have different gray value is every water body has different gray value in TM2, TM3, TM4 and TM5. Also, there is different between TM2+ TM3 and TM4+TM5.

SPM uses more bands than GTM, so more factors are involved in operations, and details of the result have better quality than GTM. In theory, SPM has superiority over GTM. In particular, it can distinguish water body from shadow, and is suitable for the extraction of water in mountain[6]. Whether SPM or GTM, both of them have certain limitation on extracting some small rivers[9].

Fig. 1. GTM

Fig. 2. SPM

Fig. 3-A. BRM of TM2/TM5

Fig. 3-B. BRM of TM1/TM7

Fig. 4. MNDWI

2.3 Band Ratio Method (BRM)

BRM is a method that uses ratio calculation to extract water body information fast according to spectral characteristic in different bands of different water bodies. There is very clear boundaries between land and water in the TM3/TM2 image, the ratio of band TM3/TM2can be selected when separating water from land. In large areas of coastal zones, the ratio of TM2/TM5 of water is greater than 1, while the ratio of land less than 1, this difference can also be used to distinguish water from land[1].Meanwhile, the ratio of TM1/TM7 prominent body of water can also highlight water body information.

This paper selects TM2/TM5 model and TM1/TM7 model to extract water body. The result image of TM2/TM5 model was shown in 3-A, and the TM1/TM7 model was shown in 3-B. The effect of extraction using the ratio model of the two-band is very similar overall. In details the TM2/TM5 model is better, because the result's error derived from TM1/TM7 model is greater than TM2/TM5 model.

2.4 Normalized Difference Water Index (NDWI)

SPM is usually complex, and BRM can't completely inhibit unrelated background information to the water body. Hence, Mcfeeters[4] put forward with NDWI to solve this problem. In fact, NDWI just improves the above phenomena. In many cases, extracted water information with NDWI is still mixed with many non-water body information, especially in the extraction of water body within the scope of urban is not very successful. Therefore, Hanqiu Xu[2] proposed modified Normalized Difference Water Index(MNDWI), which can extract urban water body information better, after Carefully analyzed the urban water body and the reflection characteristics of its background. Following equation is MNDWI considering the urban factors:

$$MNDWI = (Green - MIR) \Big/ (Green + MIR)$$

Where: MIR is Mid-infrared band, such as the fifth band of TM or ETM.

Water body displays high-brightness in the MNDWI image, corresponding with high MNDWI value. However, non-water part, such as woodland, farmland and urban and rural construction land, show dark and have low MNDWI value, with great difference from water body. The result image of MNDWI is Fig 4, water body has high MNDWI value and displays high-brightness, non-water part shows low-brightness with low MNDWI value. The brightest part of the picture is a large area of water body, according to the different MNDWI value, followed by reservoir, salt pan, Paddy field, river and other water body.

3 Comparative Analysis of Features and Precision

3.1 Precision Analysis

First, we compare whether the sampling points in the result image is in line with the TM432 and show results in the list, and then evaluating 15 sampling points with typical characteristic from each image. In the evaluation, inexact is given -1, exact is given 1, and indefinite is given 0. The sum of all values dividing 15 is used to estimate the accuracy.

From Table 1, we can see directly that the accuracy of various methods extracting water body in TM data with descending order is as follows: MNDWI, SPM, BRM and GTM. The 15 sampling points of GTM have one indefinite point and three inexact points, showing that GTM has large inaccuracy. In the two programs of BRM, the 15 sampling points of TM2/TM5 have five indefinite points and no inaccurate point, with high accuracy, and this program has less great error. Though TM1/TM7 has the same accuracy as TM2/TM5, this program has only one indefinite point and two inaccurate points. So this program has the possibility of more errors than TM2/TM5. It can be said that it is not safe enough than TM2/TM5 program. The accuracy of SPM is higher than these two methods above, all of the 15 sampling points, there are not indefinite point and 2 inaccurate points. Though accuracy of SPM was improved, it has the same inaccurate ratio as TM1/TM7 of BRM, which is they have the same possibility of big error. MNDWI has the highest accuracy of the four methods, only one inaccurate point.

3.2 Analysis of Extraction Effect

15 sampling points of every method corresponds the same water body or its boundary in every image, now we consider the situation that the same position in different methods. Table 2 compares the effect of different water body extraction. Table 3 compares the effect of different water body boundary extraction.

By the tables 2 and 3 the following conclusions can be summarized: first, the better situations in every method are port, bayou, pier and beach, and the building is difficult to judge. Second, compare the results of different method, SPM has the best effect, MNDWI is the second better, and GTM can make up the shortage of MNDWI. Third, GTM in the extraction of small water bodies has considerable shortcomings, it was not appropriate for the extraction work of small water bodies separately, and there are many noise points.

Table 1. Comparison of accuracy table

	methods and evaluation				
Method / NO	GTM	SPM	BRM(TM2/TM5)	BRM(TM1/TM7)	MNDWI
1	inexact	exact	indefinite	exact	exact
2	indefinite	exact	exact	exact	exact
3	exact	inexact	exact	indefinite	exact
4	exact	exact	exact	exact	exact
5	exact	exact	exact	exact	exact
6	exact	exact	exact	exact	exact
7	exact	exact	indefinite	exact	exact
8	exact	inexact	indefinite	inexact	inexact
9	exact	exact	exact	exact	exact
10	inexact	exact	indefinite	exact	exact
11	exact	exact	exact	exact	exact
12	exact	exact	exact	inexact	exact
13	exact	exact	exact	exact	exact
14	inexact	exact	exact	exact	exact
15	exact	exact	indefinite	exact	exact
accuracy	53%	73%	67%	67%	87%

Table 2. Comparison of different water body extraction effect

Method / position	GTM	SPM	BRM (TM2/TM5)	BRM (TM1/TM7)	MNDWI
paddy field	Not extracted	Clearly defined	Little dark	Little dark	Clearly defined
branch river	Clearly defined	Clearly defined	Clearly defined	Clear but Little dark	Clearly defined
river	Clear but Little dark	Clearly defined	The river little dark	Clear but Little dark	Clearly defined
artificial drainage	Very dark	Clearly defined	Clear but Little dark	Clear but Little dark	Clearly defined
reservoir	Clearly defined	Clearly defined	Clear but Little dark	Clearly defined	Clearly defined
small water bodies	Not extracted	Clearly defined	Clear but Little dark	Clear but Little dark	Clearly defined
salt pan	Clearly defined	Clearly defined	Clear but Little dark	Clear but Little dark	Clearly defined

Table 3. Comparison of different water body boundary extraction effect

Method / position	GTM	SPM	BRM (TM2/TM5)	BRM (TM1/TM7)	MNDWI
coastline	Clearly defined	Clearly defined	unclear boundary	unclear boundary	unclear boundary
Port	Clearly defined	Clearly defined	Clearly defined	Clearly defined	Clearly defined
bayou	Clearly defined	Clearly defined	Clearly defined	Clearly defined	Clearly defined
building	Correct	indefinite	indefinite	indefinite	indefinite
pier	Clearly defined	Clearly defined	Clearly defined	Clearly defined	Clearly defined
beach	Clearly defined	Clearly defined	Clearly defined	Clearly defined	Clearly defined

4 Concluding

After the comparative analysis, we can see there were many methods to extract water body with TM data, and of which SPM has the best comprehensive effect and MNDWI has the most accurate. Meanwhile, the results also show that no one is accurate completely, but if we combine the method reasonably according to the characteristics of data, we can achieve better results.

Acknowledgements. This project was funded by Tianjin Science and Technology Support Key Program (No.07ZCKFSF02000) and Tianjin Tanggu Science and Technology Support Program (No. 2010XQ15-01). We thank the reviewers for their critical comments.

References

1. Alesheikh, Ghorbanali, Nouri: Coastline change detection using remote sensing. Int. J. Environ. Sci. Tech. 4(1), 61–66 (2007)
2. Xu, H.-Q.: A Study on Information Extraction of Water Body with the Modified Normalized Difference Water Index (MNDWI). Journal of Remote Sensing 9(5), 589–596 (2005)
3. Zhou, C.-H., Luo, J.-C., Yang, X.-M., et al.: Geoscientific Interpretation and Analysis of electronic sensing images. The Science Press, Beijing (1999)
4. McFeeters, S.K.: The Use of Normalized Difference Water Index (NDWI) in the Delineation of Open Water Features. International Journal of Remote Sensing 17(7), 1425–1432 (1996)
5. Wang, J.-H., Zhang, Y.-B.: Application of Spectro-photometric Method in Extracting the Water Characteristics. Mine Surveying 4, 30–32 (2004)
6. Zhang, M.-H.: Extracting Water-Body Information with Improved Model of Spectal Relationship in a Higher Mountain Area. Geography and Geo - Information Science 24(2), 14–17 (2008)
7. Zhao, Y.-S., et al.: Remote Sensing Application Analytical Principles and Methods. Science Press, Beijing (2003)
8. Zhang, X.-C., Huang, Z.-C., Zhao, Z.-H.: Digital Remote Sensing Data Processing. Press of Zhejiang University, Hangzhou (1997)
9. Li, X.-M., Wang, G., Tian, J.: Study of The Method of Picking-Up Small Water- Bodies in Landsat TM electronic sensingimage. Journal of Southwest Agricultural University (Natural Science) 28(4), 580–582 (2006)

Design of the Signal Generator Based on C8051F410

Yuwen Zhai[1], Wei Liu[2], and Xiao Yang[1]

[1] College of Mechanical & Electrical Engineering, Jiaxing University, Jiaxing 314001, Zhejiang, China
[2] College of Information & Control Engineering, Jilin Institute of Chemical Technology, Jilin City 132022, China

Abstract. The design and realization of the signal generator based on C8051F410 is introduced. The signal generator can output sine wave, rectangular wave and triangular wave and can provide adjustment for signal parameters, such as frequency, amplitude and duty ratio. It has been applied in the Comprehensive Experimental System for Electronic Measurement Principle.

Keywords: C8051F410, Control, LCD, Oscillator, Signal generator.

1 Introduction

Function signal generator is widely used in circuit experiment and equipment test. This paper proposes a design and implementation of a signal generator that uses C8051F410 as its major module. It is one of the units of Comprehensive Experimental System for Electronic Measurement Principle, and has simple circuit and low-cost hardware. To meet the demand of experimental teaching, it can output sine wave, rectangular wave and triangular wave and can provide adjustment for signal waveform parameters, such as frequency, amplitude and duty ratio.

2 Interface Circuit for C8051F410 and LCD

The low power consumption SOC C8051F410 is used as the main processor, in which 24 channels of 12bit 200ksps ADC and 2 channels of 12bit DAC current output are available[1,2]. It can run in single cycle with 50MHz so as to meet the requirements of high speed data processing and wave display.

8 bits parallel interface is used to connect the main processor C8051F410 and LCD. The processor C8051F410 needs many I/O ports to implement the functions of LCD display, keyboard and wave measurement, so 74LS245 is used to perform the multiplexing for Port P2 which is connected with A1 to A8 of 74LS245. The RT128128 LCD interface and Port 2 extension circuit are shown in Fig.1.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 337–341.
springerlink.com

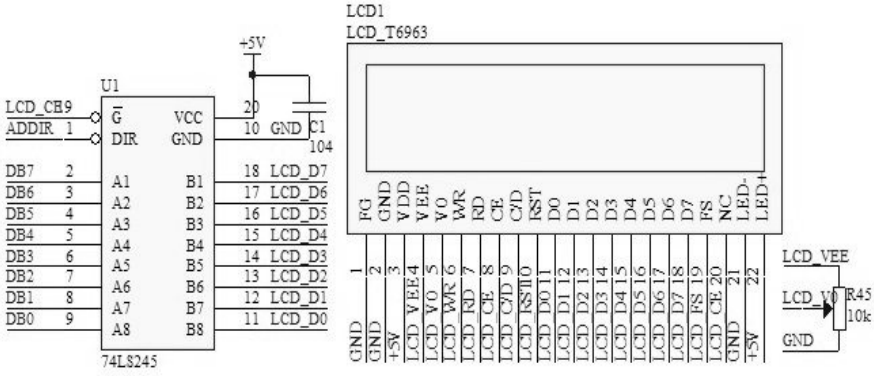


Fig. 1. LCD Interface Circuit

3 Signal Generation Circuit

3.1 ICL8038 Signal Generator

The signal generation circuit outputs sine wave, rectangular wave and triangular wave with peak value of 0 to 10V and frequency range of 10Hz to 100KHz. The duty ratio of waves can also be adjusted. The integrated chip ICL 8038 is used to design the signal generation circuit [3,4]. The frequency is controlled by adjusting the voltage of PIN 8 of ICL8038, which is from DAC0 of C8051F410. The duty ratio is controlled by adjusting the variable resistor W500. The ICL8038 signal generator circuit is shown in Fig.2.

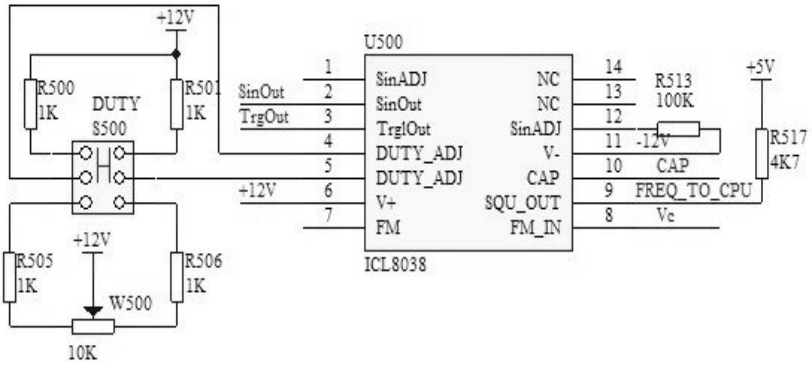


Fig. 2. ICL8038 Signal Generator Circuit

3.2 CD4051 Switching Circuit for ICL8038 Charge/Discharge Capacitor

The signal oscillation is realized by two inner Constant Current Sources in ICL8038 charging and discharging the capacitors. The frequency ratio of oscillation on the same capacitor is usually within 10 to 100. If the ratio is over 100, the ICL 8038

oscillator may stop run. In order to obtain the frequency output range of 10Hz to 100KHz, that is the frequency ratio of 10000, the analog switch CD4051 is used to perform the charge/discharge switching for capacitors C. The CD4051 charge/ discharge switching is controlled by C8051F410, shown in Fig.3. The switching circuit for charge/discharge capacitors C is shown in Fig.4.

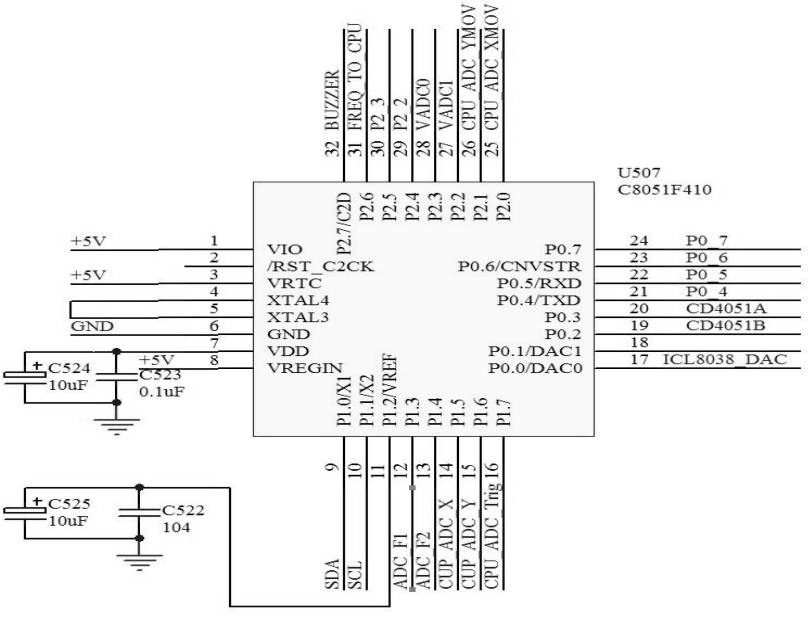


Fig. 3. CD4051 Multi-channel Switching Circuit Controlled by C8051F410

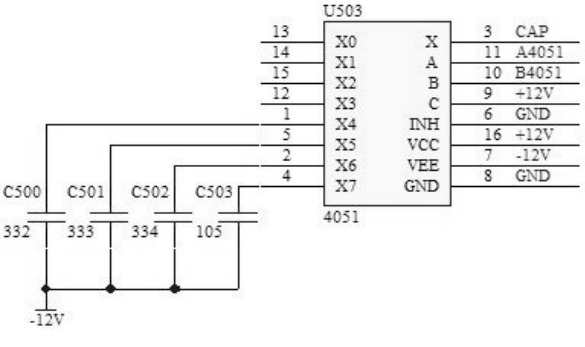


Fig. 4. CD4051 Switching Circuit for ICL8038 Charge/discharge Capacitors

3.3 Transducer Circuit of DAC Output Voltage of C8051F410

The DAC output voltage of C8051F410 is also used to adjust output frequency [5], which is within the range of 0 to 2.2V, which can not directly control the PIN 8 of

ICL8038. So a voltage transducer circuit, shown in Fig. 5, is designed between DAC of C8051F410 and PIN 8 of ICL8038. The operational amplifier LM324 performs the voltage conversion, in which U504B provides a fixed bias voltage, and U504A, a proportional operation adder, combines the bias voltage with DAC output voltage then amplifies it 5.7 times to control the frequency of ICL8038.

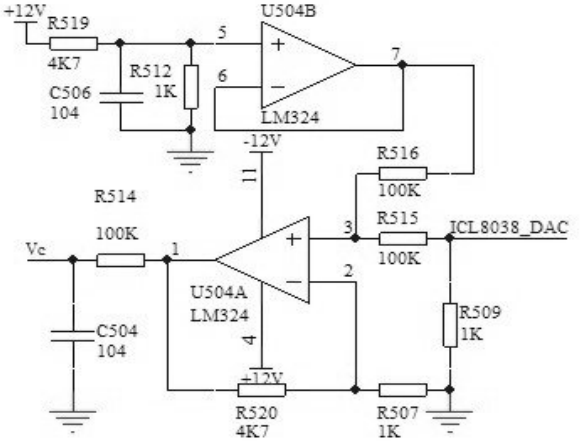


Fig. 5. Voltage transducer for DAC output voltage of C8051F410

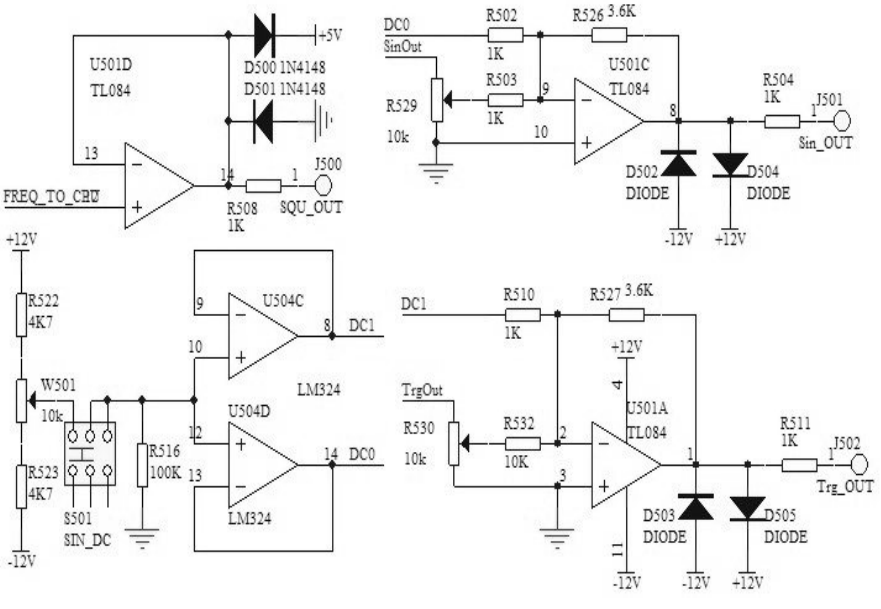


Fig. 6. Output Circuit of the Signal Generator

3.4 Output Circuit of the Signal Generator

An operational amplifier is designed to connect the signal output end of ICL8038 in order to ensure the output signal voltage range of 0 to 10V with certain ability of output driving and overload protection. The U504C and U504D are voltage followers which can change the DC voltage of the output triangular wave or sine wave by adjusting W501. The output circuit of the signal generator is shown in Fig.6 [6].

4 Conclusion

The signal generator designed in this paper can output sine wave, rectangular wave and triangular wave and can provide adjustment for the wave parameters, such as frequency, amplitude and duty ratio. The adjustable frequency range is within 10Hz to 100KHz; signal amplitude range is within 0V to 10V; duty ratio is within 15% to 85%. This signal generator has been applied in the Comprehensive Experimental System for Electronic Measurement Principle.

References

1. Silicon Laboratories (Pan, Z. (trans.)): C8051F020/1/2/3 Mixed-Signal ISP FLASH MCU Family, pp. 226–236 (February 2006)
2. Li, G., Lin, L.: 8051-Compatible High-performance High-speed MCU—C8051Fxxx. Beihang University Press, Beijing (2002)
3. Bai, B., Wu, M.: Design of Signal Generator Based on Integrated Function Generator ICL8038. Practical Electronics (03), 12–13 (2008)
4. Zhou, Y.: ICL8038 Sweep Frequency Signal Generator. Modern Electronic Technique (17), 37–38 (2003)
5. Ai, X., Ji, H., Zhang, Y.: Design of Inverter Power Based on System-on-chip MCU of C8051F410. Control and Instruments in Chemical Industry 37(8), 63–65 (2010)
6. Zhai, Y., Liang, W., Ai, X., Shi, Y.: Electronic Design and Practice. China Electric Power Press, Beijing (2005)

Comparison and Analysis of Copy-Move Forgery Detection Algorithms for Electronic Image Processing

YunZhong Sheng[1], HangJun Wang[1,2], and GuangQun Zhang[1,*]

[1] School of Information Engineering, Zhejiang A&F University, Linan 311300, China
[2] Department of Automation, University of Science and Technology of China, Hefei 230027, China
413604352@qq.com, {whj,gloria}@zafu.edn.cn

Abstract. Copy-Move forgery is one of the most common image manipulations, where a part of the image itself is copied and pasted into another part of the same image. In this paper, we firstly review several methods proposed during recent ten years, which are classified into block-matching and sift-matching. Then, we demonstrate the key techniques of two methods and analysis several different block-matching features in relation to their time complexity and robustness to common processing scaling up/down, compression, and rotation. Finally, some remaining challenges were discussed, and some future research directions were pointed out.

Keywords: Copy-Move forgery, tamper detection, block-matching, sift-matching.

1 Introduction

With the widespread use of digital cameras, powerful personal computers, and sophisticated image processing and editing software such as Photoshop, the manipulation of digital images is becoming easier. Moreover, maliciously manipulation and tamper digital images can appear real without leaving any obvious clues. Although most of people tampering digital images is only to enhance the visual effects, some people present tampered images as evidences in a court of law, as politics, as the media. Recently, there have been numerous examples of tampered images in newspapers and on magazine.

Therefore, developing techniques to verify the integrity and the authenticity of the digital images became very important. In this sense, image tamper detection is one of the primary goals in image forensics. This paper main interest is copy-move forgery which is one of the most common image manipulations .The paper is organized as follows: section 2 defines copy-move forgery detection; section 3 performs a comparison of block-matching and sift-matching techniques; section 4 discusses some remaining challenges and points some future research direction.

* Corresponding author.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 343–348.
springerlink.com

2 Copy-Move Forgery Detection

Copy-Move is a specific type of image manipulation, where a part of the image itself is copied and pasted into another part of the same image. The goal of such a tampering operation is to either cover an unwanted scene in the image, or to emphasize particular image content. In Figure 1, an example of copy-move forgery can be seen, where the left Pigeon is duplicated by cloning the right pigeon.

(a) (b)

Fig. 2. An example of copy-move forgery: (a) The original image with one Pigeon (b) The forged image with two Pigeons

Therefore, the goal of copy-move forgery detection technique is detecting the duplicated image regions. However, duplicated regions are not always exactly the same, because they may be rotated and scaled for a better effect.

Copy-Move forgery detection is defined as:

$$g(x, y) = T[f(x, y)] + n(x, y). \tag{1}$$

Where, $f(x, y)$ denotes the original region, and T denotes geometric transformations including translation, rotation, or scaling. $n(x, y)$ denotes such operations including lossy compression, blurring , or noise addition.

Hence, a good copy-move forgery detection technique should be robust to all such operations. To accomplish the task, several copy-move forgery detection techniques have been proposed.

3 Comparison and Analysis of Copy-Move forgery Detection Techniques

The simplest way to detect a Copy-Move forgery is to use an exhaustive search. But its drawback is computational expensive. Another technique is based on autocorrelation. However, it requires the duplicated region to be impractically large to perform reliably. In this paper, we focus on discussing two categories techniques: block-matching and sift-matching CMFD (copy-move forgery detection).

3.1 Block-Matching Technique

The known methods for block-matching CMFD follow a similar structure. Figure 3 shows an overview of the algorithm pipeline containing the main processing steps. After a preprocessing step, the image is divided into overlapping blocks of equal size and the feature vector is extracted from each block and inserted into a matrix, Then, these vectors are lexicographically sorted. Vectors corresponding to blocks of similar content would be close to each other in the list, so that identical regions could be easily detected.

We can see from figure 3, the selection of a suitable feature set is the core of copy-move forgery detection methods. Several authors presented to use different features to represent the image block. These features and its robustness will be discussed in 3.1.1. Another challenge is detecting the duplicated block pairs. The matching itself is most often done with so-called lexicographic sorting. Different matching techniques are described and evaluated in 3.1.2.

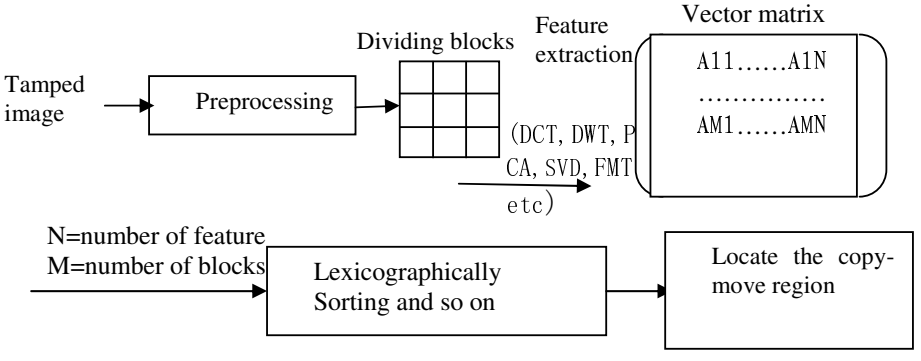


Fig. 3. Pipeline of copy-move forgery detection algorithms

3.1.1 Feature Extraction

The main difference between the various methods lies in feature extraction. On one hand a good feature can reduces computational complexity and on the other hand can be robust against some types of manipulations including Gaussian noise, lossy compression, and rotation or scaling. We considered 11 different features, which can be divided in four grouped: monment-based, dimensionality reduction-based, color-based, and frequency domain-based features (see Table 1).

All the above copy-move methods are most effective for forgery detection and robust to lossy JPEG compression, blurring, or noise addition, but very few of them are effectively robust against geometrical attacks. FMT, color2 methods can detect slight rotation and scaling. Color 3 performed well when the degree of rotation was 90,180 and 270 degree, but doesn't deal with Rotation arbitrary angles.

Table 1. Grouping of existing copy-move forgery detection methods

features	methods
Moments	Monment1: Mahdian using blur moments [6]
	Monment2: Junwen wang proposed Hu matrix [7]
	Monment3: Junwen wang proposed DWT and Zernike [8]
dimensionality	Popescu appling a principal component analysis(PCA) [9]
reduction-based	Li Guohui using DWT&SVD [10]
color-based	Color1:Luo using color information of the blocks.[11]
	Color2: Bravo-Solorio S using entropy [12 13]
	Color3: Lin Hwei-Jen using the average intensity of the block [14]
frequency domain-	Fridrich using DCT [15]
based features	Bayram using FMT [16]

3.1.2 Matching Techniques

After finding features, to detecting duplicated regions, the blocks that yield the same similar vectors have been determined. Known methods for matching are lexicographic sorting [7, 8, 9, 10, 11, 15].The biggest challenge in this step is the computation time. For lexicographically sorting, this time depends on the number of blocks, e.g. number of rows in the matrix mostly, while the number of features is also a factor for the speed. Table 2 displays the comparative results of five algorithms. For a grayscale image, sized 512×512, block sized 8×8. DCT, PCA and COLOR1 have (512-8+1)× (512-8+1)=255025 blocks, but DWT&SVD、DWT only have (512/2-8+1)×(512/2-8 +1)=62001 blocks. Because the number of blocks reduced, the number of features also reduced from 64 to 3 dimensions. This time of lexicographically sorting is reduced.

Table 2. Comparison of time Complexity

*An image of size 512*512*	*Fridrich*	*Popescu*	*Guohui Li*	*Weiqi Luo*	*Jun-wen wang*
feature	DCT	PCA	DWT&SVD	Color1	DWT low-frequency
The number of block size 8*8	255025	255025	62001	255025	62001
Feature	64	32	8	7	3

Although the time complexity of most of existing block-matching methods have been reduced, but time-consuming with a quantity of forgery databases. Also, it do not deal with all those manipulations including lossy compression, Gaussian noise, and rotation or scaling. Next, we will describe a sift-matching technique.

3.2 Sift-Matching Technology

Sift-matching technology is proposed by Huang [17]. He first extracted SIFT descriptors of an image which are invariant to changes in illumination, rotation,

scaling and are relatively low computational cost. These points (usually called keypoints) are extracted and tampered area can be determined using the set of extracted matched points. Amerini et al. [18] proposed Geometric Tampering Estimation by Means of a SIFT-based Forensic Analysis.

But similar to [17], the algorithm couldn't locate forgery region, only is to present matching points. Pan [19] present a new region duplication detection method based on the image SIFT features. This method can detect transform with continuous rotation angles.

However, SIFT can not locate keypoints on smooth background. That means it will fail to detect object concealing forgery, as shown in figure 4.

(a) tampered image (b) original image (c) key points extraction

Fig. 4. Key points extraction

4 Conclusion

As Copy-Move forgeries have become popular, the importance of forgery detection is much increased. Although many forgery detection techniques have been proposed and have shown significant promise, there are two major challenges: robustness against geometric transforms including scaling, rotation, and time complexity, especially to a quality of forgery images. In the paper, we review these algorithms and discuss its robustness and time complexity. Block-matching techniques are most effective for forgery detection and robust to JPEG lossy compression, blurring, or noise addition, but very few of them are effectively robust against geometrical attacks (rotation, scaling, distortion) and time-consuming. Sift-matching techniques still have a limitation on detection performance since it is only possible to extract the keypoints from peculiar points of the image. Therefore, how to merge two techniques is future research direction.

Acknowledgment. This work was supported by Educational Commission of Zhejiang Province of China, No. Y201121245.

References

1. Linxue winning photos are accused of fraud in CCTV 2006 (EB/OL) (2006), http://news.sohu.com/s2008/liulingyang/ (February 16, 2008)
2. Renwei tan. Pigeon tampered (EB/OL), http://epaper.nddaily.com/A/html/2008-04/04/content_431860.htm (April 4, 2008/August 8, 2008)

3. Nandu picture. in Nanning appeared People's Daily in 2009 (EB/OL) (2009), http://g1.nddaily.com/index.php?m=Photo&a=preview&uid=3254&cid=9061 (July 2, 2009/August 8, 2009)
4. Christlein, V., Riess, C., Angelopoulou, E.: A study on features for the detection of copy-move forgeries. Information Security Solutions Europe (2010)
5. Christlein, V., Riess, C., Angelopoulou, E.: On rotation invariance in copy-move forgery detection. In: IEEE Proceedings of the 2010 Second IEEE Workshop on Information Forensics and Security, Seattle, USA (2010)
6. Mahdian, B., Saic, S.: Detection of copy-move forgery using a method based on blur moment invariants. Forensic Science International 171(2), 180–189 (2007)
7. Wang, J.-W., Liu, G.-J., Zhang, Z.: Fast and Robust Forensics for Image Region-duplication Forgery. Acta Automatica Sinica 35(12), 1488–1495 (2009)
8. Wang, J.-W., Liu, G.-J., Zhang, Z.: Robust forensics for image regional duplication and forgery based on DWT and Zernike moment. Optics and Precision Engineering 17(7), 1687–1693 (2009)
9. Popescu, A., Farid, H.: Exposing Digital Forgeries by Detecting Duplicated Image Regions. Department of Computer Science, Dartmouth (2004)
10. Li, G.H., Wu, Q., Tu, D., et al.: A sorted neighborhood approach for detecting duplicated regions in image forgeries based on DWT and SVD. In: Proceedings of 2007 IEEE International Conference on Multimedia and Expo, Beijing, China, pp. 1750–1753 (2007)
11. Luo, W.Q., Huang, J.W., Qiu, G.P.: Robust detection of region-duplication forgery in digital image. Chinese Journal of Computers 30(11), 1998–(2007)
12. Bravo-Solorio, S., Nandi, A.: Passive Forensic Method for Detecting Duplicated Regions Affected by Reflection, Rotation and Scaling. In: 17th European Signal Processing Conference (2009)
13. Bravo-Solorio, S., Nandi, A.K.: Automated detection and localisation of duplicated regions affected by reflection, rotation and scaling in image forensics. In: European Signal Processing Conference (2011)
14. Lin, H.-J., Wang, C.-W., Kao, Y.-T.: Fast Copy-Move Forgery Detection. WSEAS Transaction on Signal Processing 5(5), 188–197 (2009)
15. Fridrich, J., Soukal, D., Luk, J.: Detection of copy-move forgery in digital mages. In: Proc. Digital Forensic Research Workshop, Cleveland, OH (2003)
16. Bayram, S., Sencar, H.T, Memon, N.: An efficient and robust method for detecting copy-move forgery. In: Proc. ICASSP (2009)
17. Huang, H., Guo, W., Zhang, Y.: Detection of Copy-Move Forgery in Digital Images Using SIFT Algorithm. In: Proceedings of IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application, vol. 2, pp. 272–276 (2008)
18. Amerini, I., Ballan, L., Caldelli, R.: Geometric Tampering Estimation by Means of a SIFT-based Forensic Analysis. In: Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP, Dallas, USA (2010)
19. Pan, X., Lyu, S.: Detecting Image Region Duplication Using SIFT Features. In: International Conference on Acoustics, Speech,and Signal Processing, Dallas, TX (2010)

Image Processing Technology and Its Application on Precision Stage

Xuepeng Liu and Dongmei Zhao

Zhongshan Polytechnic, P.R. China
liuxuepeng1026@yahoo.com.cn, zdmeihn@163.com

Abstract. Data smoothing, edge algorithm and binary image filtering processing is discussed in this paper. Precision stage structure is described. Then enhanced Robinson algorithm inhibits the noise. The system utilizes morphological theory to filter binary edge image. The filter image demonstrates that the noise which existed in binary edge image can be eliminated to obtain the accurate position.

Keywords: Edge algorithm, binary image filtering, precision stage.

1 Introduction

Image processing technology is widely used in industry and engineering field. There are some achievements in image division, characteristic signature extraction, edge detect, texture recognition field [1-7]. This paper analyzes the characteristic signature of precision stage Orthogonal Diffraction Grating for extracting key information. Image processing is used to extract diffraction light contour. The algorithm includes data smoothing, edge algorithm and binary image filtering processing. Firstly the color image captured by CCD is grayed, and then the data noise is smoothed. The diffraction light edge is obtained by edge detect. Finally Morphological Image Edge Detection Method is used to get characteristic feature.

2 Precision Stage System

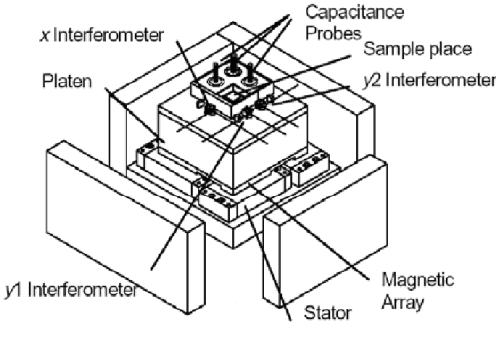


Fig. 1. Stage structure

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 349–352.
springerlink.com

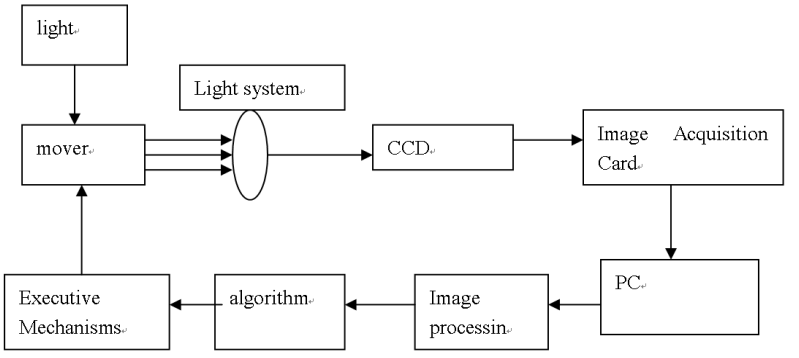


Fig. 2. System principle

3 Data Smoothing

The formula for graying is:

$$U = a_0 + a_1 \cdot t + a_2 \cdot t^2$$

According to least squares method, parameters of a0,a1,a2, is obtained by minimizing: $M = \sum_{k=-m}^{m} [u_i - (a_0 + a_1 \cdot t + a_2 \cdot t^2)]^2$. So the derivative is:

$$\begin{cases} \dfrac{\partial M}{\partial a} = 2\sum_{k=-m}^{m} (u_k - a_0 - a_1 t_k - a_2 t_k^2) = 0 \\ \dfrac{\partial M}{\partial a_1} = 2\sum_{k=-m}^{m} (u_k - a_0 - a_0 - a_1 t_k - a_2 t_k^2) t_k = 0 \\ \dfrac{\partial M}{\partial a_2} = 2\sum_{k=-m}^{m} (u_k - a_0 - a_1 t_k - a_2 t_k^2) t_k^2 = 0 \end{cases}$$

The results are:

$$A = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \left(\begin{pmatrix} 2m+1 & \sum_{k=-m}^{m} t_k & \sum_{k=-m}^{m} t_k^2 \\ \sum_{k=-m}^{m} t_k & \sum_{k=-m}^{m} t_k^2 & \sum_{k=-m}^{m} t_k^3 \\ \sum_{k=-m}^{m} t_k^2 & \sum_{k=-m}^{m} t_k^3 & \sum_{k=-m}^{m} t_k^4 \end{pmatrix}^T \begin{pmatrix} 2m+1 & \sum_{k=-m}^{m} t_k & \sum_{k=-m}^{m} t_k^2 \\ \sum_{k=-m}^{m} t_k & \sum_{k=-m}^{m} t_k^2 & \sum_{k=-m}^{m} t_k^3 \\ \sum_{k=-m}^{m} t_k^2 & \sum_{k=-m}^{m} t_k^3 & \sum_{k=-m}^{m} t_k^4 \end{pmatrix} \right)^{-1} \begin{pmatrix} 2m+1 & \sum_{k=-m}^{m} t_k & \sum_{k=-m}^{m} t_k^2 \\ \sum_{k=-m}^{m} t_k & \sum_{k=-m}^{m} t_k^2 & \sum_{k=-m}^{m} t_k^3 \\ \sum_{k=-m}^{m} t_k^2 & \sum_{k=-m}^{m} t_k^3 & \sum_{k=-m}^{m} t_k^4 \end{pmatrix}^T \begin{pmatrix} \sum_{k=-m}^{m} u_k \\ \sum_{k=-m}^{m} u_k t_k \\ \sum_{k=-m}^{m} u_k t_k^2 \end{pmatrix}$$

4 Edge Algorithm

Edge feature is an important characteristic. Most parts of information in image or signal lies in irregular structure and unstable situation. Edge detection technology enhances the contour processing algorithm to show the image. There are some difficulties in detection: one is the complexity of edge formation, the other is the choice between noise and edge.

There are some requests on edge detection algorithm:

[1], The edge detected by algorithm has much positioning precision.
[2], The response is single-pixel.
[3], The algorithm is suitable for different size edge and has no leakage.
[4], The algorithm is not sensitive to noise.
[5], The algorithm is not sensitive to edge direction.

Template matching method is used for edge detection algorithm to obtain satisfactory results. The following is the detailed Robinson algorithm:

The Robinson algorithm is made up of eight templates M. T is the A point's graying value. While J(i=0, 1, …, 7) is A point's graying value by (i+1) template processing. So

$$J_i = |M_i| \times |T|$$

The final value is:

$$J_A = \max(J_i)$$

It is concluded: when A is edge point, the graying values varies so much among the neighboring point of A, so J is the maximum value; when A is not edge point, the graying values varies so slightly among the neighboring point of A, J is nearly zero, which is much less than graying value at the edge point. So it is easy to find the image edge by Robinson algorithm and easy to mark the edge. When A is edge point, and T matches M_i, J is the maximum. So the value of point A is the maximum in the edge direction.

However, there are so many repetive calculations in Robinson algorithm, which adds up to processing time. Meanwhile, the edge direction is not needed when obtaining J. Therefore only four templates are needed for processing and to obtain absolute value. The enhanced algorithm is the following:

The one-dimension data group-D_{mxn} represents two-dimension image-MxN.

$$X_i = D_i + D_{i+1},\ Y_i = D_i + D_{i+n}, i = 0,1,2,\cdots m \times n$$

So

$$M_{oi} = X_{i-n-1} + Y_{i-n} - X_{i+n-1} - Y_{i+n}$$
$$M_{3i} = X_{i-n} + Y_{i-n+1} - X_{i+n-1} - Y_{i-1}$$

By comparing M0-M3, J is the maximum value. The test demonstrates that by the enhanced algorithm the calculation is 24% faster than that by original algorithm.

5 Binary Image Filtering Processing

The valve splitting processing is used to obtain clear edge. The valve method utilizes the difference between target and background to change the image into binary image, which is also called binary image splitting processing.

Opening binary makes image standing out of background smoother, while close binary makes background standing out of image smoother. So binary algorithm can eliminate the noise, which is the main idea of Binary morphological image edge detection method

Supposing that R is the original factor, while S is structure. F is the result of binary morphological filtering.

$$F = [(A \circ B) \bullet B + (A \bullet B) \circ B] / 2$$

The circle structure is uses to binary filtering with consideration to isotropy. The circle diameter depends on noise spatial frequency. The bigger is the partial frequency, the less is the diameter.

6 Conclusion

This paper utilizes Robinson edge detection algorithm to process mover image, which gets satisfactory contour. However, original Robinson algorithm is complicated. The enhanced Robinson algorithm inhibits the noise, which is suitable for the edge solution. In case of circumstance, mover situation and image capture device effect, there exists interference. The system utilizes morphological theory to filter binary edge image. The filter image demonstrates that the noise which existed in binary edge image can be eliminated to obtain the accurate position.

References

1. Wang, Z.: Study of Image Processing on the Automatic Vehicle Identification and Classification. China Agricuture University (2000)
2. Zhao, M., Na, L.: Study on Extraction Algorithms of the Stone Slab Surface Contour Based on Image Processing Technolog. Journal of Shenyang Jianzhu University (Natural Science) (2010)
3. Xu, D.: Machine vision-based medical heart image processing methodology. Liu Changhong, Computers and Applied Chemistry (2010)
4. Wang, W.-X., Yang, J.-M.: A geometric information on roughness of rock fractures based on image processing, Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition) (2010)

The Miniaturization of Active Vibration Control System Based on Digital Signal Processor

Ruilan Wang

School of Information and Control Enginerring, Weifang University
Weifang, Sandong, China
wrl2836@163.com

Abstract. Piezoelectric intelligent structure for the structure of the active vibration control is an important research direction. Currently, intelligent structure from the practical engineering, there is still a considerable distance, and truly has practical application value of piezoelectric smart structure hardware is also very small. Therefore developed a miniaturized, integrated characteristics to meet the multiple input and multiple output distribution control of piezoelectric smart structure active vibration control system is very meaningful. This paper introduce a kind of intelligent structure active vibration control system based on TMS320F2812, take the active vibration control experiment about flexible cantilever beam to confirm the system is feasibility.

Keywords: intelligent structure, DSP, active vibration control, auto-adapted control.

1 Introduction

With the development of space technology, intelligent structure vibration control systems towards miniaturization, intelligent direction, on the weight and size of the control system put forward higher requirements, which requires performance as high as possible, the volume as small as possible, the weight as light as possible. However, intelligent structure vibration system is composed of personal computers, data acquisition board, electric charge amplifiers and power amplifiers and other components. Therefore need to study measures for its small size, make its practical development, and to meet the needs of aerospace.

Piezoelectric material is used as sensor for piezoelectric smart structure, using it for active control of structural vibration and noise. Passive control system structure is simple, easy to implement, low cost, but the lack of flexibility of control, poor responsiveness of sudden environmental changes, tend to be applied to the structure of medium and high frequency vibration control. Compare with passive control, active vibration control rely on vibration signal as a feedback signal of structure and system, it can flexibility adapt to unforeseen external interference and the uncertainty of structure or system, with strong ability to adapt to circumstances and the ability to inhibition of

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 353–357.
springerlink.com

low frequency vibration and wide-band random vibration, therefore it has an important value and application.

2 The Composition and Working Principle of the Intelligent Structure Vibration Control System

Active vibration control system diagram shown in Fig. 1, which consists of control module, DC-DC converter module, and output vibration module.

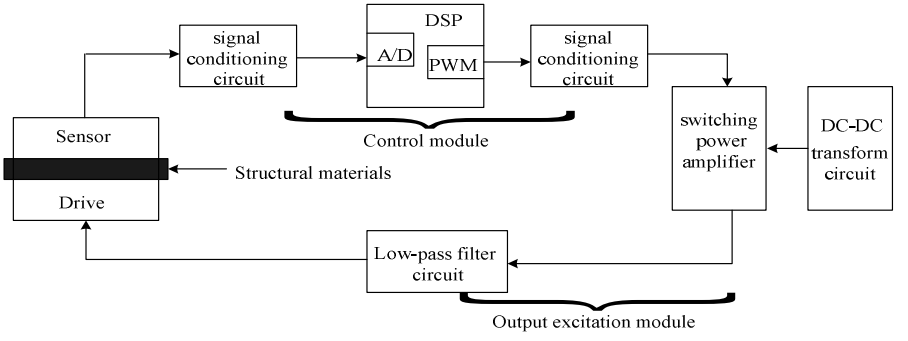


Fig. 1. Active vibration control system diagram

The operating principle of the system: the charge of sensor output is amplified to ± 5V voltage signal through charge amplifier. First this signal after conditioning as a control signal is input to the DSP control system, then is conditioning into a fixed frequency, pulse width adjustable square wave through A/D, pulse-width modulation (PWM) of EV manager of the system. By switching power amplifier, the square wave signal is amplified to square wave signal of bus voltage, after low-pass filter bring up 300V high-voltage for control signal that control of piezoelectric actuator.

3 The Hardware Design of the Intelligent Structure Vibration Control System

3.1 The Control System

The control system uses TMS320F2812 DSP as the processor, which is produced by Texas Company from the United States. It has both digital signal processing capabilities and powerful event management and embedded control. We use AD acquisition module, which is integrated on chip of TMS320F2812, to collect vibration amplitude signal of the object and reference signal (the required semaphore of adaptive algorithm adopted). After getting the result, we use PWM module to output.

3.2 DC-DC Converter Circuit

Nowadays, the requirements of miniaturization and being light of electronic devices are increasingly prominent. So the DC converter which provides DC voltage must be Miniaturized and light. And in terms of Low energy consumption, the efficiency of energy conversion must be improved. It is difficult to achieve for Conventional DC-DC converter. Especially since the existence of electromagnetic transformers, the weight and volume of DC-DC converter is even more difficult to reduce. So replacing the electromagnetic transformer with PT, the problem can be solved. This paper is the use of the advantages of the PT and design the DC-DC converter circuit. The PT model used in this article is MPT3610B10L1. It has maximum output power of 10W, the resonant frequency of 48kHz, the maximum output peak - peak voltage of 3460V and the output current of 6.5mA.

Fig. 2 shows the DC-DC converter circuit. 24VDC is the DC voltage source of, the DC-AC transforms the DC into high frequency AC excitation source to drive Piezoelectric ceramic booster PT (piezoelectric transformer).the output of PT becomes DC high voltage by Rectifier filter, which is the bus voltage of switch power amplifier.

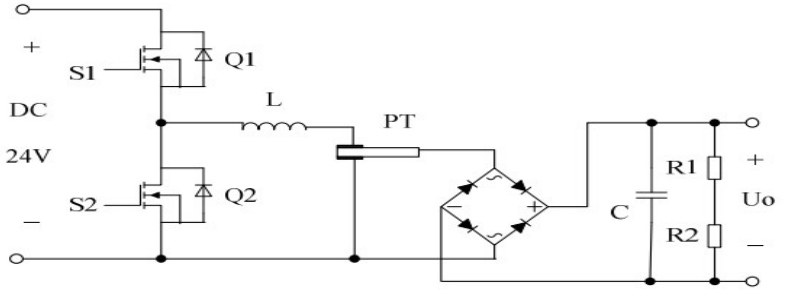


Fig. 2. DC-DC converter circuit

3.3 The Measuring Circuit of Piezoelectric Sensor

Sensors are important components of Smart structure which are in the front of the control system. First, Smart structures use sensors to measure internal and external parameters. Second, the internal and external parameters will be enlarged and sent to the processor. thirdly, it acts as the basis for the control system behavior. As the Piezoelectric sensor has high response frequency, high frequency bandwidth, good linearity, stable performance, high sensitivity, and good compatibility of composite materials, etc., so the piezoelectric sensor is a smart choice for structural vibration sensors. Piezoelectric sensor conditioning circuit is shown in Fig. 3. It has a pass band of about 0～8K and the output voltage ranges from -5V to +5 V.

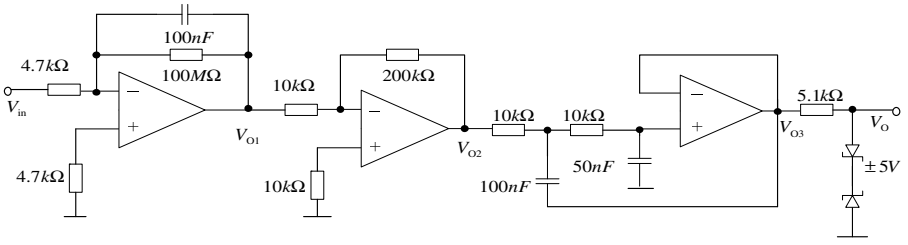


Fig. 3. The measuring circuit of piezoelectric sensor

3.4 The Drive Circuit of Piezoelectric Actuator

Switching power amplifier's main circuit diagram is shown in Fig. 4. The power switch Q3~Q6 are from IRFP460 produced by IR Company. The Electrical parameters of IRFP460 are: drain-source voltage VDSS=500V and maximum continuous on-state current ID=20A.

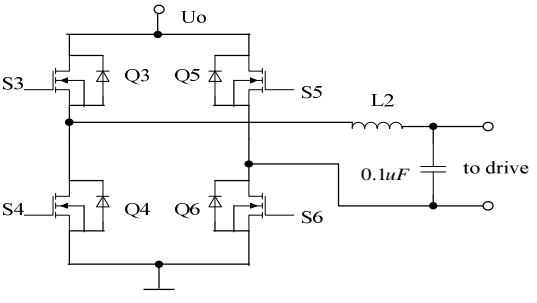


Fig. 4. The switching power amplifier's main circuit

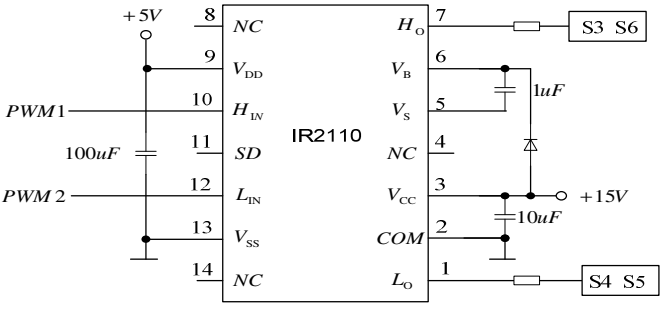


Fig. 5. The driver circuit based on IR2110

Driver circuit using IR's proprietary driver chip IR2110 is show in Fig. 5. IR2110 is a high-voltage, high power monolithic gate driver device driver. Logic input is compatible with standard CMOS or LSTTL integrated circuit. IR2110 not only meets the requirements of the power MOSFET driver circuit, but also greatly simplifies the hardware. The input control signals PWM1, PWM2 to IR2110 driver circuit are produced by the control system TMS320F2812.

4 Summary

We have conducted active vibration control for second-order bending mode of flexible cantilever plate using small active vibration control system. The experimental results show that flexible hanging arm plate vibration has been effectively suppressed through the small vibration active control system, and verify the feasibility of the control system.

References

1. Quan, W., Mao, J.: The use of smart structures and intelligent control In active vibration. Control Information and Electronic Engineering, 232–236 (2004)
2. Crawley, E.F.: A technology overview and assessment Intelligent structures for aerospace. AIAA Journal 32, 1689–1699 (1994)
3. Liu, Y.: The research of miniaturized and fast algorithm for Vibration control system. Nanjing Aeronautics and Astronautics University (2005)
4. Huang, Y., Zhou, K.: The key issues in the theory and its application of the piezoelectric transformer. Technical Acoustics, 58–64 (2003)
5. Qiu, Z.: Smart structures and active vibration control. Aerospace Control, 8–15 (2002)

Calibration of X-Ray Stereo Vision System Based on Electronic Image

Hongli Jin[1], Liqin Fu[1], and Fugen Su[2]

[1] Department of Mechanical & Electrical Engineering,
Beijing Institute of Economic Management, Beijing, 100102
[2] State Key Laboratory of Information Photonics and Optical Communications Beijing
University of Posts and Telecommunications 100876, Beijing, China

Abstract. This paper presents calibration method of X-ray system. The calibration method of exterior parameters which refer to geometric relation of imaging system is introduced for translational system and rotational system respectively. The method is implemented based on stereo image pair. Unlike it, the method of interior parameters is based on a non-linear transformation according to the calibration model obtained from the information of corresponding points on a chessboard plane. The paper also provides some experimental results which can verify the effectiveness of our method.

Keywords: System calibration, X-ray imaging, Stereo vision, Non-linear transformation.

1 Introduction

The NDT method based on radiography is used extensively and it inspects work piece automatically based on image processing techniques. The precision of inspection depends not only on image processing algorithm but also on the accuracy of system parameter. We all know that camera calibration is necessary if one is to obtain metric information from images [1,2]. For example, we need focal length and principal point of each camera in order to compute 3D coordinates from a pair of matched images [3]. Similarly, considerable efforts have been made to recover three-dimension information of object from X-ray images [4, 5]. Many stereo vision systems concern themselves only with calculating the distances to the matched objects. One must be aware that in any real application the geometric parameters of imaging system are not known with perfect accuracy. The geometrical relationship between work piece and its image can be established based on X-ray imaging model and parameter calibration.

Calibration of the system is necessary for two reasons. First, the exterior parameters of the X-ray imaging system must be determined because their precision influences the precision of depth measurement seriously. The exterior parameters refer to the geometric relation of the X-ray system which includes the distance between X-ray source and the detector, the distance between work piece and the

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 359–365.
springerlink.com

detector, and the relative location of the two X-ray source positions. The later includes the pointing angles of the two X-ray source positions, the rotations of two imaging sensors, and translational distance between two X-ray source positions as well as two imaging sensors. Second, the interior parameters calibration which influences image quality should also be implemented to solve the image distortion induced by uneven scan line-speed, delay of sensor sampling, malposition of scan row, etc.. This paper discussed the calibration method of exterior and interior parameters for X-ray stereo vision system. The exterior parameters are calibrated based on stereo image pair. Meanwhile, the method of interior parameters is based on a non-linear transformation according to the calibration model obtained from the information of corresponding points on a chessboard plane. The method has been used extensively in visible light vision field [6,7,8].

2 Exterior Parameter Calibration

2.1 Translational System

The translational system is shown in fig.1. S and S1 are the two positions of the radiant. D is the displacement of the radiant between double exposures. F is the distance between radiant and imaging plate. h is distance between defect and imaging plate. B is the displacement of defect projection. According to geometrical relation in fig.1, we can obtain

$$h = B \cdot F /(D + B) \tag{1}$$

Therefore, we can calculate the depth of a defect, h, so long as the values of F, D and B are measured. The precision of h depends on the accuracy degree of F, D and B.

On the other hand, F and D can be calibrated only if we obtain the precise B and h of several marks according to equation (1).

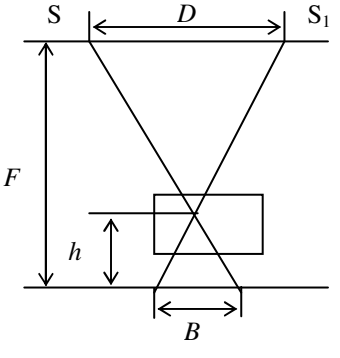


Fig. 1. Geometrical relation of translational system

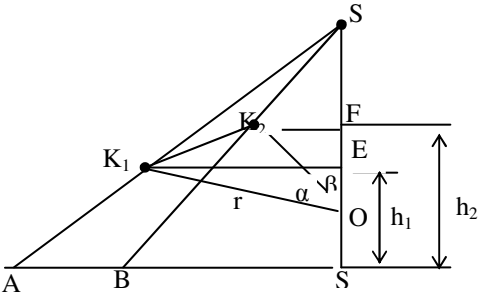


Fig. 2. The geometric relation of rotational system

2.2 Rotational System

Fig.2 describes the geometric relation of each quantity before and after rotation when the radiant rotates counter-clockwise and the positions of defect before and after rotation are in the same side of rotation axis. The meanings of symbols in fig.2 are listed as below.

O: the center of rotating axis, S: the radiant ,S1: the projection of S,OS1: the distance between O and S1 ,SS1: the distance between S and S1,K1: the position of the simulative defect before rotation, K2: the position of the simulative defect after rotation, A: the projection of K1,B: the projection of K2,AS1: the distance between A and S1,BS1: the distance between B and S1, α : the angle of the radiant rotating, h1: the depth of the defect before rotation, h2: the depth of the defect after rotation.

From Fig.2, we can see that the positions of the simulative defect before and after rotation are on the same side of line SS1. And OS1, SS1, AS1 and BS1 all can be measured directly. Let $\beta = \angle K_1OE$, $r = OK_1 = OK_2$, we can obtain equations below

$$\begin{cases} \dfrac{r\sin\beta}{BS_1} = 1 - \dfrac{OS_1}{SS_1} - \dfrac{r\cos\beta}{SS_1} \\ \dfrac{r\sin(\alpha+\beta)}{AS_1} = 1 - \dfrac{OS_1}{SS_1} - \dfrac{r\cos(\alpha+\beta)}{SS_1} \end{cases} \tag{2}$$

There are only two unknown quantity, r and β , in system of equations (2). If r and β are solved, we can obtain

$$\begin{cases} h_2 = FS_1 = r\cos\beta + OS_1 \\ h_1 = ES_1 = r\cos(\alpha+\beta) + OS_1 \end{cases} \tag{3}$$

Therefore, calibration of geometric parameter SS1. OS1 and α can be implemented according to equation (3) only if we obtain the accurate value of h1 and h2.

3 Interior Parameter Calibration

Given a sufficient number of visible points whose world coordinates are known with a high precision, as well as their corresponding observed pixel position estimate, in some optimal sense, the value of the interior parameters can be calibrated. In this paper, the interior parameter calibration is accomplished using the knowledge of corresponding points on a chessboard plane for appropriate registration of x-ray imaging system. The mesh-grid shown in Fig.3 is used to establish calibration model against geometric distortion of X-ray image. The mesh-grid provides high contrast, unambiguous points which allow a simultaneous, self-calibration of X-ray system.

The geometric distortion of image is always non-linear, but it can be regarded as linear in a small area. Therefore we can divide corresponding points in ideal image and distorted image into a series of triangle area. Take vertexes of the triangle as control points and implement linear transformation for each point in the triangle.

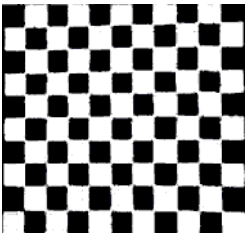


Fig. 3. Mesh-grid used to establish calibration model

Suppose the coordinates of control points are (x_1, y_1), (x_2, y_2), (x_3, y_3) in ideal image and (u_1, v_1), (u_2, v_2), (u_3, v_3) in distorted image, we can obtain:

$$\begin{cases} x_i = a \cdot u_i + b \cdot v_i + c \\ y_i = d \cdot u_i + e \cdot v_i + f \end{cases} \tag{4}$$

There, i=1,2,3. The six coefficients a, b, c, d, e and f can be obtained using known points and then calibration model is established.

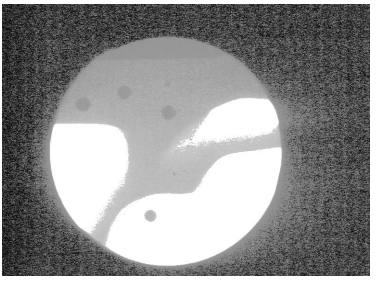
But the application of a non-linear transformation to individual image pixels will generate sub-pixel information. One way to address this issue is to subject the image to a simple affine transformation (predominantly scaling along the X- and Y-directions). These scaling factors were calculated as the average of the corresponding affine parameters computed by the calibration model.

4 Experiment and Results

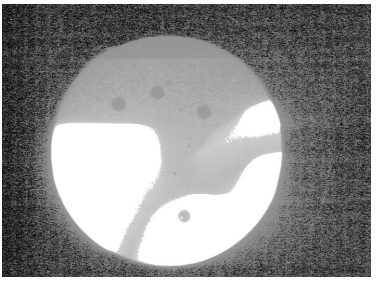
A X-ray image-intensifier is used in our experiment. The system consists of a 90 kV x-ray source, a single axis translational stage, and the x-ray image intensifier with associated electronics. X-ray image intensifier transfers weak X-ray images into images visible to human eyes. It has small volume and can be powered by a low voltage DC power supply. It is widely used in clinic diagnose and industry defects detection etc.

4.1 Parameters Calibration of Translation System

A standard test piece, of which thickness is 40mm, is used to calibrate translational system. Before imaging, four steel balls were set as marks on object surface. The steel balls, whose diameter is 4mm, are distributed roughly evenly and their depths relative to the observe side are 0mm, 0mm, 20mm, 40mm respectively. The steel balls can be distinguished from each other. The test piece was examined at a source-to-image distance of 630 cm. The distance between translational stage and image-intensifier is about 80 cm. The resulting image has a matrix of 758×576 pixels with a resolution of 2.5 line pairs per millimeter (see Fig.4).



(a) projection before translation



(b) projection after translation

Fig. 4. Image pairs for parameter calibration of translational system

Take steel balls as marks to calibrate system parameters. The relative heights of the steel balls are known with accuracy. Therefore we can establish equations according to formula (1). Finally, we obtain the precise value of source-to-image distance is $628.6cm$ and the distance between translational stage, image-intensifier is 79.4 cm and the projection shift is 12.2 cm.

4.2 Parameter Calibration of Rotational System

Take a timber wheel whose radius is 148mm as a teat piece. Make the center of the test piece be in line with rotating axis. Set three segments of thin lead thread on the test piece as simulative defects in the manner of paralleling with rotating axis. The positions of defects are shown as 1,2,3 in Fig.5, whose distances from outskirt of timber wheel are 25mm, 50mm and 80mm respectively. The angle αequals to 14 degrees and βequals to 18 degrees. Obtain the first image of test piece, as Fig.6. Then rotate the radiant 5 degrees clockwise. Obtain the second image of test piece, as Fig. 7. Fig.8 is obtained by subtracting Fig.7 from Fig.6.

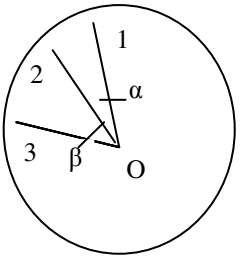


Fig. 5. Test piece with artificial focuss

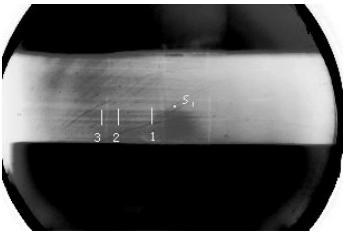


Fig. 6. Image acquired before rotation

In the experiment, the projections of defects are lines with a certain length, but points. To calculate the depth of defects, the distances between defects and S1 should be known. So we choose the upper endpoint of the projections of defects as datum marks to measure the distances. For defect 3, its measurement parameters are listed as below. AS1=44.11mm, BS_1=39.13mm, SS_1=2700mm, OS_1=204mm, $\alpha=5^0$.

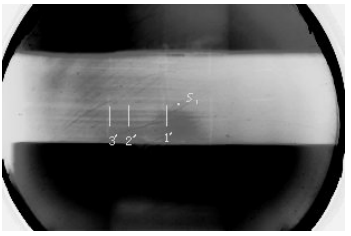


Fig. 7. Image acquired after rotation

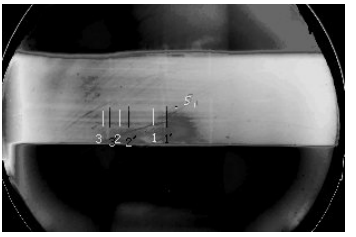


Fig. 8. The difference image

From equations (2) and (3), we can obtain h_1=254.89mm, h_2=258.17mm And the actual depths of defect 3 in two cases are h_1=258.307mm, h_2=261.667mm .So the positioning error of defect 3 is 1.34%.

Similarly, for defect 2, its measurement parameters are listed as AS1=34.12mm, BS_1=25.61mm, SS_1=2700mm, OS_1=204mm, $\alpha=5^0$. From equations (2) and (3), we can obtain h_1=290.25mm, h_2=292.57mm. And the actual depths of mark 2 in two cases are h_1=299.66mm, h_2=299.06mm. So positioning error of mark 2 is 2.18%. The depth of mark 1 also can be obtained in similar way. Its positioning error is 2.74%.

Now, we use actual values of h_1 and h_2 of three marks to calibrate SS1, OS1, α. Similarly, using equations (2) and (3), we obtain their accurate values SS_1=2702.32mm, OS_1=204.56mm, $\alpha=5.01^0$. Then recalculate the depth of focus 2 using the modified value of OS_1andα, we can get higher position accuracy.

5 Conclusion

In this paper we have provided the calibration method for X-ray stereo vision system. For interior parameters, a simple non-linear calibration scheme based on the information of corresponding points on a chessboard plane is presented. It adopts control points whose correspondences are manually established. For exterior parameters, the parameters are computed based on a small representative set of image pairs. In future, we'll study a method which can calibrate exterior and interior parameters simultaneously to simplify the calibration procedure.

References

[1] Ma, S.D.: A self-calibration technique for active vision systems. IEEE Trans. RA 12, 114–120 (2001)

[2] Pollefeys, M., Koch, R., Van Gool, L.: Self-calibration and metric reconstruction inspite of varying and unknown internal camera parameters. In: Proc. of Sixth IEEE Intl. Conf. on Computer Vision, Bombay (India), pp. 90–95 (January 2003)

[3] Shum, H.-Y., Han, M., Szeliski, R.: Interactive construction of 3D models from Panoramic Mosaics. In: Proc. IEEE Conf. Computer Vision and Pattern Recognition, Santa Barbara, pp. 427–433 (June 2003)

[4] Fu, L., Han, Y., Chen, S.: Digital Radiography Location Based on Stereo Vision. Journal of Basic Science and Engineering 2, 130–135 (2005)

[5] Tian, Y., Du, D., Hou, R., Gao, Z.: An imaging model for X- ray system and its calibration method. Transactions of the China Welding Institution 28(7), 25–29 (2007)
[6] Bookstein, F.L.: Principal warps: Thin-plate splines and the decomposition of deformations. IEEE Transactions on Pattern Analysis and Machine Intelligence 11, 567–585 (2003)
[7] Liebowitz, D., Zisserman, A.: Metric rectification for perspective images of planes. In: IEEE Conf. Comput. Vision Pattern Recog., Santa Barbara, California, pp. 482–488 (June 2004)
[8] Luong, Q.-T., Faugeras, O.: Self-calibration of a moving camera from point correspondences and fundamental matrices. Int'l J. Comput. Vision 22(3), 261–289 (2003)

Digital and Information-Processing Design of Low Power Micro Flowmeter

Xiaodan Li

Liaoning Fushun Teacher's College, Liaoning Fushun 113006, China
xdanli@126.com

Abstract. Aiming the problems, such as low minimum flow rate that can be measured, the small measuring range and low precision, a digital micro flowmeter based on MSP430 MCU and single-core structure is designed with new turbine flowmeter detection and the associated processing circuit. Watch dog, power monitoring, and filtering are combined to enhance the system reliability. Thus, the minimum measurable flow rate point is reduced with the small flow turbine flowmeter measuring range expanded. Meanwhile, protecting circuits are designed for the safety of system. Finally, to solve the noise problem, the integrated approach with reduced voltage, the FFT algorithm and power spectrum algorithm was adopted.

Keywords: Micro flowmeter, low power consumption, MCU, FFT, measurement circuit.

1 Introduction

Today, in the background of energy conservation and emission reduction, the precise metering of high precision instrument is needed, and the turbine flowmeter is widely used in various industrial fields, due to the advantages of high precision, good repeatability and small pressure loss. Meanwhile, the development of power electronics and microprocessor technologies promote the change of electronic products everyday. Specially, With the rapid development and wide application of the MCU, DSP and other digital signal processing technologies, complex signal processing algorithms are also implemented in hardware [1], [2], promoting the intellectualization and digitization of relative electronic products, including flowmeters [3]. At present, companies that have outstanding achievements in this field are Yokogawa, ABB, Foxboro, etc [4], [5].

In order to achieve low energy battery-power and reduce measurement noise, a digital micro flowmeter based on MSP430F1611 MCU of TI Company is designed in this paper. At the same time, the system has some basic measurement units such as temperature and pressure, and the two-wire current output triple intrinsic safe explosion proof protective circuit based on AM402 is used. Finally, to solve the noise problem and reduce power consumption, the integrated approach with reduced voltage, the FFT algorithm and power spectrum algorithm was adopted, at last, the digital information processing of low power micro flowmeter system is realized.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 367–373.
springerlink.com

2 Paper Preparation

As shown in Figure 1, the designed low power digital micro flowmeter consists of inductance coil and the internal circuit, stainless steel septum, blades, bearing, etc. Figure 2 shows the corresponding equivalent circuit schematic.

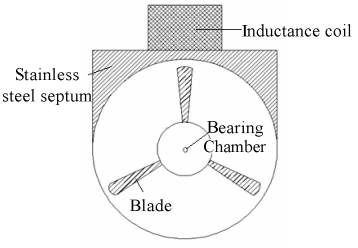


Fig. 1. Diagram of eddy current flowmeter sensor

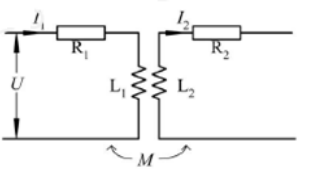


Fig. 2. Equivalent circuit of eddy current flowmeter sensor

Figure 3 shows the designed circuit structure of the system according to the eddy current flow sensor in Figure 1, it can be divided into three sections: pre-signal processing circuit, temperature and pressure measurement circuit and the MSP430F1611 master control chip and its peripheral circuit.

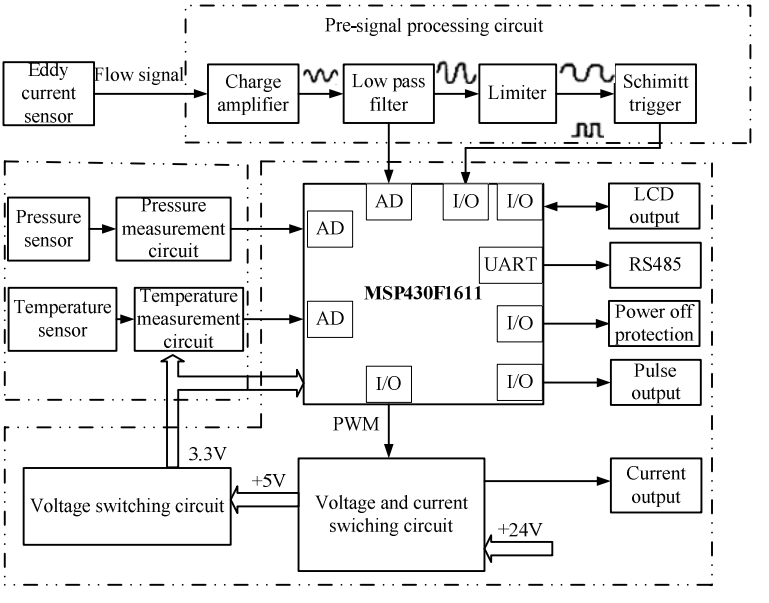


Fig. 3. System hardware design diagram

2.1 Temperature Measurement Circuit

The system setting temperature measurement ranges from -40℃ to 260℃, considering the wide temperature measuring range, high requirement of reliability in industrial applications and the requirements of low power consumption and cost factors, etc, we choose the PT1000 thermal resistance temperature sensor. The PT1000 platinum thermal resistance was choosed for two reasons: (1) its temperature measurement range is -200℃～850℃, wider than the copper's; (2) the indexing number of PT1000 is higher than PT100 in each degree, so the measurement accuracy is higher than PT100.

The temperature measurement circuit is designed as shown in Figure 4, it uses constant voltage power supply bridge structure, and can be directly drived by the 2.5V vlotage supplied by MCU, making it easy to control the circuit, thus, the power consumption is reduced and the design of the system is simplified.

As shown in the figure, the MCU output driving voltage 2.5V can be controlled by the program, when measuring the temperature, the 2.5V voltage driving measurement circuit is on, while after the measurement, the reference voltage is cut to reduce power consumption. When the medium temperature is -45.713℃, the resistance of PT1000 is 820Ω, and the output voltage difference of two bridges is 0, corresponding to the zeropoint of bridge. The input signal of instrument amplifier INA321 at PT1000 arm is VIN+, and the input signal at 820Ω resistance arm is VIN-, then the INA321 magnifies the voltage difference of VIN+ and VIN-, the amplification factor is G=5+5(W5/R1). Also, when the medium temperature is up to 260 ℃, the INA321 on both ends of the bridge makes the Temp_Signal fully output 2V by using W5 to adjust the amplification factor.

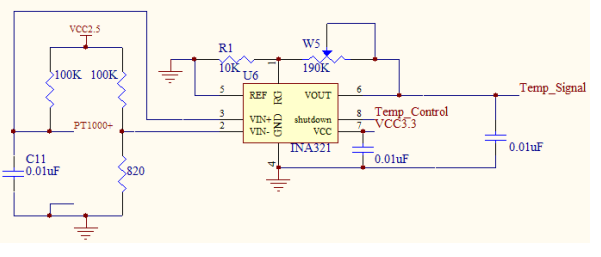


Fig. 4. Temperature measurement circuit

Table 1 shows the circuit test results of different platinum thermal resistance. In order to reduce the error produced from the nonlinearity of bridge structure temperature measurement circuit, the platinum thermal resistances are calculated according to the resistance parameters of the three arms and the amplification factor determined by adjusting the full degrees, then the nonlinearity brought by the circuit and the degree form are eliminated through checking the degree form to calculate the temperature. So this bridge must choose precise resistance. Through the test results in table 1, it can be seen that the full error is 0.5%.

Table 1. Temperature measurement data

Thermal resistance testing point (Ω)	Bridge voltage difference (V)	amplifier output voltage (V)	Sampling resistance (Ω)
1	*2370*	*0.0592*	*2.01*	*2370*
2	*2170*	*0.0528*	*1.793*	*2175*
3	*1990*	*0.0459*	*1.560*	*1988*
4	*1780*	*0.0390*	*1.320*	*1783*
5	*1580*	*0.0316*	*1.070*	*1586*
6	*1290*	*0.0205*	*0.695*	*1298*

2.2 Pressure Measurement Circuit

To reduce power consumption, the pressure sensor is also supplied by constant pressure source, the minimum voltage provided by the MCU is 2.5V, and the experimental results proved the current of pressure sensor was only 0.6mA.

Figure 5 shows the pressure measurement circuit, similar to the temperature measurement circuit, according to the measurement need, the 2.5 V voltage is on every once in a while to to collect the temperature and pressure signals. When not working, the reference voltage is cut to reduce power consumption. The output voltage difference is the input of the instrumentation amplifier, and its output voltage is input to the MCU, then the corresponding medium pressure is calculated.

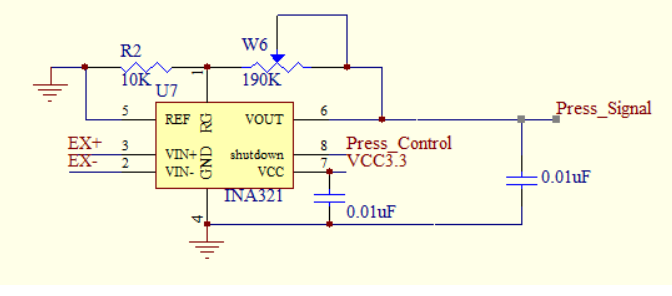


Fig. 5. Pressure measurement circuit

Table 2 shows the calibration results of pressure sensors on the pressure calibrator. Because it is hard to judge the balancepoint between the weights and oil pressure when the piston pressure calibrator is used to calibrate big pressure, it will cause big personnel errors, therefore only 0.1 ~ 0.7 MPa are calibrated.

Table 2. Pressure measurement data

Absolute pressure (Mpa)	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**
Output voltage difference (mV)	*49.326*	*98.676*	*148.20*	*198.56*	*248.86*	*298.50*	*347.82*

2.3 Intrinsic Safe Explosion Design

The paper designed and experimented the explosion proof measures according to the technical requirements of ib grade[6], [7]. Figure 6 shows the triple intrinsic safe protective design of voltage switching circuit.

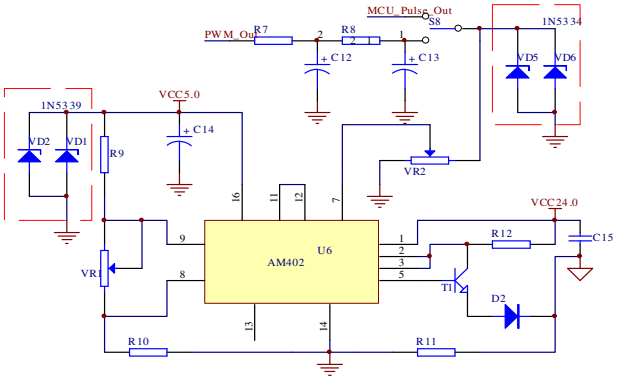


Fig. 6. Safe voltage conversion circuit

From the figure we can see that two stabilovolts parallel at the 5 V voltage output of AM402, and the stabilovolts VD1, VD2 and that of AM402 are combined to form the triple voltage stabilization.

In the signal input section of AM402, if S8 is connected to the output part of PWM_OUT, the filtered output PWM wave is converted into smooth voltage signal which is connected to the input of AM402; while if S8 is connected to the output part of MCU_OUT, the 3.3V pulse signal is input into the AM402 chip. Therefore, the dual voltage protection is used at the PWM wave output side of MCU. The output voltage of MCU is 3.3V, and the maximum input voltage signal of AM402 is also 3.3 V, so choose the 1N5336 stabilovolt with maximum stabilovolt voltage 4.3V.

3 The Main Program Design

The main program flowchart of micro flowmeter based on the MSP430 MCU is shown in Figure 7. By the combination of hardware circuitry, anti-jamming

technology, FFT and modular programming ideas [8], [9], the entire system can be divided into the following modules: the initialization module, the frequency calculation module, the flow calculation and the compensation module, the power-down protection and parameter-saving module, etc. The flowchart of FFT-based frequency capturing and calculation processing is shown in Figure 8.

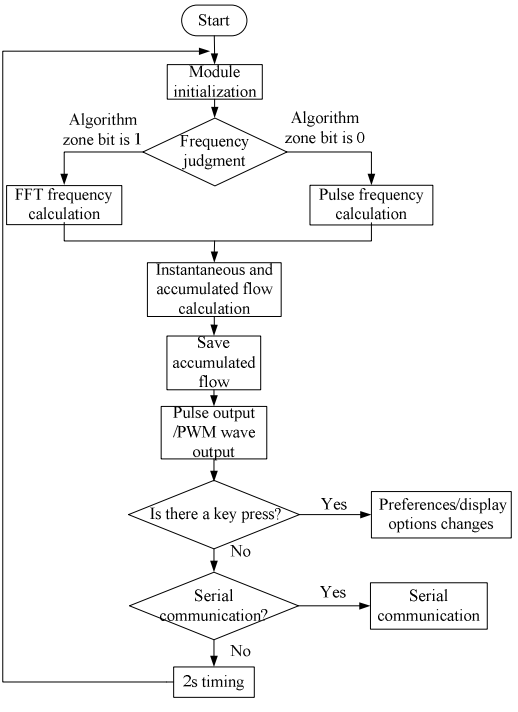


Fig. 7. Main flowchart

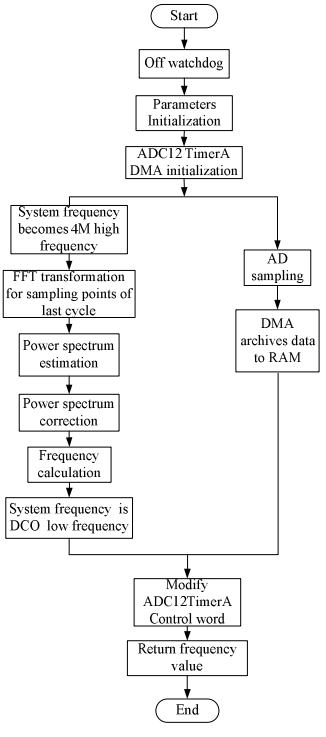


Fig. 8. FFT calculation

4 Results

Adopt the above hardware and software design, a turbine flowmeter with DN4mm diameter is selected randomly, now use the electrical eddy signal detector to do the experiments respectively. The results of measuring the micro turbine flowmeter based on FFT are shown in Figure 9 and Figure 10. Flow of 40 L / h and 20 L / h is respectively shown in Figure 9, Figure 10. From the figures we can see that the method can detect micro electrical eddy signals no matter the flow is high or low. Meanwhile, the experiment also demonstrates that the upper limit of the method can reach 200L/h, so the measurement range can be up to 10:1.

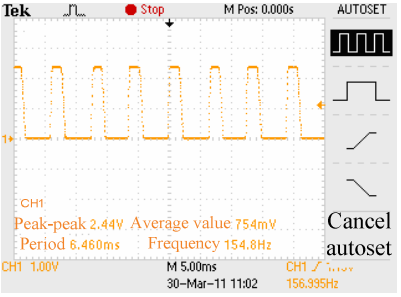


Fig. 9. Experimental results for DN4 mm with 40 L/h

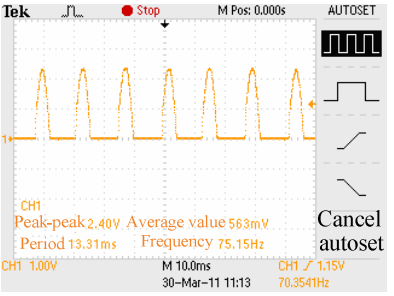


Fig. 10. Experimental results for DN4 mm with 20 L/h

5 Conclusion

Based on the MSP430 MCU, the digital low power micro flowmeter is designed; it has fast response speed, better anti-jamming capability and complete function. Experiments showed that the flowmeter has lower power consumption and wider measurement range, meanwhile, by reducing the sensor supply voltage from 12V to 3.3V. At the same time, the reliability and flexibility of the system are improved by using the FFT algorithm for spectral analysis of the micro flow signal.

References

1. Xia, C.: Permanent magnet Brushless DC Motor Drives and Control. John Wiley & Sons Singapore Pvt. Ltd., Singapore (2012)
2. Wei, X.: Interface technology and system design examples of MSP430 MCU. Beijing University of Aeronautics and Astronautics press, Beijing (2002)
3. Liang, G., Cai, W.: The flow measurement technology and instruments. Mechanical industry press, Beijing (2002)
4. Cai, W.: Current situation of micro flow measurement. The World Instrument and Automation 8(1), 26–29 (2004)
5. Gu, J., Tang, W.: Research of high precision electric eddy current micro-measuring system. Jiangsu University of Science and Technology Journal 2, 73–77 (2006)
6. Wang, L., Xu, G.: Explosion-proof measures of battery powered intrinsically safe equipments. Electric Explosion Protection 3, 7–9 (2008)
7. Yang, M., Xu, Y.: Analysis of explosion protection measures of electrical equipments. Wuxi Commercial Professional Technology Institute Journal 12, 96–98 (2009)
8. Zhang, Y., Shen, N.: Fast Fourier transform and Walsh transformation. Aviation industry press (1989)
9. Song, Y.: Practical design based on the software anti-jamming of MCU. Software Guide 4, 11–12 (2010)

Image and Video Processing on GPU: Implementation Scheme, Applications and Future Directions

Roberto Di Salvo and Carmelo Pino

Department of Electrical, Electronics and Computer Engineering,
Catania University, Viale Doria 6, Catania, Italy
{roberto.disalvo,carmelo.pino}@dieei.unict.it

Abstract. Most of the recent computer graphic applications are essentially based on multicore general-purpose processors architectures, which include CPUs made up of parallel processors with high elaboration capability. Due to the rapid turning towards high definition multimedia, much more memory space and computational resources are needed to achieve better performance. Recently the GPGPUs (which stands for General-Purpose computing on Graphic Processing Units) architectures have been adopted. Since the GPUs are becoming increasingly programmable and provide a vast number of simple, data-parallel, deeply multithreaded cores and high memory bandwidths, they have been used in several research fields such as image and video processing, medical and human applications. This work aims at providing a meticulous review of the most recent image and video processing techniques and significant applications, highlighting both the advantages of using CUDA in terms of efficiency and the guide line for future development by means of these emerging techniques.

Keywords: CUDA platform, Parallel programming, Multicore systems, Image processing, Medical applications.

1 Introduction

Since the massive parallelism, high memory bandwidth, and general purpose instruction sets (including single/double precision IEEE floating point arithmetic) of Computed Unified Device Architecture (CUDA) systems, it makes convenient its application in several fields [1], especially for graphical applications where the power of the graphic processing units fully takes advantage from the multi-threaded capabilities of the multi-core structure [2].

A lot of applications have achieved intense speedup using GPGPU (which stands for General-Purpose computing on Graphic Processing Units) implementation concerning the development of parallel primitive to simplify the project of GPGPU applications [3]. In particular, the GPUs have recently used both for solving general purpose problem such as linear algebra issues [4], cloud dynamics simulation using partial differential equations [5] and in the field of computer vision, e.g., [6-7] especially for all those operations characterized by multiple iterations and high memory requisite, making this new technology increasingly competitive as regards price, speed and programmability. For these reasons, CUDA systems are taken into

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 375–382.
springerlink.com

account more and more as the reference hardware architecture for parallel processing created by NVIDIA.

In this paper we illustrates the effectiveness of CUDA as a model to provide parallel computation with several sets of performance features on GPUs by using some small programming examples and relevant CUDA applications.

In particular, the paper first discusses the advantage of CUDA programming on GPUs starting from simple examples and giving the limits of CPUs multicore technologies, then an analysis is carried out to point out the effectiveness of the CUDA model on the basis of implementation scheme of the most recent applications for GPU video and image processing in medical fields and for video-surveillance applications. Future directions are given in the last section.

2 The CUDA Programming Model

Fig.1 shows a reference schema for CUDA threads model to point out how simple C code may be executed very fast by thousands of parallel invocations of threads in the CUDA model. The data processed by the GPU are stored in the graphics board's memory and data transfer is done by DMA (Direct Memory Access).

Kernels are defined as the sequence of work to be done in each thread at a single point in the domain. This one is defined with a structure composed by a grid of 3D thread blocks. Fig.2 shows two different block distributions, when GPU has two or four multiprocessors [8]. Synchronization within a thread block is managed by allowing a kernel to complete and starting a new kernel.

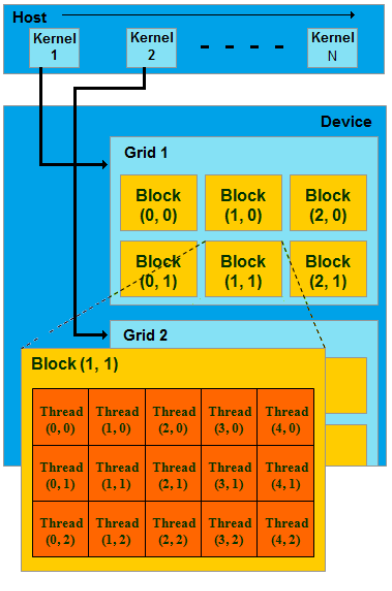


Fig. 1. CUDA model: a program consists of kernels assigned to a grid of blocks, each block contains threads

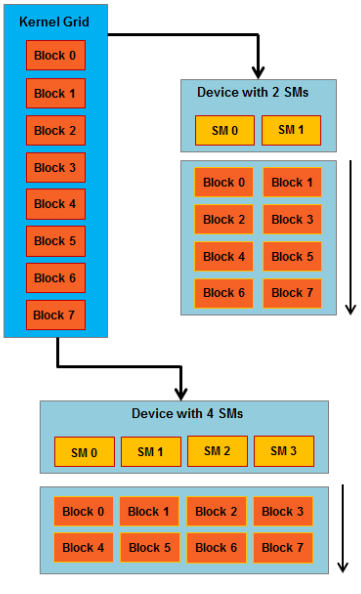


Fig. 2. An example of block distribution among GPUs with different number of multiprocessors

All thread within the thread block are live all together and are temporally multiplexed onto the processing elements in a fine-grained, time-sliced manner, but their resources cannot be reclaimed until the entire block of threads completes. Each block could use its own shared memory (SM) (besides the global memory) visible only for the thread and allow thread block to cooperate by sharing data among themselves with low latency.

Fig.3 shows an example of CUDA procedure to transfer two matrices to the GPU and calculate their sum. Once the device vector have been allocated, the input matrices, read from files, are copied into the blocks of the device (see fig.2). Then, the main kernel, activated by the *run kernel*, is executed by producing the sum in parallel of the corresponding blocks of the two matrices to be summed. Finally the result is passed to the host and the memory is released.

```
/* Define kernel */
__global__ void sumMatrix(int* a_matrix, int *b_matrix, int *c_matrix, int rows, int cols) {
        // Compute element coordinates
        int x = blockIdx.x*blockDim.x + threadIdx.x;
        int y = blockIdx.y*blockDim.y + threadIdx.y;
        // Check limits
        if(x >= cols || y >= rows){
        return; }
        // Compute sum
        c_matrix[x + y*cols] = a_matrix[x + y*cols] + b_matrix[x +  y*cols];
}

/* Allocate device vector */
  error = cudaMalloc(&d_a_matrix, rows*cols*sizeof(int));
  if(error != cudaSuccess) {
        printf("Device malloc failed: %s\n", cudaGetErrorString(error));
        return 1; }
  // The same code for other matrices

/* Create the input matrices */

/* Copy matrices to device */
error = cudaMemcpy(d_a_matrix, h_a_matrix, rows*cols*sizeof(int), cudaMemcpyHostToDevice);
if(error != cudaSuccess) {
        printf("Matrix copy error: %s\n", cudaGetErrorString(error));
        return 1;}
  // The same code for other matrices

/* Run kernel */
sumMatrix<<<grid_size, block_size>>>(d_a_matrix, d_b_matrix, d_c_matrix, rows, cols);

/* Wait for kernel execution to finish */
error = cudaThreadSynchronize();
…
/* Copy data from device to host */
error = cudaMemcpy(h_c_matrix, d_c_matrix, rows*cols*sizeof(int), cudaMemcpyDeviceToHost);
…

/* Free memory */
```

Fig. 3. An example of CUDA programming: sum of matrices

3 CUDA Applications and Implementation Scheme

Before the advent of general purpose languages for GPGPU, GPU implementations could only be achieved using existing 3D-rendering APIs such as OpenGL or DirectX. In fact, in the earliest versions, the GPU based applications were implemented via a graphics language which limited its flexibility. The syntax, the problems concerning the context of polygon rasterization and the restrictions imposed by pixel independence all made this approach uncomfortable.

On the other hand, in the last few years the CPUs architecture have been increasingly improved with "many-core" processors, providing high parallelism, but a lot of operations could be executed in a tiny time, especially in research fields such as i) video-surveillance for vehicles [9-10], animals [11-13] and humans [14-16], ii) medical image analysis [17-21], and iii) image restoration [22-24]. For this reason, GPU vendors added driver and hardware support to use the highly parallel hardware of the GPU without the need for computation to proceed through the entire graphics pipeline and without using 3D APIs at all.

Parallel implementations of such applications have been developed more and more using GPU systems with the aim of achieving high speedups ranging from 10× to 100× and above. With the advent of NVIDIA's CUDA platform a hugely multithreaded general purpose architecture is provided, with up to 512 processor cores and thousands of threads executed at the same time. CUDA is programmable in C and capable of millions of floating-point operations per second [25]. This allowed the CUDA technology to be adopted across several domains ranging from image and video processing applications [26-30], computational chemistry [31], astrophysics [33], biomedical purposes including for example gene sequencing [33], CT reconstruction [34], and medical image segmentation [35-36]. Recently, even in video-surveillance applications the CUDA technology has been successfully adopted giving an important contribution in the detection and tracking tasks such a fast pedestrian localization, pedestrian recognition for automotive and security applications and face detection and recognition,e.g. [37]. Fig. 4 shows the typical compilation schema for CUDA programs in which the integrated GPU and CPU program is compiled by NVCC, and then the GPU code and CPU code are separated.

To give an idea of how the implementation schema changes when passing from sequential to CUDA programs, in the following we show the parallelization of four emblematic algorithms belonging to the mentioned application fields.

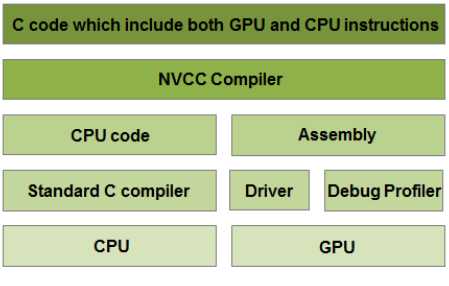


Fig. 4. CUDA compilation process

Fig.5 (left) shows the CUDA workflow of the parallelization scheme used in Pyramidal implementation aiming at computing a dense and accurate velocity field according the well-known GPU pyramidal algorithm of Lucas & Kanade, where we show that the main part of the algorithm (included in the dotted box) may be executed on a CUDA platform with a 100× speed-up. A speed-up of the same magnitude may be achieved also when executing CUDA applications to carry out the mentioned user applications on medical and surveillance fields since they are based on image processing techniques that can be easily parallelized on a CUDA platform.

This is illustrated with the workflows of the two methods drawn in fig.5, i.e.: a) one in fig.5 (center) method proposed by Gu et al. in [38] to treat the requirement of rapid image segmentation for the clinical implementation of online adaptive radiation therapy (ART), and b) the other method in fig.5 (right) dealing with the sequence of steps of the proposed implementation for fast pedestrian localization framework that integrates the cascade-of-rejectors approach with the Histograms of Oriented Gradients (HoG) features, selecting the most informative histogram blocks with the AdaBoost algorithm, in a parallel architecture [39]. The workflows confirm that the gradient calculations to be carried out in both methods, pointed in the dotted boxes, may fully exploit the CUDA parallelism thus achieving a relevant speed-up.

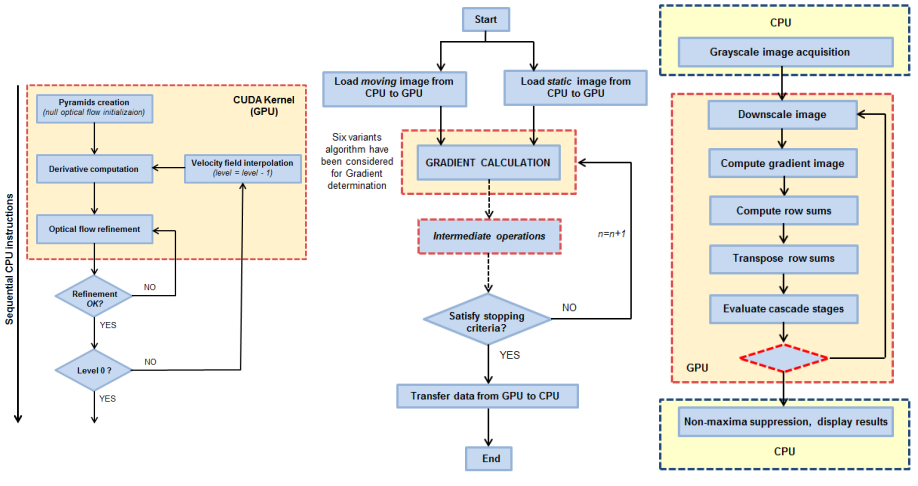


Fig. 5. CUDA implementation scheme to speed up relevant algorithms: optical flow for motion detection, image segmentation for radiation therapy, people detection and tracking.

4 Conclusions and Future Directions

CUDA based methods for image and video processing in different domains have been illustrated pointing out that they achieved very high time performance maintaining the same accuracy of the sequential versions of these methods. The main part of the available CUDA based approaches dealt with the parallelization of generic-purpose image processing tasks.

However, in future we expect that CUDA based techniques should allow to improve medical and engineering applications that involve complex diagnostic and

control tasks such as parallel clustering methods for literature mining, e.g., [40], real time control of traffic and people flows, e.g.[41], knowledge discovery from biodata, e.g. [42], [43]. Indeed, these systems are very time-consuming making their application on real life environment a challenge.

It is also expected the porting to GPU systems of existing algorithms or entire image processing environments to facilitate the processing and interpretation of complex images and videos, e.g., the porting of image processing platform proposed in [44], by using either the available CUDA systems or the incoming GPU-equipped computers [45]. This trend will be further consolidated if the programmability of the CUDA platform will be improved.

References

1. Che, S., Boyer, M., Meng, J., Tarjan, D., Sheaffer, J.W., Skadron, K.: A performance study of general-purpose applications on graphics processors using CUDA. J. Parallel Distrib. Comput. 68, 1370–1380 (2008)
2. Pan, J., Manocha, D.: GPU-based Parallel Collision Detection for Real-Time Motion Planning. STAR, vol. 68, pp. 211–228 (2011)
3. Sengupta, S., Harris, M., Zhang, Y., Owens, J.D.: Scan primitives for GPU computing. In: Proc. of the ACM SIGGRAPH/EUROGRAPHICS Conf. on Graphics Hardware (2007)
4. Krüger, J., Westermann, R.: Linear algebra operators for GPU implementation of numerical algorithms. ACM Transactions on Graphics 22(3), 908–916 (2003)
5. Harris, M.J., Baxter, W.V., Scheuermann, T., Lastra, A.: Simulation of cloud dynamics on graphics hardware. In: Proc. of the ACM SIGGRAPH/EUROGRAPHICS (2003)
6. Mairal, J., Keriven, R., Chariot, A.: Fast and Efficient Dense Variational Stereo on GPU. In: Proceedings of International Symposium on 3D Data Processing, Visualization and Transmission, pp. 97–704 (2006)
7. Fung, J., Man, S.: OpenVIDIA: Parallel GPU Computer Vision. In: Proceedings of ACM International Conference on Multimedia, pp. 849–852. ACM (2005)
8. NVIDIA CUDA Programming Guide version 2.3 (January 7, 2009), `http://developer.download.nvidia.com/.../cuda/.../NVIDIA_CUDA_Programming_Guide_2.3.pdf`
9. Faro, A., Giordano, D., Spampinato, C.: Integrating location tracking, traffic monitoring and semantics in a layered ITS architecture. IET Journal Intelligent Transportation Systems 5(3), 197–206 (2011)
10. Faro, A., Giordano, D., Spampinato, C.: Evaluation of the traffic parameters in a metropolitan area by fusing visual perceptions and CNN processing of webcam based images. IEEE Transactions on Neural Networks 19(6), 1108–1129 (2007)
11. Faro, A., Giordano, D., Spampinato, C.: Soft-computing agents processing webcam images to optimize metropolitan traffic systems. In: Computer Vision and Graphics, Computational Imaging and Vision, vol. 32, pp. 968–974. Springer Book Series (2006)
12. Spampinato, C., Giordano, D., Di Salvo, R., Chen Burgher, J., Fisher, R.B., Nadarajan, G.: Automatic Fish Classification for Underwater Species Behavior Understanding. In: Proc. of the first ACM Int. Workshop on Analysis and Retrieval of Tracked Events and Motion in Imagery Streams, ARTEMIS 2010. ACM (2010)
13. Spampinato, C.: Adaptive Objects Tracking by using Statistical Features Shape Modeling and Histogram Analysis. In: Proceedings of the 7th IEEE International Conference on Advances in Pattern Recognition, ICAPR 2009. IEEE (2009)

14. Spampinato, C., Chen Burger, J., Nadarajan, G., Fisher, R.B.: Detecting, Tracking and Counting Fish in Low Quality Unconstrained Underwater Videos. In: Proceedings of VISAPP 2008, Madeira, Portugal (2008)
15. Devrari, K., Vinay Kumar, K.: Fast Face Detection Using Graphics Processor. International Journal of Computer Science and Information Technologies 2(3), 1082–1086 (2011)
16. Tang, S., Goto, S.: Multi Scale Block Histogram of Template Feature for Pedestrian Detection. In: Proceedings of 2010 IEEE 17th International Conference on Image Processing. IEEE (2010)
17. Faro, A., Giordano, D., Spampinato, C.: An Automated Tool for Face Recognition Using Visual Attention and Active Shape Models Analysis. In: 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2006, pp. 4848–4852. IEEE (2006)
18. Giordano, D., Leonardi, R., Maiorana, F., Spampinato, C.: Cellular Neural Networks and Dynamic Enhancement for Cephalometric Landmarks Detection. In: Rutkowski, L., Tadeusiewicz, R., Zadeh, L.A., Żurada, J.M. (eds.) ICAISC 2006. LNCS (LNAI), vol. 4029, pp. 768–777. Springer, Heidelberg (2006)
19. Faro, A., Giordano, D., Spampinato, C., Ullo, S., Distefano, A.: Basal ganglia activity measurement by automatic 3D Striatum segmentation in SPECT images. IEEE Transactions on Instrumentation and Measurement, Special Issue on Medical Measurements, 3269–3280 (2011)
20. Giordano, D., Spampinato, C., Scarciofalo, G., Leonardi, R.: An Automatic System for Skeletal Bone Age Measurement by Robust Processing of Carpal and Epiphysial/ Metaphysial Bones. IEEE Transactions on Instrumentation and Measurement 59(10), 2539–2553 (2010)
21. Leonardi, R., Giordano, D., Maiorana, F., Spampinato, C.: Automatic Cephalometric Analysis: A Systematic Review. The Angle Orthodontist 78(1), 145–151 (2007)
22. Giordano, D., Leonardi, R., Maiorana, F., Scarciofalo, G., Spampinato, C.: Epiphysis and metaphysis extraction and classification by adaptive thresholding and DoG filtering for automated skeletal bone age analysis. In: Proc. 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 2007, vol. 2007, pp. 6551–6556 (2007)
23. Faro, A., Giordano, D., Spampinato, C.: Neural Network Combined with Fuzzy Logic to Remove Salt and Pepper Noise in Digital Images. In: Applications of Soft Computing: Advances in Soft Computing, vol. 36, pp. 23–33. Springer, Heidelberg (2006)
24. Faro, A., Giordano, D., Scarciofalo, G., Spampinato, C.: Bayesian Networks for Edge Preserving Salt and Pepper Image Denoising. In: First Workshops on Image Processing Theory, Tools and Applications, IPTA 2008, pp. 1–5 (2008)
25. Cannavò, F., Nunnari, G., Giordano, D., Spampinato, C.: Variational Method for Image Denoising by Distributed Genetic Algorithms on GRID Environment. In: 15th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, WETICE 2006, pp. 227–232. IEEE (2006)
26. NVIDIA. CUDA Programming Guide 1.1 (2007), http://developer.download.nvidia.com/compute/cuda/1_1/NVIDIA_CUDA_Programming_Guide_1.1.pdf
27. Yu, Q., Medioni, G.: A GPU-based implementation of Motion Detection from a Moving Platform. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp. 1–6 (2008)
28. Bilgic, B., Horn, B.K.P., Masaki, I.: Efficient Integral Image Computation on the GPU. In: Intelligent Vehicles Symposium (IV), pp. 528–533. IEEE (2010)

29. Catanzaro, B., Su, B.-Y., Sundaram, N., Lee, Y., Murphy, M., Keutzer, K.: Efficient, high-quality image contour detection. In: IEEE 12th International Conference on Computer Vision, pp. 2381–2388 (2009)
30. Colic, A., Calva, H., Furht, B.: Exploring NVIDIA-CUDA for Video Coding. In: MMSys 2010 Proceedings of the First Annual ACM SIGMM Conference on Multimedia systems. ACM (2010)
31. Yang, Z., Zhu, Y., Pu, Y.: Parallel Image Processing Based on CUDA. In: 2008 International Conference on Computer Science and Software Engineering, pp. 198–201 (2008)
32. Stone, J., Phillips, J., Hardy, D., Freddolino, P., Trabuco, L., Schulten, K.: Accelerating molecular modeling applications with graphics processors. Journal of Computational Chemistry 28(16), 2618–2640 (2007)
33. Belleman, R.G., Bédorf, J., Portegies, S.F.: High Performance Direct Gravitational N-body Simulations on Graphics Processing Units II: An implementation in CUDA, arXiv:0707. 0438v2 [astro-ph] (July 2007)
34. Schatz, M.C., Trapnell, C., Delcher, A.L., Varshney, A.: MUMmerGPU: High-throughput sequence alignment using Graphics Processing Units. BMC Bioinformatics 8, 474 (2007)
35. Scherl, H., Keck, B., Kowarschik, M., Hornegger, J.: Fast GPU-Based CT Reconstruction using the Common Unified Device Architecture (CUDA). In: 2007 IEEE Nuclear Science Symposium Conference Record, October 26-November 3, pp. 4464–4466 (2007)
36. Pan, L., Gu, L., Xu, J.: Implementation of medical image segmentation in CUDA. In: International Conference on Information Technology and Applications in Biomedicine, pp. 82–85 (2008)
37. Faro, A., Giordano, D., Spampinato, C.: An automated tool for face recognition using visual attention and active shape models analysis. In: Conf. Proc. IEEE Engineering in Medicine and Biology Society, EBCS 2006, vol. 1, pp. 4848–4852 (2006)
38. Ruiz, A., Jun, K., Ujaldon, M., Boyer, K., Saltz, J., Gurcan, M.: Pathological image segmentation for neuroblastoma using the GPU. In: 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 296–299. IEEE (2008)
39. Bilgic, B., Horn, B.K.P., Masaki, I.: Fast human detection with cascaded ensembles on the GPU. In: 2010 IEEE Intelligent Vehicles Symposium (IV), pp. 325–332 (2010)
40. Faro, A., Giordano, D., Maiorana, F.: Mining massive datasets by an unsupervised parallel clustering on a grid: Novel algorithms and case study. Future Generation Computer Systems 27(6), 711–724 (2011)
41. Faro, A., Giordano, D., Spampinato, C.: Adaptive background modeling integrated with luminosity sensors and occlusion processing for reliable vehicle detection. IEEE Transactions on Intelligent Transportation Systems 12(4), 1398–1412 (2011)
42. Faro, A., Giordano, D., Maiorana, F., Spampinato, C.: Discovering genes-diseases associations from specialized literature using the grid. IEEE Transactions on Information Technology in Biomedicine 13(4), 554–560 (2009)
43. Faro, A., Giordano, D., Spampinato, C.: Combining literature text mining with microarray data: advances for system biology modeling. Briefings in Bioinformatics 13(1), 61–82 (2012)
44. Crisafi, A., Giordano, D., Spampinato, C.: GRIPLAB 1.0: Grid Image Processing Laboratory for Distributed Machine Vision Applications. In: Proc. Int. Worshop on Enabling Technologies: Infrastructure for Collaborative Enterprises, WETICE 2008. IEEE (2008)
45. Owens, J.D., Luebke, D., Govindaraju, N., Harris, M., Krüger, J., Lefohn, A.E., Purcell, T.: A Survey of General-Purpose Computation on Graphics Hardware. Computer Graphics Forum 26(1), 80–113 (2007)

The Research on Liquor Identifications Based on Digital Signal Processor

Lianjun Hu[1], Hong Song[2], and Xiaohui Zeng[2]

[1] School of Mechanical Engineering, Sichuan University of Science & Engineering, 643000, Zigong, China
[2] School of Automation and Electronic Information Engineering, Sichuan University of Science & Engineering, 643000, Zigong, China
{hlj28288,cgx8287,sh8887878}@sina.com, xh-z@sohu.com

Abstract. A liquor identification system based on DSP is designed in the paper, in which neural networks are introduced into the liquor identification system to make the system more scientific. Gas sensor arrays and neural networks are used for realizing pattern recognition, and liquors are identified through simulated biological olfactory.

Keywords: Neural network, Liquor identification, Sensor array, DSP.

1 Introduction

China is a white wine production and consumption country. Compositions of liquors are mainly alcohol and water. Other components are 1~2 percentage of the total, 342 kinds of which can be detected out, including methanol, polyols, carboxylic acids, aldehydes, esters, etc. Components detection and control is most important in the liquor quality control. Different odors can be identified and recognized through the olfaction. Gas sensor arrays and neural networks are used for realizing pattern recognition, and liquors are identified through simulated biological olfactory[1-2].

2 The Structure of the Liquor Identification System

The liquor identification system is based on TMS320C6713 DSP chip, working with gas sensors, A/D transformer circuits, interface circuits, a keyboard, display devices, memory circuits and a PC. The modular design method is adopted in the design of the system structure, as shown in figure 1.

Along with the development of microelectronic technologies, the calculation abilities and processing speeds of various kinds of CPU are getting more and more powerful. Product TMS320C6713 of C6000 series by the TI Corporation is selected as the CPU of the system. TMS320C6713 is one of high-speed floating-point DSP products with relatively powerful calculation ability introduced by the TI Corporation. The maximum working main frequency of the DSP chip is up to 300MHz and its processing speed is up to 2400MI/s. It has optimal instruction sets specialized for

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 383–387.
springerlink.com

algorithms and high-level languages. There are 264 KB×8 bits memories. The DSP chip also has abundant peripheral resources, including two sets of I^2C buses, two McASP, two McBSP, two 32-bit general timer, a GPIO, a 16-bit host interface HPI. The characteristics of the chip make itself an ideal DSP processor with advantages of small size, high performance and low cost[2-3].

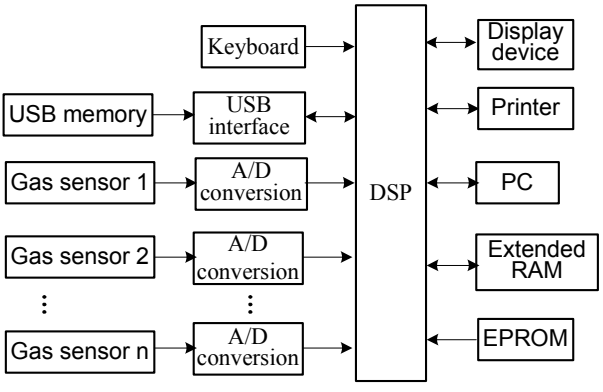


Fig. 1. The global construction of the multi-axes control system

AD7329 is adopted as A/D converter to convert analog signals into digital signals. AD7329 is a 12-bit plus sign successive approximation ADC with 8 channels and it has true bipolar input ranges. The analog input channels on the AD7321 are programmed to be single-ended, true differential, or pseudo differential. All analog channels can be programmed independently. The ADC has a high speed serial interface that can operate at throughput rates up to 500 kSPS. The ADC contains a 2.5 V internal reference and it also allows for external reference operation.

3 Gas Sensor Array

Sense of smell is a chemical feeling, a nerve impulse of olfactory nerve fibers when olfactory cells are stimulated and it is a physiological response of the biological olfactory system to materials dispersed in the air. A receptor cell may be sensitive to more than 10 kinds of different smells, showing a very wide response spectrum.

A gas sensor is a device which can detect specific components and concentrations in gases and converts them into appropriate signals. Gas sensors are used to stimulate the sense of smell. In the liquor identification, the selection and application of gas sensors and their array is essential. At present, gas sensors commonly used are conductivity type semiconductor gas sensors, semiconductor FET gas sensors, surface acoustic wave gas sensors, conductivity type organic polymer membrane gas sensors, LB films gas sensors, quartz crystal microbalance gas sensors, electrochemical type gas sensors, optical fiber gas sensors, etc[4].

It is not satisfied in sensitivity, the resolution of a complex mixture and adaptability of a changing environment when individual gas senor is used in detections of gas mixtures with complicated compositions.

A gas sensor array is a simulated biological olfactory system using the feature that a variety of gas sensors response to complex compositions of gases differently to form feature vectors, with greater information processing capability, higher resolution accuracy and affordability of the harsh environment than artificial olfactory systems. Multiple gas sensors with different selectivity are used in the gas sensor array, in order to improve its sensitivity to different gases and achieve the analysis of gas mixtures, and to overcome the cross-sensitivity problem of a single gas sensor[2-5].

For extracting enough information of samples so that the analysis results may be as accurate and reliable as possible, gas sensors are used as much as possible. But the number of gas sensors in the design of the array should be determined in accordance with the specific circumstances. Because too much gas sensor for a particular gas will take up a lot of storage space and computing time, resulting in unnecessary waste and delay. Therefore, gas sensors in the array are not required to have a very strong selectivity, but they should have a broad-spectrum response and a preferential response to certain types of gases or odours in order to obtain the optimal gas sensor array.

The gas senor array designed consists of eight gas sensors, a humidity sensor and a temperature sensor. All sensors are arranged on a gas sensor array circuit board. Outputs of sensors are converted from analog signals into digital signals by the A/D converter, and then transmitted into DSP for subsequent processing. Eight gas sensors selected should be sensitive to gases with different sensitivities. This allows the sensor array to form a high-dimensional response pattern for gases being measured in general. And it can also use the selectivity of each sensor to emphasize composition information of a certain gas, which may result a more accurate and reliable analysis.

4 Pattern Recognition

Pattern recognition methods commonly used in liquor identifications are statistical methods and artificial intelligence methods. Statistical methods often used are cluster analysis and principal component analysis. Artificial intelligence methods commonly used are artificial neural networks and wavelet transformation theories.

The human brain is a highly complex, nonlinear, parallel processing information processing system connected complicatedly by a large number of basic units(neurons). The artificial neural network is to mimic the human brain from the perspective of artificial intelligence, making use of modern neurobiology and cognitive science to explore a new method of information representation, storage and processing. The artificial neural network is a network of neurons interconnected. Compared to the traditional pattern recognition techniques, the artificial neural network pattern recognition has advantages of good fault tolerance and greater ability to adapt, parallel processing and self-learning and reasoning abilities, distributed storage of information, environmental adaptability and so on. Therefore, the artificial neural network is adopted to achieve pattern recognitions[1].

A multilayer feed-forward neural network(BP neural network) is used in the system designed. It is composed of three layers: input layer, hidden layer and output layer. The output of the previous layer of neurons is the input of the next layer of neurons. Neurons of each layer will couple with that of other layers while no coupling

between neurons in the layer. According to network characteristics, forward and back propagations are used to adjust weights of neurons so that a good mapping relationship is obtained between the input and the output, as shown in figure 2.

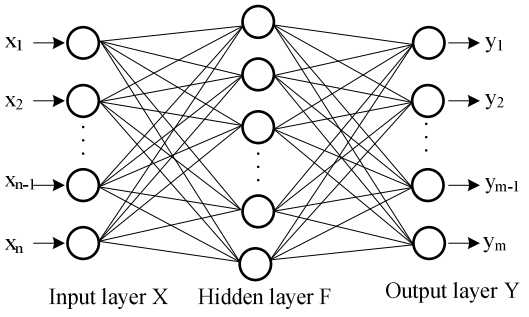


Fig. 2. Three-layer feed-forward neural network

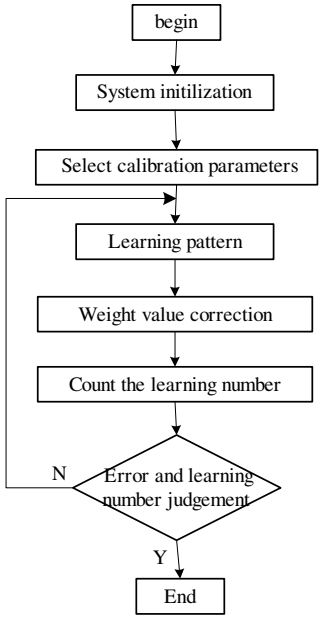


Fig. 3. The algorithm flowchart of the BP neural network

The number of the input layer neurons of the BP neural network could be determined based on the problems to be solved and the way of data representations. The number of the output layer neurons could be decided according to coding methods. Insufficient number of hidden units results that the network may not be able to meet the requirements of trainings and to identify samples, and the fault tolerance is poor. Too much hidden units will prolong the training time either. Hidden

units are set enough at the beginning, and then useless hidden units will be removed through the learning until no units can be cut out any more. The learning process is composed of two processes – the forward propagation of signals and the back propagation of errors. Through these two processes, weights are adjusted continuously until the output error drops in a range or the learning number reaches its predetermined value, as shown in figure 3.

5 Summary

The artificial network is used in achieving liquor identifications. The number of input layer neurons of the network is the number of sensors in the gas sensor array, and the number of output layer neurons of the network is the number of parameters to be identified in samples. Data acquired are used in trainings and tests of the system. Artificial identification system based on the gas sensor array and the feed-forward neural network can identify the main components of white wines and tell the quality evaluation and the category of the liquor.

References

1. Shi, Z., Li, J., Ma, Q., Cui, D., Zhu, M.: Electronic Nose and Its Application to Identification of Distilled Spirits. Instrument Technique and Sensor (1), 34–36 (2000)
2. Huang, Y.-C., Wang, W.-L., Liu, Q., Yan, S.-H.: Design of USB data storage system based on TMS320C6713. International Electronic Elements (10), 51–53 (2008)
3. Xiong, J., Wu, Z., Hu, M.-Q.: Design of multi - function power quality analyzer based on DSP TMS 320C6713. Electric Power Automation Equipment 26(5), 79–83 (2006)
4. He, G., Liu, M.: Gas Sensors and Array based on Electronic Nose. Sensor World (7), 6–9 (2008)
5. Li, J.: The Research on Electronic Nose in Liquor Identifications. Master's Thesis of Dalian University of Technology (2009)

A 3-d Slices Stereo Microscope Based on Electronic Image

Xuepeng Liu and Dongmei Zhao

Zhongshan Polytechnic, P.R. China
liuxuepeng1026@yahoo.com.cn, zdmeihn@163.com

Abstract. A 3- stereo slices microscope based on electronic image based on electronic image is designed, including process control, the stable bar, the feeding system, the control process flow, principle of work. Edge Algorithm and Binary Image Filtering Processing are discussed

Keywords: 3-d slices stereo microscope, 2-d slices image, process control, Micro feeding system, electronic image.

1 Introduction

In the biological medicine researchers often need to understand biology internal 3 d structure, such as the sample for a part of the internal space position, size, and shape and around and the spatial relationship of tissue. Biological structure of many of the features brings a lot of difficulties. Usually, most of the instrument of the style of observation still only stay in 2 d images, can't direct observation of its stereo image, often only with personal experience in the observation a lot of 2 d images in your mind up after reconstruction. This no doubt will influence the credibility of the final. 3 d reconstruction to the biology technology is the internal structure of the study provides great convenience; improve the accuracy and scientific research [1-4].

Continuous biopsy of the computer reconstruction is to research objects by continuous thin section, the use of photography system for each section 2 d image, and then call this series of 2 d image through the computer to carry on the processing and get the three-dimensional structure form. This method of characteristics, one of, can obtain two-dimensional fault, by CCD camera image resolution of the decision, but the present resolution CCD camera resolution has enough to satisfy the needs of the study. Second, the number of slice, the sampling period can be a lot more short, slice thickness can reach 1 u m, even can also make ultra-thin slices (thickness of the nanometer range) to restore microstructure. Thirdly, can study form smaller organisms.

2 3-d Microscope Reconstruction Design Scheme Based on Electronic Images

The whole shading and embedding good sample in spring and holding mechanism, fixed by the drives the intermittent up for linear motion, every time the size of the

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 389–393.
springerlink.com

displacement is thickness can be in control program slice of specified. Whenever the spindle a week to complete a cutting, sensor test to position signals to control card, and the program continuously scan control card each port to the position of the spindle, detection signal namely later captured picture to grab the current program called fault and stored in hard disk image, then to step motor stepping into a pulse control to capture picture, thus finish in a slice is a process of slice-catching-feeding.

When finished, need to style cutting computer signals and stop all operations. To achieve this goal, you can use a computer count * step into the distance to reach; In addition, because we make style in are in a certain size, therefore, can be in the chip with a process control on board. When the distance to when fixed in the slot through the switch, a high level assolenoid style, to end the signal operation.

Step motor due to feed into the distance or a long, will produce the instability, can be in the ball screw side with a stable bar, with a guide set of connection ball nuts and stable bar, and increase the stability.

3 Micro Feeding System

The ball screw and stepping motor driver is used. The ball screw is in the screw and nut put right amount, between the balls to make it by sliding friction into the rolling friction. The ball screw transmission device is composed of screw, nut, ball and ball cycle returning device.

The 3 d stereoscopic microscope slice of control is mainly to the stepping motor and drive and control of the CCD head, so it can in the spindle's gyration a cycle to complete the step motor respectively in and image collection. Therefore, we need to know the location of the spindle. We in the spindle and a position sensor, so that when the spindle's gyration to a certain position, which should not affect the normal knife row in and image acquisition, through the control of instruction, so that in the less than a rotary tool cycle of time, respectively in complete and image acquisition. The control process flowcharts, as shown in figure (1) shows:

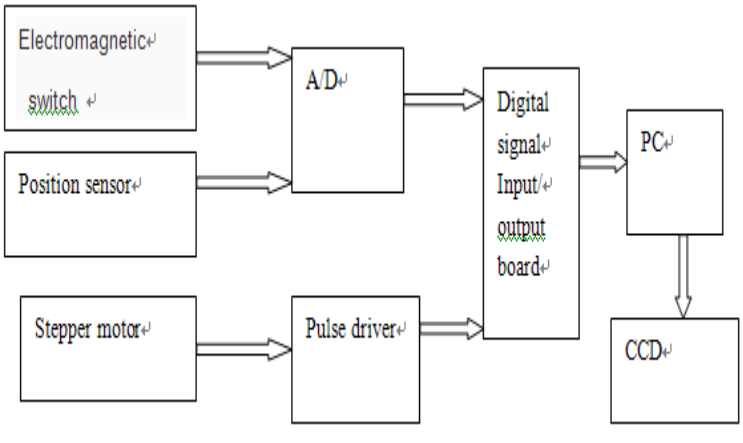


Fig. 1. Control system flow chart

The microscope control is mainly to the cloth into the motor and the CCD control. We adopt HY-6160 digital quantity signal input/output of 74 LS224 board. 4 a 48 BIT (32 channel) buffer type the digital quantity input channel, 4 piece of 74 LS273 constitute a 48 BIT (32 channel) hosting type digital output channel. The entire output channel reset when power is low level.

Interrupt signal to apply for board INT three states in the door input, when the interrupt signal for low electricity allowed at ordinary times, allowing external events make interrupt signal for. Through the hardware options, sure jump line interrupt level. As shown in the Fig.2 below.

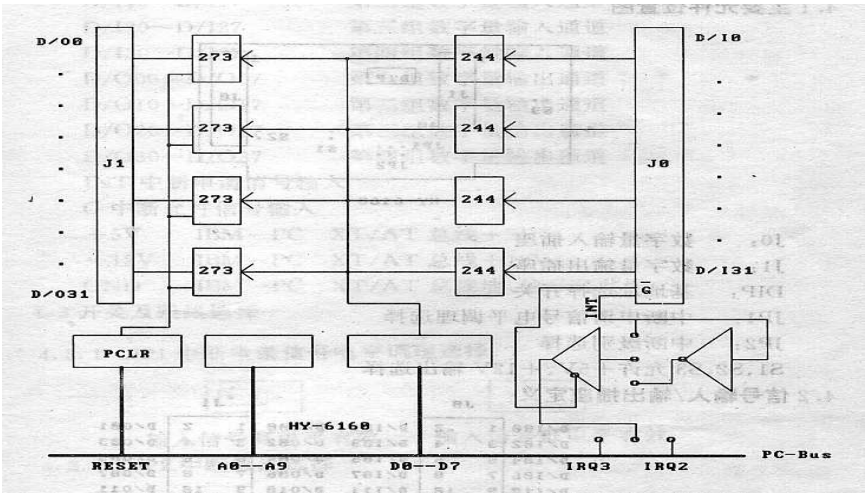


Fig. 2. Working principle figure

4 Edge Algorithm

Edge feature is an important characteristic. Most parts of information in image or signal lies in irregular structure and unstable situation. Edge detection technology enhances the contour processing algorithm to show the image. There are some difficulties in detection: one is the complexity of edge formation, the other is the choice between noise and edge.

There are some requests on edge detection algorithm:

[1], The edge detected by algorithm has much positioning precision.
[2], The response is single-pixel.
[3], The algorithm is suitable for different size edge and has no leakage.
[4], The algorithm is not sensitive to noise.
[5], The algorithm is not sensitive to edge direction.

Template matching method is used for edge detection algorithm to obtain satisfactory results. The following is the detailed Robinson algorithm:

The Robinson algorithm is made up of eight templates M. T is the A point's graying value. While J(i=0, 1, …, 7) is A point's graying value by (i+1) template processing. So

$$J_i = |M_i| \times |T|$$

The final value is:

$$J_A = \max(J_i)$$

It is concluded: when A is edge point, the graying values varies so much among the neighboring point of A, so J is the maximum value; when A is not edge point, the graying values varies so slightly among the neighboring point of A, J is nearly zero, which is much less than graying value at the edge point. So it is easy to find the image edge by Robinson algorithm and easy to mark the edge. When A is edge point, and T matches M_i, J is the maximum. So the value of point A is the maximum in the edge direction.

However, there are so many repetive calculations in Robinson algorithm, which adds up to processing time. Meanwhile, the edge direction is not needed when obtaining J. Therefore only four templates are needed for processing and to obtain absolute value. The enhanced algorithm is the following:

The one-dimension data group-D_{mxn} represents two-dimension image-MxN.

$$X_i = D_i + D_{i+1},\ Y_i = D_i + D_{i+n}, i = 0,1,2,\cdots m \times n$$

So

$$M_{oi} = X_{i-n-1} + Y_{i-n} - X_{i+n-1} - Y_{i+n}$$
$$M_{3i} = X_{i-n} + Y_{i-n+1} - X_{i+n-1} - Y_{i-1}$$

By comparing M0-M3, J is the maximum value. The test demonstrates that by the enhanced algorithm the calculation is 24% faster than that by original algorithm.

5 Binary Image Filtering Processing

The valve splitting processing is used to obtain clear edge. The valve method utilizes the difference between target and background to change the image into binary image, which is also called binary image splitting processing.

Opening binary makes image standing out of background smoother, while close binary makes background standing out of image smoother. So binary algorithm can eliminate the noise, which is the main idea of Binary morphological image edge detection method.

Supposing that R is the original factor, while S is structure. F is the result of binary morphological filtering.

$$F = [(A \circ B) \bullet B + (A \bullet B) \circ B] / 2$$

The circle structure is uses to binary filtering with consideration to isotropy. The circle diameter depends on noise spatial frequency. The bigger is the partial frequency, the less is the diameter.

6 Conclusion

This paper utilizes Robinson edge detection algorithm to process mover image, which gets satisfactory contour. However, original Robinson algorithm is complicated. The enhanced Robinson algorithm inhibits the noise, which is suitable for the edge solution. In case of circumstance, mover situation and image capture device effect, there exists interference. The system utilizes morphological theory to filter binary edge image. The filter image demonstrates that the noise which existed in binary edge image can be eliminated to obtain the accurate position.

References

[1] Jiang, Z.: 3-d imaging technology Based on optical microscope. Infrared and Laser Engineering 35(z4) (2006)
[2] Hao, L.: Design of a 3-D internal structure microscope. Optics and Precision Engineering 10(5) (2002)
[3] Yu, X.-M.: High-Voltage Electron Microscopy For Observing “True” Cellular Morphology. Acta Academiae Medicinae Qingdao Universitatis 39(2) (2003)
[4] Yang, G.: Observation of neuromuscular junction by serial electron microscopy and 3D-reconstruction. Journal of Chinese Electron Microscopy Society 23(2) (2004)

An Experimental Study on the Auto-association Neural Networks with Electronic Image Processing

Jin Zhou and Shunfu Gao

School of Electronics and Electrical Engineering,
Shanghai Second Polytechnic University Shanghai, P.R. China

Abstract. Based on the mechanism of auto-association, this paper proposed a new type of auto-association neural networks which can classify images and restore distorted images. Meticulous experiments were designed with Matlab language to test the networks' performance on image recognition and image restoration. In addition to random noise test, occlusion noises were included in the experiments too. The comparison of different number of recurrent layers settings in neural networks was discussed in details.

Keywords: auto-association neural network, image recognition, image restoration.

1 Introduction

Neural networks are massively parallel processors that have the ability to learn patterns through a training experience. Because of this feature, they are often well suited for pattern recognition. An auto-association network model is one where an input pattern is associated with itself[1-2]. Pattern association makes a connection between an input set of data to an expected output value. The idea behind this type of network is that part of input pattern can trigger the recognition of the whole pattern. Therefore, auto-association neural networks can also be used for distorted image restoration.

This paper proposed an experimental study on the auto-association neural networks. A recurrent layer with the same inputs and targets served as auto-associate memory. Then *hardlim* function[3] in Matlab was used in the following layer to separate target values into 0 and 1.

2 Experiment Design

Given 15 faces of a multimedia game (See Fig. 1), each of them was stored as 20×20 binary matrices (i.e. having values 1 and 0). The objectives of this experiment were to:

1. Design one or more networks capable of distinguishing them;
2. Recover distorted face images;
3. Evaluate the noise performance of the network(s).

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 395–401.
springerlink.com

Fig. 1. The 15 schematic faces

Note that in this paper, the 15 faces were always sequenced as:

1 2 3 4
5 6 7 8
9 10 11 12
13 14 15

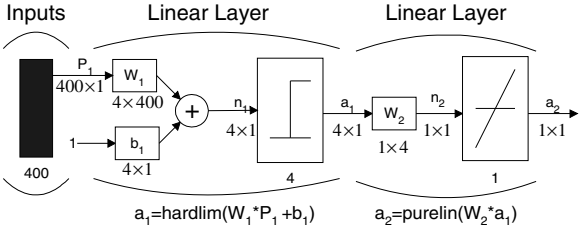


Fig. 2. Basic network to distinguish 15 faces

Firstly we introduced a basic classification network for comparison which is shown in Fig. 2. And target matrix T of the first layer was chosen as: T= $\begin{bmatrix} 0&0&0&0&0&0&0&1&1&1&1&1&1&1&1 \\ 0&0&0&1&1&1&1&0&0&0&0&1&1&1&1 \\ 0&1&1&0&0&1&1&0&0&1&1&0&0&1&1 \\ 1&0&1&0&1&0&1&0&1&0&1&0&1&0&1 \end{bmatrix}$. In order to minimize overall error $\left\|T - W_1 P_1\right\|^2$, pseudoinverse rule was implemented in first layer with $W_1=TP^+$, $P^+=(P^TP_1)^{-1}P_1T$. The output of first layer was a 4-digit binary code of number 1~ 15,

for example, output 0111 represented the Face 7. Second layer worked to decode the outputs of first layer into decimals with W_2=[8, 4, 2, 1] and b_2=0. b_1=[-0.5, -0.5, -0.5, -0.5] is the best choice to separate target values into 0 and 1. The reason was that float computations in Matlab only generates approximate zeros, and this tiny approximation would make a big difference when it is an input of *hardlim* tranfer function.

This basic network worked successfully to identify all 15 faces images. However, it was sensitive to noise contamination. We modeled random noise as randomly choosing points in a face image and changing their color to its opposite, i.e. from white to black or black to white. In our noise test, we found that the basic network could only recognize faces 100 percent correct with noise level less than 10%.

In order to improve the network's noise performance, the network shown in Fig. 3 was developed, in which one recurrent layer (the first layer) was added before the basic network we developed in previous section. Pseudoinverse rule was applied in the first layer again. But this time we used auto-association memory to store input patterns, so the weights and biases in the network were initialized as $W_1=P_1P^+$, $P^+=(P_1TP_1)^{-1}P_1T$, $W_2=TP^+$, W_3=[8, 4, 2, 1], b_2=[-0.5, -0.5, -0.5, -0.5], b_1 and b_3 are zero vectors.

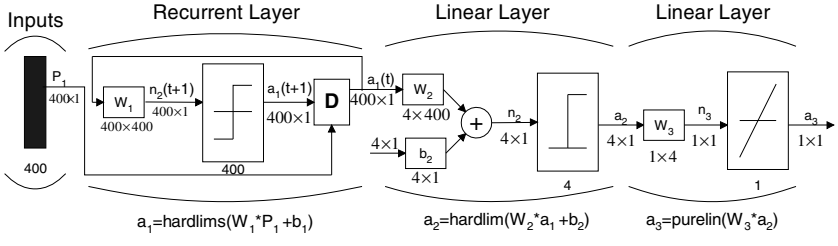


Fig. 3. Improved network to recover distorted faces

By controlling the network structure and times of recurrences, four cases were developed for comparison:

(i) Basic network, no auto-association memory, no recurrence
(ii) Auto-association memory, no recurrence (time of recurrence = 0)
(iii) Auto-association memory, one recurrence
(iv) Auto-association memory, two recurrences

Fig. 4 showed an example result, where the settings were auto-association memory, no recurrence network at noise level 50%. Though the input almost could not be recognized by naked eyes, in the final output, except Faces 3, 7, and 8, all the rest were correctly classified. Apparently adding auto-association memory can restore noisy images quite well. But could recurrence layer works better? If so, how many times of recurrence are needed for this 15 faces problem?

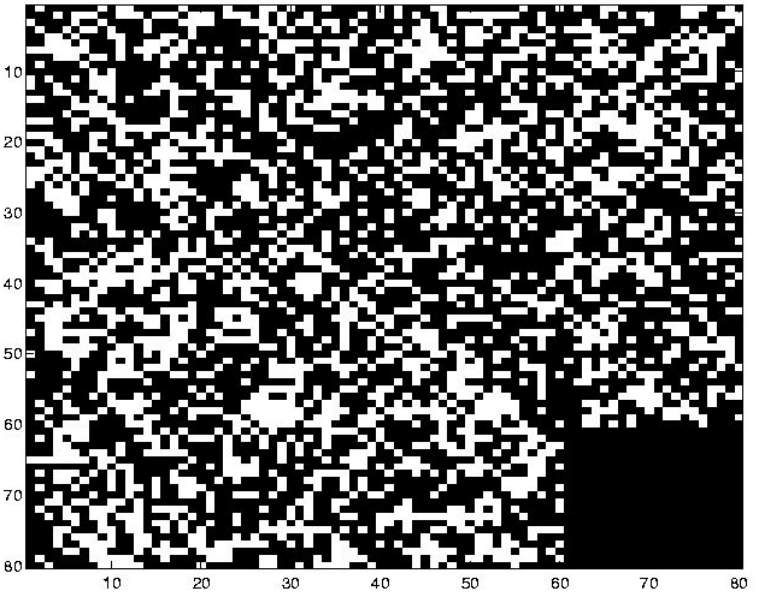


Fig. 4(a). Input with random noise level = 50%

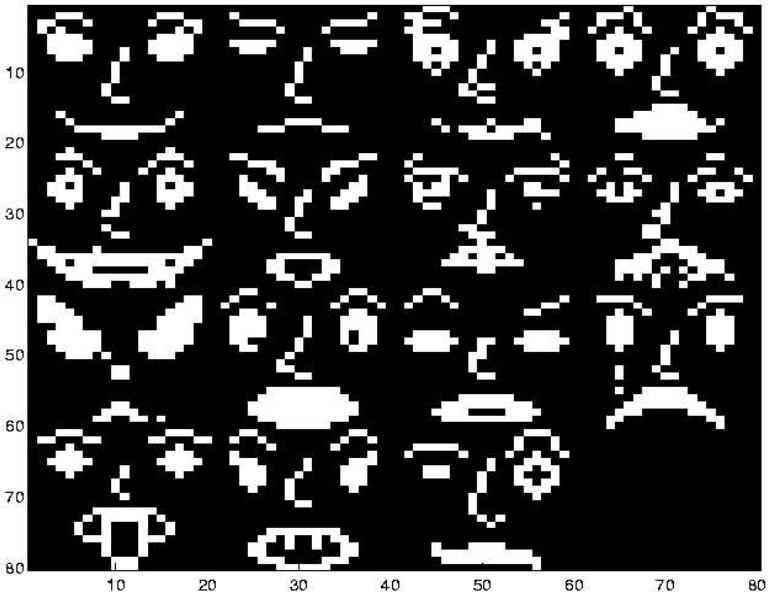


Fig. 4(b). Output of recurrent layer with input shown in Fig. 4(a). (Noise Level=50%)

3 Experiment Result

3.1 Random Noise Experiment

Keep changing noise level, average error rates of 20 runs (20*15=300 faces) under the same scenarios were shown in Fig. 5. In this paper, error rates were calculated by the number of wrong classifications divided by the number of face inputs.

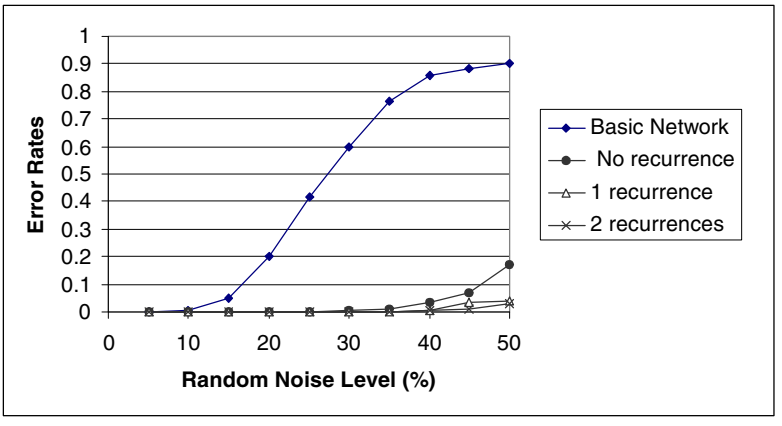


Fig. 5. Average error rates of different networks

As the results showed, adding auto-associate memory layer to the basic network lowered error rates greatly. 100 percent correctness even occurred at noise level 25%. Also, although the basic network almost could not recognize any face at noise level 50%., with auto-associate memory, half of the faces could be recognized correctly.

Furthermore, with more recurrences, we could restore the faces even more successfully. In case (iii), (iv), 100 percent correctness occurred at noise level 35%. And at noise level 50%, the error rate was even lowered to 0.03, this means only 9 out of 300 faces were wrongly classified. Since the error rates in Case (iv) were the lowest, network in Case (iv) was suggested to be the best one out of the 4 cases. In next experiment, we will use it to deal with occlusion noise.

3.2 Occlusion Noise Experiment

In this experiment, we tried to find out the capability of the improved network restoring distorted images from occlusion noise (Here, the network we used follows Case (iv) in the previous experiment). Occlusion noise could be generated from four directions: left, right, upper, and lower. Again, keep changing noise level, average error rates of 20 runs (20*15=300 faces) under the same scenarios were shown in Fig. 6.

Note that large increases in error rates always correspond to that the informative portions of face images were being cut off. It could also be observed that with noise level less than 50%, the network performs much better in cases of left and right occlusions than upper or lower occlusions. The reason was that most face images were horizontally symmetric, so even almost one half was occluded, the whole face could be restored by the rest information. An extreme case was although faces were occluded nearly left halves (45%), the network we developed was able to recognize all 15 faces. However, even small upper portions (15%) were cut off, error occurred.

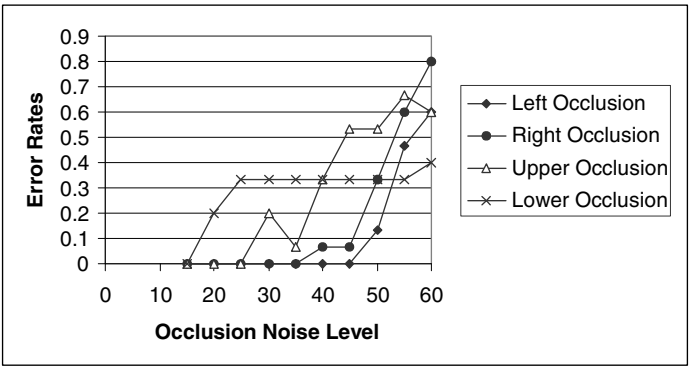


Fig. 6. Average error rates caused by different occlusion noises

4 Conclusion

Based on the experiment result, following conclusions are made:

1. Even a basic one-layer network is able to identify 15 different faces. But the basic network is sensitive to noise contamination, 100 percent correctness only happens at random noise level less than 10%.
2. In most cases, non-zero biases are necessary in practical network realization.
3. After adding auto-association memory, the network's noise performance is improved greatly. The network can overcome successfully random noise at noise level 25%.
4. Recurrence layer can improve noise performance even further. 100 percent correctness even occurred at random noise level 35%. And the error rate could be lowered to 0.03 at random noise level 50%.
5. The network we developed is also capable of restoring occluded images. However the capability depends on which part of the images is most informative, and whether the images are symmetric.

Acknowledgement. This study is sponsored by the Leading Academic Discipline Project of Shanghai Municipal Education Commission (No. J51801).

References

[1] Hagan, M.T., Demuth, H.B., Beale, M.H.: Neural Network Design. China Machine Press, Beijing (2006)
[2] Hou, Xuan: Induction and Application of Supervised Association Artificial Neural Network Optimization Algorithms. International Electronic Elements 16(8), 6–8 (2008)
[3] Chang, H., Yuan, G., Chang, M.: Simulation Software Textbook-Multisim and MATLAB. Tsinghua University Press, Beijing (2006)

The Research of Digital Color Image Quality Metrics

Xiangyang Xu*, Qiao Chen, and Yuanhong Zhu

Media & Communication School, ShenZhen Polytechnic, ShenZhen, China 518055
{xxy0906,qiaochen,zyh551111}@szpt.edu.cn

Abstract. One goal of image quality metrics is to predict human judgments of the difference between image pairs. The analysis of color image quality usually includes three aspects: color difference, sharpness, contrast. The calculating of color difference has been put forward, S-CIELAB has already been used to calculate color difference. In order to measure the sharpness difference and contrast difference of digital color image pairs, a new framework of color image quality metrics is constructed on the basis of CIE color difference metrics. In this research, three arguments of color images can be calculated to evaluate the difference between digital color image pairs.

Keywords: Image quality, Color image differences, Image contrast, Vision modeling.

1 Introduction

With the development of network technology and new media technology, the digital image is closely linked to our life. Not only the higher requirement is required to process and display of digital images, but also the evaluation of image quality has became a hot topic to measure image difference between the disposed images and original images. For example, to quantify the image difference between the compressed image and the original image, to determine the image difference before and after the gamut mapping algorithm etc. At present there is not an accurate and complete mathematical model, which can predict the perceptual image difference of human visual system. There are only a few simple models to use to quantitative analysis, and using the predetermined scale implement visual assessment by some observers to image pairs. It is lack of objectivity, and very tedious and expensive.

Recently some objective evaluation methods are usually used to measure digital image quality, such as the root mean square error (RMSE) and peak signal-to-noise (PSNR). Without considering the human visual characteristic, the result is not according to the subjective feelings of the human eyes [1-2]. So several new methods [3-9] were come up with by some reaserchers referring to the psychophysical experiment results which are related to human visual characteristic, and they greatly improved prediction accuracy. Most of the digital image distortion metrics in the literature apply to grayscale images, only a few studies address color image

* Xili Lake, Nanshan District, Shenzhen, Guangdong Province, P.R. China 518055
Xiangyang Xu (xxy0906@szpt.edu.cn)

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 403–407.
springerlink.com

difference, the most representative is the S-CIELAB[10], which can evaluate the image difference between image pairs. In a sense the S-CIELAB can only reflect the image's chromaticity properties, but the characteristic of image contrast and sharpness property cannot be reflected. In the study, we will extend the image difference framework of S-CIELAB to compute the contrast and sharpness difference between image pairs.

2 The Digital Color Image Quality Metrics

The color image includes three characteristics: chroma, sharpness and contrast. Judging the difference between image pairs, the three aspects should be included. The chroma attribute consists of lightness, saturation and hue. The sharpness represents the image edge and texture. The contrast represents the distribution of image tones. In order to calculate the difference of image pairs, we have proposed a metric of image quality based on S-CIELAB. Simultaneously we have proposed two algorithms of the correlation analysis of image sharpness difference and the loss degree of image contrast.

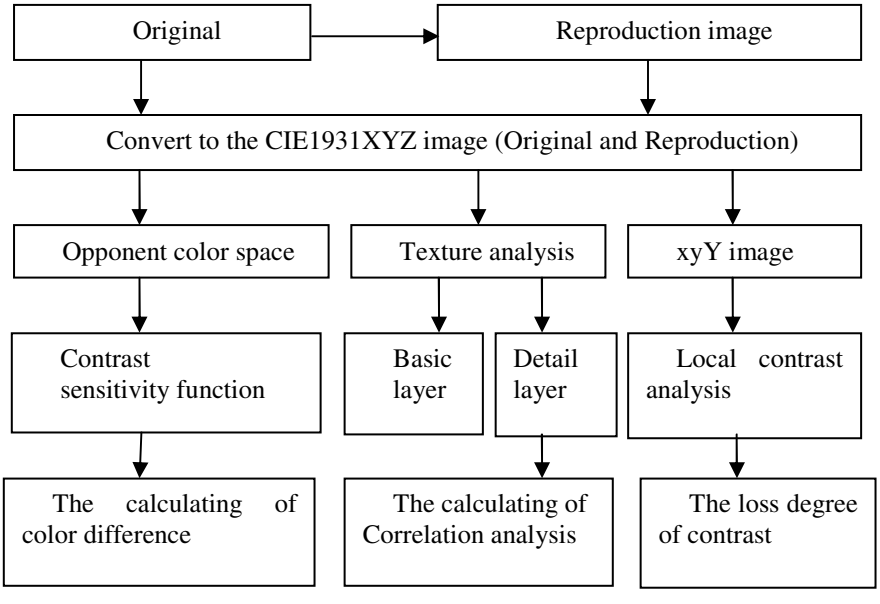


Fig. 1. The digital color image quality metric

The image difference is closely related to the visual difference. The ultimate goal of the image color difference reaserch is expecting the calculating result in accordance with the human eye perception difference. To resolve this fundamental issue, we have to establish the visual difference model of the complex color stimuli based on plenty of visual physiological and psychological experiment results. In our model, we used

the S-CIELAB model to calculate the color difference of image pairs, because the contrast sensitivity function of S-CIELAB was established according to the law of the human vision.

The parameter of image sharpness represents the difference of image edge and image texture. Firstly we obtained the texture image by the principle of gray-scale image filtering. The original and reproduced images are divided into the two images: one including all details is called detail layer, the other including basic information is called basic layer. To calculate the difference of image sharpness, we only analysis the correlation of the detail layer between image pairs. The correlation analysis adopted the theory of covariance. The covariance is defined as the average of the sum of the product of the dual variable deviation from mean value. Covariance is a coefficient to measure the collaborative variation between two variances, which represents the influence between them: the larger the covariance value, the more influence between two variances. For the images before and after processing, the changes of image sharpness are two correlation variances, so the covariance can be used to represent their correlation, and the larger the value means the closer of two image's sharpness. The covariance is calculated by equation (1).

$$\mathrm{cov}(X,Y)=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y}) \tag{1}$$

Contrast is the parameter that represents the distribution of image tones. It is well known that we can obtain the distribution of image tones by image gradient analysis. Comparing to the contrast difference between the processed image and the original, it is equivalent to the statistics analysis of the D-value of the image pixel in different regions. First we calculate the image gradient of the image pairs respectively, then calculate the relative values of the two gradient images. This reflect the loss degree between the processed image and the original image. The larger the value, the greater the contrast difference between images pairs. The algorithm is to find the mean gradient value of a center pixel with its neighbor pixels, which represents the local contrast. Therefore we can calculate the ratio of the absolute gradient value, then we obtain the mean value of the ratios. This is used as the loss degree between image pairs. The loss degree of contrast is calculated by equation (2).

$$con=\frac{1}{n}\sum\frac{\sum_{p,q\in[1,8]}Y'_{(p)}-Y'_{(q)}}{\sum_{p,q\in[1,8]}Y_{(p)}-Y_{(q)}} \tag{2}$$

3 Experience

We have selected an image from the ISO standard test image datasets as the original one. Then we obtained five calculating images by using different JPEG compression ratio. The five images were given number 1 2 3 4 5 according to compression ratios in descending order. The visual difference is obvious between the compressed images and original image. Therefore those images are served as test data sets.

3.1 The Calculating of Color Difference

The color difference of S-CIELAB between the five compressd images and the original are shown in Fig 2.

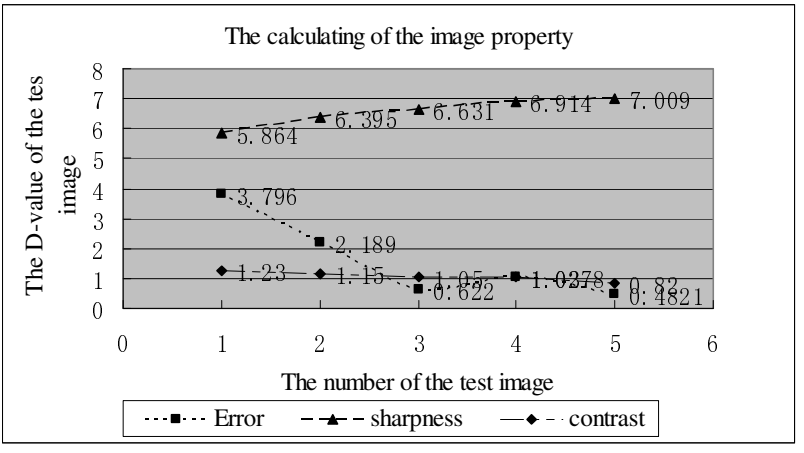


Fig. 2. The horizontal axis represents the number of five images respectively, while the vertical axis represents the D-value of the test image corresponding attribute. From the Fig2, the trend of "Error" curve is down, this explain that the color difference of S-CIELAB is consistent with the image quality.

3.2 The Correlation Analysis and Calculating

Firstly we execute the texture analysis to the test images. After extracting all images' detail layer, we perform the calculation of the correlation analysis, the results are shown in Figure 2. The trend of the "sharpness" curve in Figure 2 is up, which shows that the fifth image is more closely to the original image, because it's correlation parameters is the maximum. The trend of the sharpness curve have confirmed that the correlation analysis algorithm is in line with the objective facts, it can be applied to the sharpness measurement between image pairs.

3.3 The Calculating of the Loss Degree of Image Contrast

The results of calculating the gradient loss degree between the five processed images and the original image are shown in Figure 2. The trend of the contrast curve in Figure 2 is down, which shows that the fifth image' contrast loss is the minimum, it is the most closely to the original image. The first image' contrast loss degree is the maximum, it is the largest difference comparing with the original one. The contrast curve trend has confirmed that the algorithm of contrast loss degree is consistent with objective facts, it can be applied to the measurement of contrast loss between image pairs.

4 Conclusion

The experimental results have confirmed that the three measurement algorithms in the image quality metric accord with objective fact. On the whole, those curves' trend are consistent with the objective judgment of human eyes. The model can be applied to the metric of the image color difference, the sharpness, and the contrast between image pairs. But for one component, to ensure the consistency between the calculating value and the visual perception difference in magnitude, further researches are needed to establish the visual difference model of the image attribute metric.

Acknowledgments. This research work was supported by the National Natural Science Foundation of China (61108087). The study of spectral reflectance model in the area of cross-media color reproduction).

References

1. Eckert, M.P., Bradley, A.P.: Perceptual quality metrics applied to still image compression. Signal Process. 70(3), 177–200 (1998)
2. Wang, Z., Bovik, A.C., Lu, L.: Why is image quality assessment so difficult? In: Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing, Orlando (2002)
3. Bradley, A.P.: A wavelet visible difference predictor. IEEE Transactions on Image Processing 5(8), 717–730 (1999)
4. Heeger, D.J., Teo, P.C.: A model of perceptual image fidelity. In: Proc. IEEE Int. Conf. Image Proc., pp. 343–345 (1995)
5. Li, B., Meyer, G.W., Klassen, R.V.: A comparison of two image quality models. SPIE Coni on Human Vision and Electronic Imaging 3299(2), 98–109 (1998)
6. Wang, Z., Sheikh, H.R., Bovik, A.C.: Objective video quality assessment. In: Furht, B., Marques, O. (eds.) The Handbook of Video Databases: Design and Applications. CRC Press (2003)
7. Daly, S.: The visible difference predictor: An algorithm for the assessment of image fidelity. In: Proc. SPIE, vol. 1616, pp. 2–15 (1992)
8. Lubin, J.: A visual discrimination mode for image system design and evaluation. In: Peli, E. (ed.) Visual Models for Target Detection and Recognition, pp. 207–220. World Scientific, Singapore (1995)
9. Chen, M.J., Bovik, A.C.: No-reference image blur assessment using multiscale gradient. EURASIP Journal on Image and Video Processing (2011)
10. Zhang, X.M., Wandell, B.A.: A Spatical Extension to Cielab for Digital Color Image Reproduction. In: Society of Information Display Symposium Technical Digest, vol. 27, pp. 731–734 (1996)

The Research on New Efficient Image Editing Method Based on Texture Regularity and Texture Synthesis

Qiu Yong and Gan Gang

Xichang College

Abstract. This paper proposes a new method of texture regularity analysis and texture synthesis techniques for efficient image fill. The method can automatically extract the image foreground region and a reasonable recovery of the original content of the occluded region. For foreground extraction, the general practice in order to get accurate results often requires the introduction of interoperability. In order to automate the extraction, this paper focuses consider the prospects for the approximation the rules texture (NRT) of the scene region. Approximate rules texture has a strong signal to the rules, the use of analysis of this rule can be effective and accurate segmentation of the foreground. Recovery issues for background and texture synthesis, we proposed a new background to fill the algorithm, and by the introduction of synthetic priority concept of image restoration algorithm to come. Therefore, this method has the advantages of the two methods.

Keywords: Image filled, texture regularity, texture synthesis, approximate rules texture, image rendering, foreground segmentation, fill.

1 Introduction

The images filling the problem is an important problem in computer graphics and image processing. This problem usually consists of two steps, fill the foreground extraction and content. About the prospects for extraction, has been a lot of interaction based method proposed. But it also means a fairly elaborate and tedious work. The automatic extraction methods are suitable for all occasions do not seem likely. Therefore, our work focuses on a class of prospects content for the approximation the rules texture of the scene [1], [3]. Approximate rules texture just a special kind of texture, but it can establish a model for a lot of texture in the real world. This article extends the James Hays, and other work [1] proposed an automatic detection algorithm to extract the target element. Image fill studies usually have two types of methods: texture synthesis algorithm [4-10] and image rendering technology [2], [11]. Texture synthesis technique is suitable for the area that used to fill the static texture regional and structural texture. Image rendering technology is commonly used in the restoration and repair of small and medium-sized defects of the image, such as cracks, blemishes and so on. Our approach attempts to combine the two together, both to ensure the efficiency and effectiveness of the synthesis algorithm based on the sample, taking into full account the structural information contained in the image.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 409–412.
springerlink.com

The remainder of this article is organized as follows. Statements related to the work of the Section 2, 3 and 4 gives the algorithm for basic theory and algorithm implementation details, section 5 shows the test results, and finally in Section 6, we give summary.

2 Related Work

Related to the problem of image fill generally have two aspects: First, the texture regularity analysis, texture synthesis algorithm.

Texture regularity analysis is a very difficult and very important issue. In addition to the interactive method [3], automated analysis is always the focus of the study and difficult. Lobay and Forsyth, [1] to solve this problem through the establishment of a global pattern prime structure. The above method is different; Schaffalitzky et al. [1] Turning et al. [1] assume that texture completely rule and established a very strong conversion model in a plane. After considering the advantages and disadvantages of the above methods, James et al. [1] proposed a high-level feature matching algorithm. The advantage of this algorithm lies in its use of the basic structure of the texture, but not too many restrictions on allowable deformation.

Texture synthesis method can be divided into two categories, one is the process of texture synthesis, and the other is based on the texture of the samples synthesized. While the latter can avoid texture mapping and procedural texture synthesis method defects become more popular. Drawing texture synthesis method is a method of Efros and Leung [4] proposed in 1999, non-parametric feature matching on Markov Random Field (MRF) model. Wei and Levoy [5] extend the work of Efros / Leung, based pyramid [7], [8] synthesis methods greatly accelerate texture synthesis rate., Ashiknmin stronger coherence search [9], its significance is to reduce the search space and access to interactive rate. In 2001, Efros [6] proposed a block-based texture synthesis method, while dramatically speeding up the synthesis rate to maintain good results.

3 Rule Model of the Three Textures and the Texture Synthesis with Priority

3.1 Approximate Rules Texture Analysis Model

Approximate rules texture is defined as a strict rule texture after a slight transformation of the geometry, light, or color texture [3].: $p = d(p_r)$, Denote $d = d_{color} \times d_{light} \times d_{geo}$ the color, light, geometric transformations, and combinations thereof, transform, corresponding to the strict rules texture. Plane mosaic theory [12], the translational symmetry, (t_1, t_2) texture generated by a pair of the shortest vector. The trajectory of the vector form of two-dimensional quadrilateral mesh, this quadrilateral mesh defines the topology between all the 2D pattern elements and pattern factors. Similar approximate rules texture it corresponds to the (t_1, t_2) strict rules texture generated vector representation. Based on the above theory, texture analysis

model can be obtained by repeating the following four steps: raised pattern prime distribution grid, to improve the grid and thin plane spine convolution. Based on the work of James et al, the algorithm of this paper focuses on approximate rules texture as the research object. We apply a clustering algorithm, the texture synthesis algorithm and the vector mask to modify the automatic mask extraction can be completely and accurately cover the target area. The algorithm details will be detailed in the fourth quarter.

3.2 With the Priority of the Texture Synthesis

In this paper, the pyramid-based multi-resolution texture synthesis algorithm [7], [8]. At the same time, taking into account the characteristics of the problem itself, we have introduced a priority calculation in order to better maintain the structure of texture synthesis algorithm.

The texture synthesis model is: $I_{syn} = \sum_i M\left[P\left(\underset{d(I_i, I_{syn})<\zeta}{S}(I_i)\right)\right]$ where is I_{syn} the need for synthetic images, M the boundary fusion operator, P is to reset the operator, I_i the sample texture in accordance with a certain setting to remove the basic synthesis unit, S the search operator d is the image distance ζ is visual similarity threshold. You can see there are four factors can be adjusted: the basic synthesis unit, the search rules, restructuring, and boundary integration method in this framework. The synthetic sequence is the key of the texture pass structural information. The general synthesis algorithm is synthesized in accordance with the order of the scan line; this does not get the best results in our problem. By Criminisi [2] inspired by the work given different priority for each candidate block, we can easily control more suitable for the synthesis of the order filling problems.

As mentioned in [2], high priority should be given to the strong continuity on the boundary, and surrounded by high degree of confidence pixel candidate block. The priority calculation defined as: $P(p) = C(p)D(p)$ on behalf $C(p)$ of confidence, $D(p)$ referred to as data items, used to describe the point on the boundary continuity.

$C(p)$ And $D(p)$ defined as follows: $C(p) = \frac{\sum_{q \in \psi_p \cap (I-\Omega)} C(p)}{|\psi_p|}$ $D(p) = \frac{\left|\nabla I_p^{\perp} \bullet n_p\right|}{\alpha}$ where is the area $|\psi_p|$ is ψ_p a normalization factor, α is a point orthogonal to the boundary of the unit orthogonal vectors, n_p pointing out p that the $\delta\Omega$ orthogonal operation. We calculate the priority for each boundary candidate $P(p)$.

4 Conclusions

This paper presents a simple and effective image fill, automatic mask extraction algorithm and priority-based synthesis algorithm. We selected the target object contains the approximate rules texture structure information. The main contribution of

this paper is two things: for the automatic extraction of the foreground object approximate rules texture of type and priority-based texture synthesis to fill. Experiments show that this algorithm can effectively and automatically extract the corresponding mask on the target object. Into the priority computation can effectively be passed to the linear structure and the 2D texture target area in the pyramid-based multi-resolution texture synthesis algorithm. In addition, the size of the sampling window and the search range is very important. Experiments show that the window parameters used in this article should be relatively small, is proportional to the size of the target object.

References

1. Hays, J., Leordeanu, M., Efros, A.A., Liu, Y.: Discovering Texture Regularity as a Higher-Order Correspondence Problem. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3952, pp. 522–535. Springer, Heidelberg (2006)
2. Criminisi, A., Pèrez, P., Toyama, K.: Region filling and object removal by exemplar-based image inpainting. IEEE Transactions on Image Processing 13, 1200 (2004)
3. Liu, Y., Lin, W.C., Hays, J.: Near-regular texture analysis and manipulation. ACM Transactions on Graphics (SIGGRAPH) 23, 3 (2004)
4. Efros, A., Leung, T.: Texture synthesis by non-parametric sampling. In: Proc. Int. Conf. Computer Vision, Kerkyra, Greece, pp. 1033–1038 (1999)
5. Wey, L.W., Levoy, M.: Fast texture synthesis using tree-structured vector quantization. In: Proc. ACM Conf. Comp. Graphics (SIGGRAPH), pp. 479–488 (2000)
6. Efros, A., Freeman, W.T.: Image quilting for texture synthesis and transfer. In: Proc. ACM Conf. Comp. Graphics (SIGGRAPH), Eugene Fiume, pp. 341–346 (2001)
7. de Bonet, J.S.: Multiresolution sampling procedure for analysis and synthesis of texture images. In: Proc. ACM Conf. Comp. Graphics (SIGGRAPH), pp. 361–368 (1997)
8. Heeger, D.J., Bergen, J.R.: Pyramid-based texture analysis / synthesis. In: Proc. ACM Conf. Comp. Graphics (SIGGRAPH), Los Angeles, CA, pp. 229–233 (1995)
9. Ashikhmin, M.: Synthesizing natural textures. In: Proc. ACM Symposium on Interactive 3D Graphics, pp. 217–226. Research Triangle Park, NC (2001)
10. Liang, L., Liu, C., Xu, Y.-Q., Guo, B., Shum, H.-Y.: Real-time texture synthesis by patch-based sampling. ACM Transactions on Graphics 20, 127 (2001)
11. Bertalmio, M., Sapiro, G., Caselles, V., Ballester, C.: Image inpainting. In: Proc. ACM Conf. Comp. Graphics (SIGGRAPH), New Orleans, LU, pp. 417–424 (2000)
12. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. IEEE Transactions on Pattern Analysis and Machine Intelligence 24, 509 (2001)

Study on Passive Tracking Based on Electronic Chart Information

Jian Kang, Guosheng Rui, and Shicai Yu

Naval Aeronautical and Astronautical University, Yantai, 264001, P.R. China

Abstract. To avoid the huge computation burden and poor tracking accuracy of the particle filters in maneuvering target tracking, Particle Filter based on electronic chart information (ECI-PF) is introduced in this paper. According to the prior information of nearby island, when the target boat is closed to the island, the distribution, the number and the weight of particle clouds are adjusted properly, unwanted particles are discarded. Simulation results show that with the participant particles decrease, the computational loads are descended, the algorithm can be implemented in real time.

Keywords: electronic chart information, maneuvering target tracking, particle filter.

1 Introduction

Because of the characteristics of little observation equipments and superior concealment, bearing-only passive location and tracking problems of single observer have increasingly attracted comprehensive attentions. But the non-linear models in state equation and non-Gaussian noise assumption may often lead to poor performance in the traditional Kalman filter (KF) and its various modified algorithms. Particle filter (PF) can be applied robustly to any nonlinear/non-Gaussian situations. To avoid the degeneracy phenomenon of the particle filters, it is often adopted a lot of particle numbers in the iterative procedure, so that generates huge computation burden and poor tracking accuracy. Particle filter based on electronic chart information (ECI-PF) is introduced in this paper to solve above questions. With the prior geographic information of nearby island, when the target boat is closed to the island, the distribution, the number and the weight of particle clouds are adjusted properly, unwanted particles are discarded, so that effective particle numbers can be reduced. We compare this constrained filter with the standard particle filter in simulations, the result shows that the incorporation of the prior information can effectively reduce computation burden and improve the accuracy of the bearing-only maneuvering target tracking.

2 System Model

The location and tracking situation based on electronic chart for bearing-only problem is shown as Fig.1.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 413–418.
springerlink.com

Fig. 1. Location and tracking situation based on electronic chart

In Cartesian position coordinates, we assume $\mathbf{X}_T(k)=[x_T(k),y_T(k),v_{Tx}(k),v_{Ty}(k)]^T$ to represent the target boat state vector, where $x_T(k)$, $y_T(k)$, $\boldsymbol{v_{Tx}(k)}$ and $v_{Ty}(k)$ is the position and the velocity of target boat at x and y axis. $\mathbf{X}_O(k)=[x_O(k),y_O(k),v_{Ox}(k),v_{Oy}(k)]^T$ represents the tracking boat state vector, where $x_O(k)$, $y_O(k)$, $\boldsymbol{v_{ox}(k)}$ and $\boldsymbol{v_{oy}(k)}$ is the position and the velocity of tracking boat at x and y axis.

The state equation is given by

$$\mathbf{X}_{k+1}=\mathbf{F}_k\mathbf{X}_k+\mathbf{v}_k=\begin{bmatrix}1&0&T&0\\0&1&0&T\\0&0&1&0\\0&0&0&1\end{bmatrix}\mathbf{X}_k+\mathbf{v}_k \tag{1}$$

Where $\mathbf{X}_k=\mathbf{X}_T(k)-\mathbf{X}_O(k)=\left[x_k,y_k,v_{xk},v_{yk}\right]^T$, T is the sampling period.

The measurement equation is modeled as

$$\mathbf{Z}_k=h(\mathbf{X}_k)+\mathbf{w}_k=\tan^{-1}(y_k/x_k)+\mathbf{w}_k \tag{2}$$

Where $\mathbf{v}_k,\mathbf{w}_k$ is zero mean independent noise with covariance $\boldsymbol{Q_k}$ and R_k.

3 Particle Filter Based on Electronic Chart Information (ECI-PF)

Particle filters (PF) perform sequential Monte Carlo estimation based on point mass (or "particle") representation of probability densities. Let x_k denote the state of the

observed system and $Y_k = \{y_i\}_{i=1}^{k}$ be the set of observed measurements. The Monte Carlo filter approximates the density $p(x_k|Y_k)$ by a large set of N particles $\{x_k^i, w_k^i\}_{i=1}^{N}$, where w_k^i are the weights sum to unity, i.e., $\sum_{i=1}^{N} w_k^i = 1$. The location and weight of each particle reflect the value of the density in the region of the state space, and the joint posterior density can be approximated as follows:

$$\hat{p}(x_k / Y_k) = \frac{1}{N}\sum_{i=1}^{N} \delta\ (x_k - x_k^i) \tag{3}$$

The particle filter updates the particle and corresponding weights recursively with each new observed measurement. From above, we can see the particle filter is not restricted by the linear and Gaussian assumption, it often follows next steps:

① Initialize the particles and the importance weights: generate N samples $\{x_0^i\}_{i=1}^{N}$ from the initial distribution $p(x_0)$ and assign the respective importance weights $\tilde{w}_0^i = 1/N$, $i = 1,...N.$

② Propagate: set $k = 1$ and predict new particles $\{x_{k|k-1}^i\}_{i=1}^{N}$, i.e., $x_{k+1}^i = f(x_k^i, v_k)$, $i = 1,..N.$

③ Measurement update: calculate the acceptance probability $\tilde{w}_k^i$.

④Resampling: Generate a new set $\{x_{k|k}^i\}_{i=1}^{N}$ according to $Pr(x_{k|k}^i = x_{k|k-1}^j) = \tilde{w}_k^j$, $i = 1,..N.$

⑤ Increase k and iterate to step 2.

The block diagram of particle filters is shown as Fig.2.

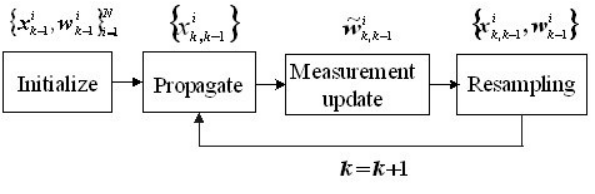


Fig. 2. The block diagram of particle filters

In practical maneuvering target tracking on the sea, we can easily achieve the information of nearby land or island by electronic chart. Namely, we can constrain the trajectory region of the target boat on some conditions, so that when the target boat move to or close to the land or island, the particle clouds is adjusted, avoid the target boat is beyond the island.

Assume the land or island region is similar to rectangle shape, the border for X axis is $\boldsymbol{D}_1 = [(\boldsymbol{x}_1, \boldsymbol{x}_2)]$, the border for Y axis is $\boldsymbol{D}_2 = [(\boldsymbol{y}_1, \boldsymbol{y}_2)]$, here the border coordinate can achieved recur to electronic chart. When the different distance $\boldsymbol{d}_i (\boldsymbol{i} = 1,2,3,4)$ between the island and the target boat is smaller than the bound distance $\boldsymbol{S}$, add particle selection procedure to the common particle filter, discard the particles out of constrain conditions, set their weight zero and readjust the distribution, the number and the weight of particle clouds properly.

The particle filter based on electronic chart information is shown as Fig.3.

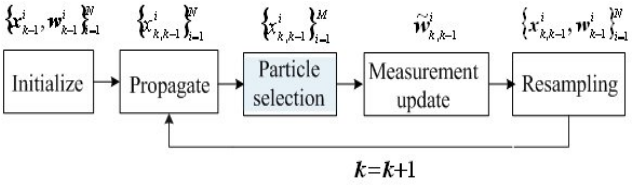


Fig. 3. The block diagram of ECI-PF

The key of particle selection is to confirm the constrained conditions. We can achieve the distance between the island and the target boat in all directions:

$$\begin{aligned} \boldsymbol{d}_1 &= \sqrt{(\hat{\boldsymbol{x}}_T - \boldsymbol{x}_1)^2 + (\hat{\boldsymbol{y}}_T - \boldsymbol{y}_1)^2} \\ \boldsymbol{d}_2 &= \sqrt{(\hat{\boldsymbol{x}}_T - \boldsymbol{x}_2)^2 + (\hat{\boldsymbol{y}}_T - \boldsymbol{y}_1)^2} \\ \boldsymbol{d}_3 &= \sqrt{(\hat{\boldsymbol{x}}_T - \boldsymbol{x}_1)^2 + (\hat{\boldsymbol{y}}_T - \boldsymbol{y}_2)^2} \\ \boldsymbol{d}_4 &= \sqrt{(\hat{\boldsymbol{x}}_T - \boldsymbol{x}_2)^2 + (\hat{\boldsymbol{y}}_T - \boldsymbol{y}_2)^2} \end{aligned} \tag{4}$$

Furthermore, we choose the bound distance:

$$\boldsymbol{S} = \alpha \sqrt{\hat{\boldsymbol{v}}_{Tx}^2 + \hat{\boldsymbol{v}}_{Ty}^2} \cdot \boldsymbol{T} \tag{5}$$

here $\boldsymbol{T}$ is sampling period; $\alpha > 1$ is a constant, can be decided by sampling period, locating precision and the prior information of island.

If $\boldsymbol{d}_i < \boldsymbol{S}$, the particle selection procedure will be executed, it will lead to the deduction of particles and the variation of relevant weight. Assume through selecting, the numbers of particles reduce to $\boldsymbol{M}$, among these particles, we sort the particle $(\boldsymbol{x}_k^a, \boldsymbol{w}_k^a)$ with the highest weight, replicate it with $N - M - 1$ times, so that the weight $w_k^i = \dfrac{1}{N - M - 1} w_k^a$.

4 Simulations

A bearing-only maneuvering target tracking problem like Fig.1 is studied in the simulation. With the help of the electronic chart information, we can get the

geographical position of island 1 and 2. The target boat is located in the direction of 45^0 with velocity of 30knots. The tracking boat is moved through maneuvering path with velocity of 40knots. The distribution of the measurement noise is Gaussian, with deviation 1°. The sampling period is $\boldsymbol{T} = 4\boldsymbol{s}$, α is chosen to be $\alpha = 2$, adopt $N = 5000$ particles.

The simulation result is shown as Fig.4-Fig.6. In Fig.4, the tracking performance of PF, ECI-PF is compared with the true trajectory of target boat. As seen that the ECI-PF method is superior to the PF because the constraints are incorporated in the tracking process, especially at places nearby the island 1 and 2.

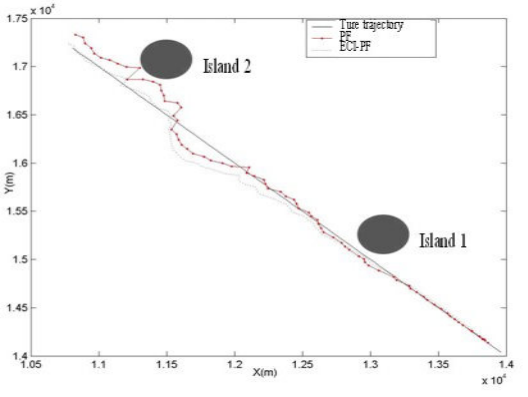


Fig. 4. Comparison of tracking trajectory

The position RMSE for x and y axis is presented in Fig.5, Fig.6.

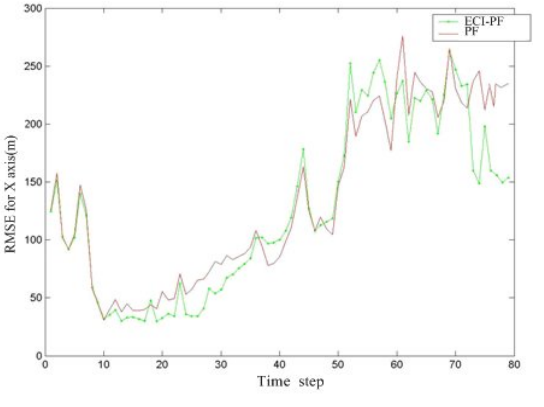


Fig. 5. Position RMSE for X axis

It is obvious that when the target boat near the island, at time step 25 and 62, ECI-PF method improved tracking performance compared to the PF. This may be caused by particle selection. During this procedure, the outrange particles are discarded and the replicated particles are closed to the true position, the particle clouds will move towards the direction of decreasing the tracking errors.

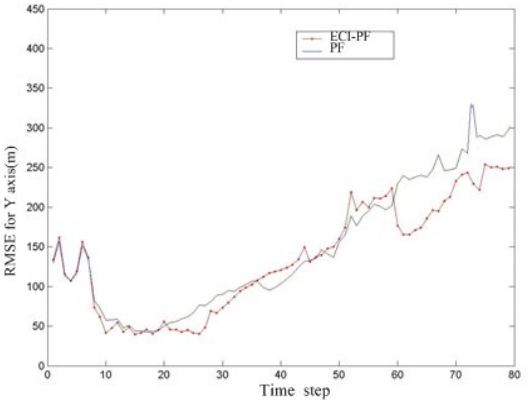


Fig. 6. Position RMSE for Y axis

5 Conclusions

In order to solve huge computation burden and poor tracking accuracy, particle filter based on electronic chart information (ECI-PF) is introduced in this paper. With the prior geographic information of nearby island, when the target boat is closed to the island, the distribution, the number and the weight of particle clouds are adjusted properly. Comparing this constrained filter with the standard particle filter, the simulation result shows that the incorporation of constrains can effectively improve the accuracy of the target tracking. In practice, other constraints can also be imposed on maneuvering target tracking problem, for example, hard constraints on target velocity or acceleration, etc. It is obvious that these constrains will reduce the effective number of particles, so that the operating speed of the algorithms can be highly enhanced.

References

1. Haug, A.J.: A tutorial on Bayesian Estimation and Tracking Techniques Applicable to Nonlinear and Non-Gaussian Process. MITRE Corporation, Mclean (2005)
2. Arulampalam, S., Maskell, S.: A tutorial on particle filters for no-line/no-Gaussian Bayesian tracking. IEEE Transactions of Signal Processing 50(2), 174–188 (2002)
3. Cheng, Y., Singh, T.: Efficient Particle Filtering for Road-Constrained Target Tracking. IEEE Transactions on AES 43(4), 1454–1469 (2007)
4. Deng, X.-L., Xie, J.-Y., Guo, W.-Z.: Bayesian target tracking based on particle filter. Journal of Systems Engineering and Electronics 16(3), 545–549 (2005)
5. Hall, P.: A Bayesian Approach to Map-Aided Vehicle Positioning. Master Thesis. Linköping University (2001)
6. Johansson, H.: Road-constrained target tracking using particle filter. Master Thesis. Linköpings universitet (2008)

Study of Speed Identification for Vector Control System Based on the Adaptive Variable Step-Size BP Algorithm

Yunhai Hou, Hongwei Zhang, and Kai Mei

School of Electrical & Electric Engineering, Changchun University of Technology, 130012 Changchun, China
houyunhai@mail.ccut.edu.cn, zhanghongwei@126.com

Abstract. A field oriented control system of speed sensorless based on Digital Signal Processor TMS320LF2407 is designed in this paper. According to the ANN model, a two-layered neural network is used to identify the speed. And a kind of adaptive variable step-size BP algorithm is put forward to adjust the weight of the neural model, making the energy of error function least, avoiding the shortcomings of traditional BP algorithm such as slow convergence rate. The experimental results show that the control system can identify the speed exactly and has better stability. It can achieve both good dynamic and static performance.

Keywords: List Vector control, Speed identification, Variable step-size BP algorithm, Neural network.

1 Introduction

The three-phase asynchronism motor is a complex object controlled with many variables and strong coupling. At the same time, it has the factors such as variational parameters, nonlinearity and uncertainty, and so on. So it is very difficult to obtain the good control performance. It promoted the development of AC drives with the sensorless vector control technique was put forward. And the characteristics of AC drives can compare beauty with DC drives on the basis of which the VC control combine with modern control technique and intelligent control strategy. In the paper, an AC motor speed regulating system with adaptive variable step-size BP algorithm for vector control was designed based on the close-loop vector control strategy with magnetic flux and rotate speed, pulse width modulation technique of SVPWM and neural network technique.

This paper uses flux, speed closed-loop vector control strategies, combined pulse width modulation SVPWM technique, speed sensor technology and neural network technology to design a speed identification of motor speed control system based on adaptive variable step size BP algorithm. The results show that the control system has speed response, high recognition accuracy and anti-interference ability.

2 Speed Identification Method

The core issue of sensorless vector controlis is the estimation of the motor flux and the observation of rotor speed, the quality of control system performance depends on the right combination of reasonable control program and speed identify links.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 419–426.
springerlink.com

In this paper, it uses reference adaptive speed identification method based on BP neural network model. Neural network model reference adaptive speed identification system block diagram is shown in Fig. 1. Estimator compares the output EMF of voltage model (reference model) and the output EMF of current model (neural network model can be adjusted), it improved BP algorithm to train the neural network, adjust the weights, then estimated EMF tracks given EMF, so that the estimated speed tracks actual speed. The system uses back-EMF signal to replace the traditional model output used by rotor flux as a design speed, and by means with the characteristics of neural networks, such as self-organizing self-learning, adaptive, improved the estimation performance greatly.

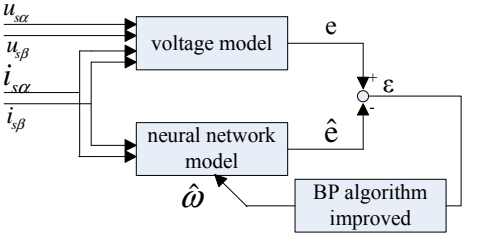


Fig. 1. Speed identification schematic based on BP neural network

2.1 Voltage Model

According to the theory of vector transform, in the two-phase cartesian coordinate system α / β, back-EMF component output of the stator can be expressed as:

$$e_{s\alpha} = \frac{L_r}{L_m}[u_{s\alpha} - R_s i_{s\alpha} - \sigma L_s \frac{di_{s\alpha}}{dt}]. \tag{1}$$

$$e_{s\beta} = \frac{L_r}{L_m}[u_{s\beta} - R_s i_{s\beta} - \sigma L_s \frac{di_{s\beta}}{dt}]. \tag{2}$$

Then the output of voltage reference model:

$$e = e_{s\alpha} + j e_{s\beta}. \tag{3}$$

That it can be strike through the inherent parameters detected of stator voltage, current and electrical.

2.2 Neural Network Model

Accordance the estimation method of current model, the estimation of stator back-EMF can be written as:

$$\frac{d\hat{e}_{s\alpha}}{dt} = -\frac{1}{\tau_r}\hat{e}_{s\alpha} - \hat{\omega}\hat{e}_{s\beta} + \frac{L_m}{\tau_r}\frac{di_{s\alpha}}{dt}. \tag{4}$$

$$\frac{d\hat{e}_{s\beta}}{dt} = \hat{\omega}\hat{e}_{s\alpha} - \frac{1}{\tau_r}\hat{e}_{s\beta} + \frac{L_m}{\tau_r}\frac{di_{s\beta}}{dt}. \tag{5}$$

Then

$$\hat{e} = \hat{e}_{s\alpha} + j\hat{e}_{s\beta}. \tag{6}$$

In which

R_s - motor stator resistance;

L_m - mutual inductance between stator and rotor;

L_s , L_r - stator, rotor inductance;

τ_r - rotor time constant, $\tau_r = L_r / R_r$;

σ - magnetic flux leakage coefficient, $\sigma = 1 - L_m{}^2 / L_s L_r$.

Neural network model uses two current sense input $i_{s\alpha}$, $i_{s\beta}$ and a volume and speed estimate $\hat{\omega}$, When the speed estimate $\hat{\omega}$ is inconsistent with the actual speed ω, the reference model and adaptive model of the EMF output will generate the error signal, which is

$$\varepsilon = e - \hat{e} = k_e(\omega - \hat{\omega}). \tag{7}$$

The system consider bias ε as the adjustment factor of BP algorithm, adjust the corresponding weights of neural networks by real-time online training, then achieve identification and non-poor regulation of closed loop speed control.

2.3 The Improved Design of BP Neural Network

Composite the formula (4), (5) aboved, it uses the forms of expression vectors may include:

$$\frac{d\hat{e}}{dt} = (-\frac{1}{\tau_r}A + \hat{\omega}B)\hat{e} + \frac{L_m}{\tau_r}\frac{di_s}{dt}. \tag{8}$$

In which

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Discretization can obtain:

$$\begin{aligned}\hat{e}(k) &= (1-\frac{T}{\tau_r})A\hat{e}(k-1)+T\hat{\omega}(k)B\hat{e}(k-1)+A\frac{L_m}{\tau_r}[i_s(k)-i_s(k-1)] \\ &= w_1x_1+w_2x_2+w_3x_3+w_4x_4\end{aligned} . \quad (9)$$

It takes $x_1 = A\hat{e}(k-1)$, $x_2 = B\hat{e}(k-1)$, $x_3 = Ai_s(k)$, $x_4 = -Ai_s(k-1)$ as the four nodes input variables of BP network, it can be obtained through the previous estimate of back-EMF and current value in this and previous detection. $w_1 = (1-\frac{T}{\tau_r})$, $w_2 = T\hat{\omega}(k)$, $w_3 = w_4 = \frac{L_m}{\tau_r}$ are the corresponding weight of four input nodes, in which, w_1 , w_3 , w_4 can be determined in advance by the motor parameters and the sampling time, only w_2 depends on the estimate of speed $\hat{\omega}(k)$, so it can adjust the network weights of speed identification online by error , so that gets the error to a minimum, then achieves the exact speed of identification. The two layers neural network model consisted by formula (9) is shown in Fig. 2.

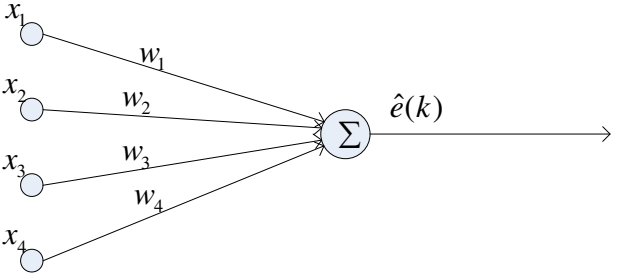


Fig. 2. Network structure of BP

Energy function is defined as:

$$E = \frac{1}{2}[e(k)-\hat{e}(k)]^2 = \frac{1}{2}\varepsilon^2(k) . \quad (10)$$

It strikes the right value of the increment by formula (9) and (10), namely:

$$\begin{aligned}\triangle w_2(k) &= -\frac{\partial E}{\partial w_2} = -\frac{\partial E}{\partial \hat{e}(k)}\bullet\frac{\partial \hat{e}(k)}{\partial w_2} . \\ &= [e(k)-\hat{e}(k)]^T B\hat{e}(k-1)\end{aligned}$$

Through introduce the learning efficiency to the minimum value of energy function to fix the iterative weights.

$$w_2(k) = w_2(k-1)+\eta\triangle w_2(k) . \quad (11)$$

In the traditional algorithm, the learning efficiency η namely the step size is constant, resulting in slower convergence, the system settling time longer, in particular, may also lead to network into a local optimum.To improve the convergence speed of the system, the system selected η as dependent variables in error, that takes:

$$\eta = 1 \Big/ d(k) \cdot$$

$$d(k) = (1-\alpha)d(k-1) + \varepsilon(k) . \tag{12}$$

In which: η is inversely proportional to $d(k)$, $d(k)$ is calculated by the error $\varepsilon(k)$ and the previous value $d(k-1)$ through formula (12) recursively. The system selectes coefficient α between 0 and 1, recursives initial value $d(0) = 0$ .As $d(k)$ changing with k increasing, and with $\varepsilon(k)$ proportional and with η to inversely proportional. In the training process, when E is small, it indicats that it adjusts in the right direction at this time, so the step increase; When E is large, it indicats that it adjusts in the wrong direction at this time, so the step size reduces. This will ensure that the right correct changes in the direction of making error ε always decreases, so it speeds up the convergence rate of the system.

With the variable step size improved BP algorithm, it overcomes the blindness to selecte step, adjustments the step value according to the error derived timely, prevents the oscillation effectively and speeds up the convergence rate of the network.

3 The Implementation and Control Process of System

Based on the BP neural network algorithm improved, the vector control system diagram of speed sensorless induction motor is shown in Fig. 3.

The entire system uses AC - DC - AC VVVF structure. The main circuit is composed by the non-controlled rectifier bridge, filter capacitor filter, SVPWM inverter circuit; The control circuit uses the high-performance digital signal processor TMS320LF2407 of U.S. Texas Instruments as the master unit. It has the feature of a computing speed, high precision, resource-rich in internal and external, cost performance, it can meet the requirements of real-time control. The control circuit establish communications with the host computer by RS485.

When system is working, it detects current and voltage firstly, through the appropriate A/D conversion and 3/2 transformation, then get it as the voltage model and the input neural network model, Then it dose back-EMF estimation in the neural network model, changes steps η and corrects the weights w_2 of BP based on the size of the error in real time, and then identify the motor speed $\hat{\omega}$,And then through PI regulator control of speed and current controller, get the wave voltage $u_{s\alpha}$, $u_{s\beta}$ used to generate the magnetic field space vector PWM, Finally, the six-way PWM wave created in the digital signal processor drives the inverter by circuit driven, then achieves speed closed-loop adjustment control of speed sensorless.

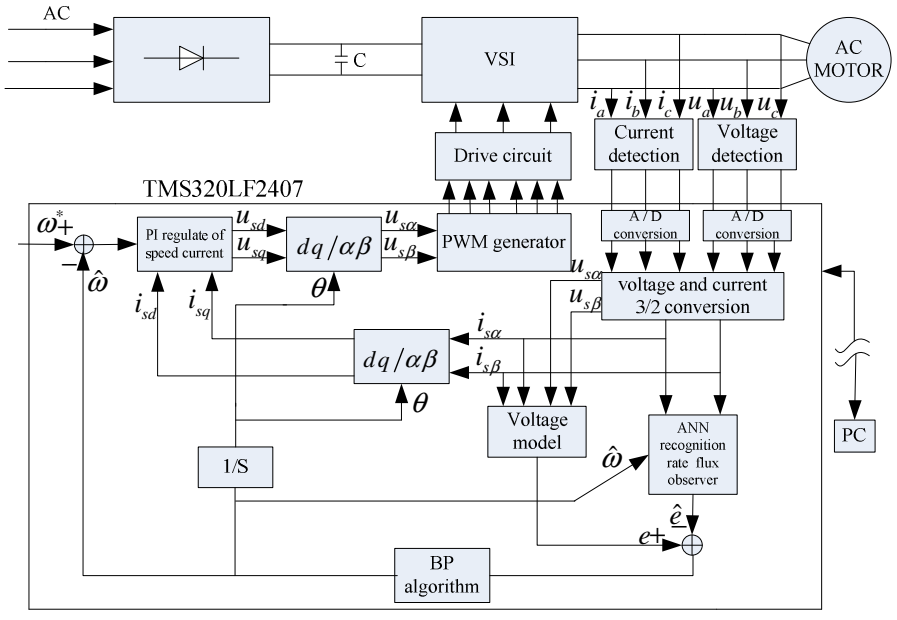


Fig. 3. The vector control system diagram of speed sensorless induction motor

4 The Results

According to figure 3, for the motor parameters in table 1, it builds the system and takes the precision test of motor speed control and interference test. the experimental results is shown in Fig. 4 and Fig. 5.

Fig. 4 shows the record curve in adjustment process of the motor speed given becomes from 900r/min to 1400r/min, it can be seen that the system accept given change from the moment $t = 46.6^{"}$, only costs to reach at 1381 r/min at $t = 48.5''$, then it uses the regulatory control of the system converges to a given speed 1400 r/min gradually.

Fig. 5 is the recording curve in the regulation process when the sudden load disturbance stable at the original 1400 r/min run.

The disturbance is getted suddenly at the moment $t = 82.3^{"}$, the speed decreases; After a quick adjustment of variable step size, it reached 1376 r/min after $t = 1^{"}$, after the further adjustment by the system in order to speed still closer to the given value, It shows that the system convergences speedly, has performance dynamic, its speed accuracy estimated has been achieved to good performance.

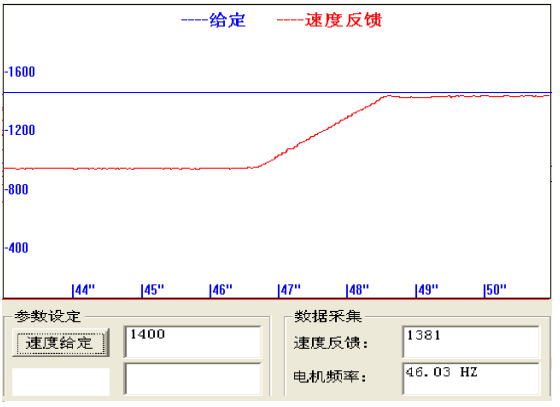


Fig. 4. Test curve of system speed control accuracy

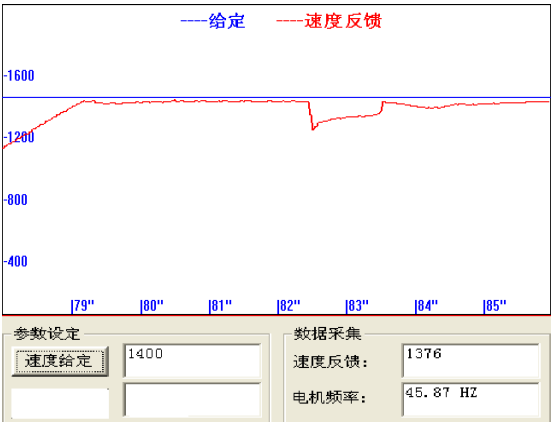


Fig. 5. Test curve of system interference

Table 1. Electrical parameters table

Rated voltage （V）	380
Power （kW）	0.37
Speed （rpm）	1400
Rated current （A）	1.94
Stator resistance （Ω）	1.01
Stator inductance （H）	0.25
Rotor resistance （Ω）	0.6
Rotor inductance （H）	0.2

5 Conclusion

In this paper, it combins the improved BP neural network algorithms and model reference adaptive methods, it designs a vector control system of induction motor speed sensorless based DSP, it uses software calculate Method instead of the sensing devices of speed and position, it also proposes a adaptive variable step size algorithms and solves the problems of slow convergence in traditional BP algorithm effectively, the results shows that the control system has performance dynamic, it can track the changes of motor speed accurately.

References

[1] Chen, B.: Electric drive control system. Mechanical Industry Press, Beijing (2000)
[2] Li, Y.: Digital control system of AC motor. Mechanical Industry Press, Beijing (2002)
[3] Schander: Adaptive Speed Identification for Vector of Induction Motors without Rotational Transducers. In: IAS 1989, pt.1, pp. 493–499 (1989)
[4] Hirokazu, T., Hori, Y.: Speed Sensorless Field-orientation Control of the Induction Machine. IEEE Trans. on IA 29(1), 175–180 (1993)
[5] Hagan, M.T.: Neural network design. Mechanical Industry Press, Beijing (2002)
[6] Cao, Y., et al.: The simulation of current-type inverter system of brushless DC motor. Beijing: Computer Information 20(12), 40 (2004)

Sleepiness Recognition Based on Time-Domain Analysis of ECG and Pulse Signals

AiHua Zhang[1], Qian Liu[1], and JinHui Bi[2]

[1] College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China
[2] College of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China
lutzhangah@163.com

Abstract. In order to find a simple and quick method of sleepiness recognition, the experiments were conducted and 26 healthy volunteers with good sleep quality were selected as the subjects. Electrocardiogram (ECG) and pulse signals in awakening and sleepiness states were synchronously acquired respectively and some time-domain features were extracted. The results showed that in sleepiness state the amplitude of T wave in ECG, the height of dicrotic wave, and pulse transit time decreased very significantly (p<0.005), heart rate reduced significantly (p<0.05), and the time from starting point of pulse wave to the pulse main wave point increased very significantly (p<0.005), comparing with awakening state. The classification accuracy rate was up to 100% using Support Vector Machines. This work demonstrated that time-domain features of ECG and pulse signals could reflect human sleepiness state objectively and be expected as indicators of sleepiness.

Keywords: sleepiness, ECG, pulse, Support Vector Machines.

1 Introduction

In sleepiness state, the cognitive ability of human is greatly weakened, and human reaction and the accuracy of various task performances decreased. In driving task and high-risk operation, workers must not be allowed to be sleepiness. Pilots, shift work doctors, bus drivers, power and steel industry workers, security guards, and so on, all need to work in high-awake-level state [1-2]. The sleepiness in work might make the efficiency decrease and cause serious accidents, even endanger people's life and property security. Therefore, it's necessary to detect sleepiness in practical life.

At present, the main methods of sleepiness detection contain the ways by measuring human behavior action, physical reaction, and physiological signal. For instance, some researchers detect sleepiness through drivers' behavior such as steering control and lane keeping, and some detect sleepiness by learning from speech and video signals [2-4]. There are also some scholars monitoring sleepiness by using physiological signals such as galvanic skin reflex, electrooculogram (EOG), electrocardiogram (ECG) and electro encephalograph (EEG), and so on[5-6]. Although the methods of human behavior action

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 427–431.
springerlink.com

and physical reaction can be carried out more easily, the accuracy rate of sleepiness recognition is lower. The detection of EEG may cause staff action inconvenience and be vulnerable to outside interference, so the practical application is highly restricted. The sleepiness detection research through ECG is relatively less, and it mainly concentrates on heart rate variability.

ECG and pulse contain rich physiological information in each human subsystem, and are convenience and real-time in measuring. The two signals were synchronously collected when experimental subjects were in awakening and sleepiness states, some time-domain features were extracted, and Support Vector Machines (SVM) was applied to classify features, so as to provide a basis for evaluating sleepiness.

2 Materials and Methods

BiopacMP150 instruments and Acqknowledge4.0 data acquisition software were used to acquire ECG, EEG and pulse signals synchronously, when 26 experiment subjects were in awakening and sleepiness states. Materials contain ECG100C amplifier, PPG100C amplifier, EEG100C amplifier, TSD200 sensor, EEG cap, 2 LEAD110S shielding wires, 4 LEAD100 shielding wires, 3 one-time patch electrodes, and so on.

Experiment process: 1) Select experiment subjects. According to the self-rating sub-health scale (score should be less than 45 points) and Pittsburgh sleep quality index (score should be less than 7 points), 26 cases of healthy subjects with good sleep quality were selected, and all experiment subjects have naps habit during lunch time. All subjects attended this experiment were required to be familiar with experimental procedures, operating method and attention matters beforehand. 2) Acquire signals. The experiment was in a well ventilated and quiet laboratory with comfortable temperature (21.2±4.2℃) and moderate humidity. Signals being awakening were acquired during early morning 9:00-9:30, and finished sleepiness signals collection when the subjects were sleepiness during lunch time. Put ECG electrodes in the same level place on the left anterior axillary line with the fifth rib on the left collarbone midline, clamp pulse transducer on left middle finger, set subject wearing the brain electrical cap. V_5 of ECG, pulse signal and EEG during awakening and sleepiness states can be synchronously collected. Distinguished whether the subject was in sleepiness state, according to the staging criteria standardized in 1968 by Rechtschaffen, A. and Kales, A. [7], and there were real-time videos as a reference.

3 Signal Processing

3.1 Signal Preprocessing

The main noises interfering physiological signals are: baseline drift, the frequency is less than 1 Hz; power frequency interference with the fixed frequency in 50 Hz; muscle power interference with the frequency ranging from 5 Hz to 2000 Hz.

For ECG, zero phase filtering method was used to remove baseline drift and elliptic trapped wave filter was applied to reduce power frequency interference, then muscle power interference could be filtered out with wavelet transform and smooth filter. The

collected pulse was mainly affected by baseline drift, because power frequency interference and majority of muscle power interference had been filtered by the acquisition hardware circuit in MP150 system. Baseline drift in pulse signals were removed by wavelet transform. Fig.1 shows the results of signal preprocessing.

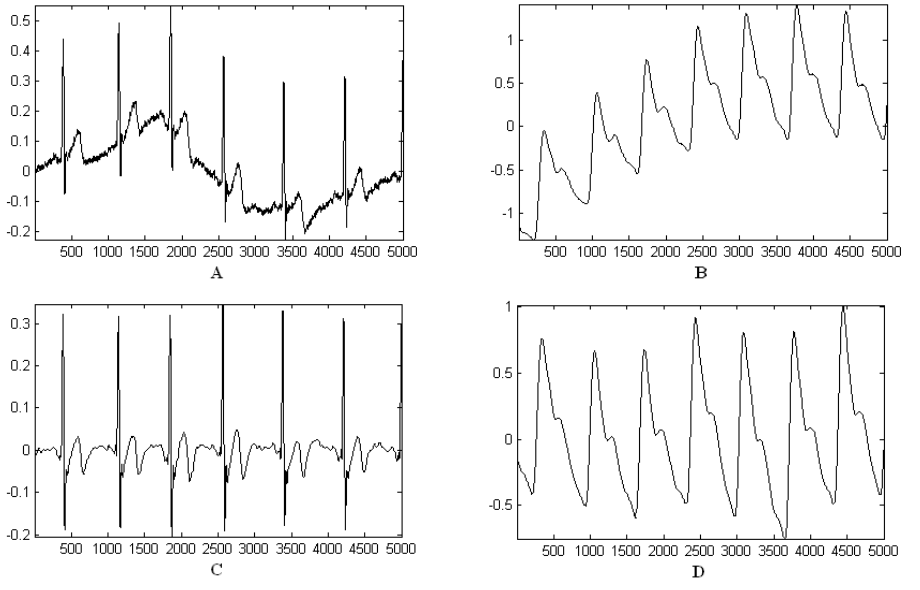


Fig. 1. Figures A and B show ECG and pulse original signals, respectively. And figures C and D show the corresponding ECG and pulse signals after signal preprocessing, respectively.

3.2 Feature Extraction and Analysis

Feature Extraction. Time-domain features were extracted from 60 seconds ECG and pulse signals and the sampling frequency is 1000 Hz. R wave is the highest wave in ECG forward direction. In this paper, Mexican-hat wavelet was used to find R wave position. R wave detection rate by detecting the extreme value point of the wavelet transformation reached 99.9% for the standard ECG data in MIT-BIH database [8], and the accuracy rate was more than 99% for the collected data in this experiment. The negative maximum points within certain length window were found around R wave in ECG. The point before R wave corresponds to Q wave and the point after R wave is S wave peak. Then RR interval sequence, RMSSD, Pnn50 and heart rate (HR) can be obtained through the difference of R wave peak points. After determining the maximum value point as the peak of T wave in ECG during 0.55 RR interval periods behind S wave, the difference between the wave amplitude and the value of ECG baseline can be calculated, namely AT. The values of R wave, Q wave, and S wave, QS interval, QT interval and so on, can also be figured out.

Pulse difference signal $y(n)$ can be got from once difference of pulse signal $x(n)$. The corresponding time of T wave peak point in ECG could be regarded as the starting point, and the first null point in signal $y(n)$ after the stating point can be found

out. The maximum point in $x(n)$ which is the main wave peak point in pulse is corresponding to the first null point time in $y(n)$. The time corresponding to the maximum point just mentioned was made once difference treatment to get the pulse main wave period sequence $P(i)$. On the left of each pulse main wave peak, the minimum point in a 300 sampling points window was regarded as the pulse wave starting point. The time t_1 from the pulse wave starting point to the main wave peak point and Pulse Transit Time (PTT) which was from R wave peak point in ECG to the pulse wave starting point can be easily calculated. In $y(n)$, $M(i)$ was the maximum points behind the main wave peak points in a range of 0.1 $P(i)$ and 0.4 $P(i)$, and the time of the latest null zero after each $M(i)$ in $y(n)$ did correspond to the position of dicrotic wave peak point in $x(n)$. Being limited by the equipment, the amplitude of dicrotic wave defined traditionally was not easily measured, so the height of dicrotic wave (HDW) which was the height value from pulse valley baseline to the dicrotic wave peak point was used to analyze in this paper. Pulse main wave value, and time from pulse main wave to dicrotic wave, also can be calculated.

Signal Analysis. After *t* test on all features extracted above, there are five features had significant or very significant difference between awakening and sleepiness states, as is shown in Table 1. Comparing with the awakening state, the amplitude of T wave (AT) in ECG, the height of dicrotic wave (HDW) and Pulse Transit Time (PTT) reduced significantly ($p<0.005$), and heart rate (HR) reduced obviously ($p<0.05$), the time from the pulse wave starting point to the pulse main wave point (t_1) increased significantly ($p<0.005$), in sleepiness state. Then, leave-one-out method is used to train samples, and SVM is suitable to solve the finite sample machine learning problem, so it's very applicable in this paper [9]. The RBF was chosen as the kernel function, and the punishment factor controlling the compromise between the sampling bias and machine generalization ability is set to 0.25 and the nuclear parameter is set to 0.0625 through the grid optimization.

Table 1. Results of *t* test and single feature classification

Feature	Awakening	Sleepiness	Result of *t* test	Classification accuracy Rate
AT(mV)	0.05±0.02	0.04±0.02	$p<0.005$	100%
HR	71.44±8.83	69.96±7.92	$p<0.05$	98.08%
HDW(mV)	3.52±1.35	2.11±1.02	$p<0.005$	92.31%
t_1(ms)	260.21±33.99	282.74±24.02	$p<0.005$	94.23%
PTT(ms)	80.80±45.52	55.48±37.18	$p<0.005$	100%

The classification accuracy rates of single features are all higher than 92.31%, and the accuracy rates of AT and PTT are 100%, as shown in Table 1.

After combining two or three features which can be selected from Table 1, the classification accuracy rates were calculated. The result shows that all classification accuracy rates of feature combination were up to 100%, except the combination HDW and t_1 is 98.08%.

4 Conclusions and Discussions

ECG and pulse signals of 26 experimental subjects were synchronously collected in awakening and sleepiness states, and some time-domain features were extracted, then the classification accuracy rates of selected features were calculated. The results show that feature combination is helpful to improve classification accuracy rate, and the accuracy rate of some features classification can reach 100% by using SVM method.

The characteristics of ECG and pulse represent certain human physiological information. Comparing with awakening state, the peak value of T wave in ECG significantly decreases in sleepiness state and this may be related with electrical activity of ventricular wall cells weakened, and HR, HDW and PTT decrease during sleepiness state which may be related with the excitement of human vagus nerve. Time t_1 is corresponding to the rapid ejection period of left ventricular. These time-domain features have significant or very significant difference between awakening and sleepiness states, so they can be used to distinguish the two states simply and quickly. This is helpful to realize real time sleepiness recognition.

In this experiment, all signals were acquired when experiment subjects were in quiet condition and the sample quantity is limited. So the result has a certain distance with clinical use. In future research, we plan to increase the samples number, extract features which more effectively reflecting sleepiness, and develop classification accuracy rate by combining features, hoping to find more optimal discrimination method of human sleepiness state.

References

1. Barbara, J.W., Anja, W., Wilhelm, D., Christian, H., Gerhard, O.: Objective and quantitative analysis of daytime sleepiness in physicians after night duties. Int. J. Psychophysiol. 72, 307–313 (2009)
2. Mortazavi, A., Eskandarian, A., Sayed, R.A.: Effect of drowsiness on driving performance variables of commercial vehicle drivers. Int. J. Automot. Techn. 10, 391–404 (2009)
3. Krajewski, J., Batliner, A., Golz, M.: Acoustic sleepiness detection: Framework and validation of a speech-adapted pattern recognition approach. Behavior Research Methods 41, 795–804 (2009)
4. Flores, M.J., Armingol, J.M., de la Escalera, A.: Real-time warning system for driver drowsiness detection using visual information. J. Intell. Robot. Syst. 59, 103–125 (2010)
5. Yamamoto, Y., Isshiki, H.: Instrument for controlling drowsiness using galvanic skin reflex. Med. Biol. Eng. Comput. 30, 562–564 (1992)
6. Kim, Y.S., Baek, H.J., Kim, J.S., Lee, H.B., Choi, J.M., Park, K.S.: Helmet-based physiological signal monitoring system. Eur. J. Appl. Physiol. 105, 365–372 (2009)
7. Rechtschaffen, A., Kales, A.: A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects. Public Health Service. US Government Printing Office, Washington (1968)
8. Homaeinezhad, M.R., Ghaffari, A., Atyabi, S.A.: Design of a Unified Framework for Analyzing Long-Duration Ambulatory ECG: Application for Extracting QRS Geometrical Features. Biomed. Eng. Lett. 1, 116–128 (2011)
9. Chong, Z., Xiao, L.Y.: Estimating mental fatigue based on electroencephalogram and heart rate variability. Pol. J. Med. Phys. Eng. 16, 67–84 (2010)

Design of a Low-Power Content-Addressable Memory Using Double-Feedback Match-Line Sense Amplifiers

Meng-Chou Chang, Ming-Hsun Hsieh, and Shih-Ju Tsai

National Changhua University of Education, Department of Electronic Engineering,
No. 2, Shida Rd., Changhua 50074,
Taiwan, R.O.C.
mchang@cc.ncue.edu.tw

Abstract. This paper presents the design of a low-power ternary content-addressable memory (TCAM) based on the proposed double-feedback match-line sense amplifier (DF-MLSA). DF-MLSA achieves power savings by reducing match-line voltage swing and search-line switching activity in a TCAM and employs the double positive-feedback network to boost the search speed of the TCAM. We have employed Hspice to evaluate various match-line sensing circuits using the Berkeley Predictive Technology Model (PTM) for 65 nm transistor model with supply voltage of 1.1 V and temperature of 25 °C. The simulation results show that the proposed DF-MLSA can achieve a search time of 0.49699 ns and an energy consumption of 0.201474 fJ/bit/search. The measured results show that the proposed DF-MLSA can reduce the energy consumption by 85.481%, 21.195%, and 10.545% compared to the conventional precharge MLSA, the stability-based MLSA, and the positive-feedback MLSA, respectively.

Keywords: Content-addressable memory (CAM), match-line sense amplifier.

1 Introduction

Content-addressable memory (CAM) compares the input search data in parallel against every stored data in the entire CAM array and then returns the address of the matching data. CAMs can be employed in a wide variety of applications requiring high-speed parallel search. These applications include pattern recognition, data compression, and network address translation [1]. However, the parallel search operation in CAMs consumes considerable power due to frequent switching of highly capacitive match-lines (MLs) and search-lines (SLs). For example, in a conventional precharge-high NOR-type CAM, all MLs are first precharged to V_{DD} and most of them, except for the few matching MLs, are eventually discharged to GND in every search cycle. Moreover, one of the two SLs in every SL pair first transits to HIGH to represent a valid search data and eventually transits to LOW to release the search data in every search cycle. In recent years, various match-line sensing schemes have been proposed to reduce the power dissipation of CAMs [2-6]. This paper proposes a novel match-line sense amplifier, called DF-MLSA, which achieves power savings by reducing ML voltage swing and SL switching activity in a TCAM and employs the double positive-feedback network to boost the search speed of the TCAM.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 433–438.
springerlink.com

2 The Proposed Double-Feedback Match-Line Sensing Scheme

Figure 1 shows the TCAM architecture for the proposed double-feedback match-line sensing scheme. This TCAM is composed of (1) a search key register, which holds the 144-bit search key, (2) 64 TCAM memory words, each of which stores one 144-bit data word and has one associated match-line used to signal whether the stored data word matches or mismatches the search key, (3) a dummy TCAM memory word, which mimics a matched word and is used to provide a timing signal to other memory words, and (4) 65 double-feedback sense amplifiers (DF-MLSAs).

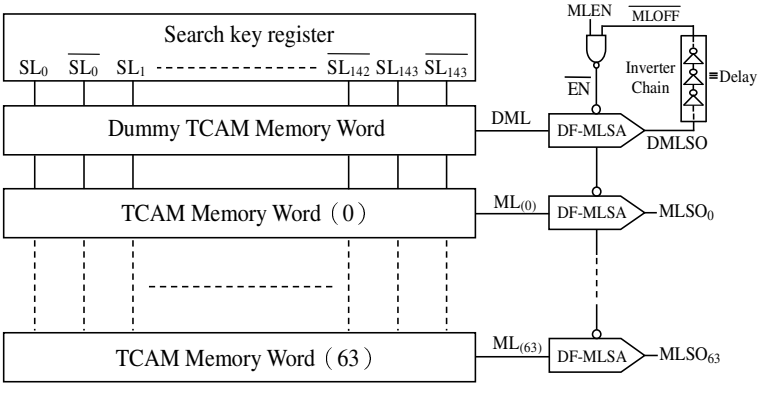


Fig. 1. The TCAM architecture for the proposed double-feedback match-line sensing scheme

In the proposed TCAM, each TCAM memory word is composed of 144 NOR-type TCAM cells, and the structure of a NOR-type TCAM cell is shown in Fig. 2. A TCAM cell can store a logic '0' (with D = 0 and $\overline{D}$ = 1), a logic '1' (with D = 1 and $\overline{D}$ = 0), or a *don't care* 'X' (with D = 1 and $\overline{D}$ = 1). Each pair of transistors (N_1,N_2) and (N_3,N_4) implements a pulldown path from the match-line to ground. A mismatch of the stored bit (represented by D and $\overline{D}$) and the search bit (represented by SL and $\overline{SL}$) enables one of the pulldown paths, connecting the match-line to ground. Since the match-line is shared by all the TCAM cells on the same memory word, the match-line

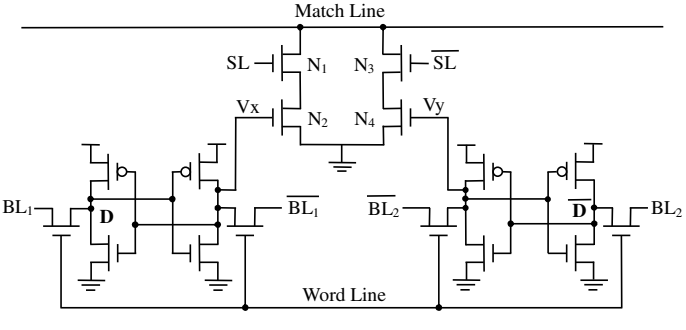


Fig. 2. The NOR-type TCAM cell

is connected to ground if any TCAM cell on the memory word has a mismatch. We use ML_n to refer to a match-line with *n*-bit mismatches. For example, ML_0 denotes a fully matched match-line, and ML_1 denotes a match line with 1-bit mismatch.

Figures 3 and 4 show the structure of the proposed double-feedback sense amplifier (DF-MLSA) for an ordinary CAM word and the dummy CAM word, respectively. The operation of the proposed TCAM is explained as follows.

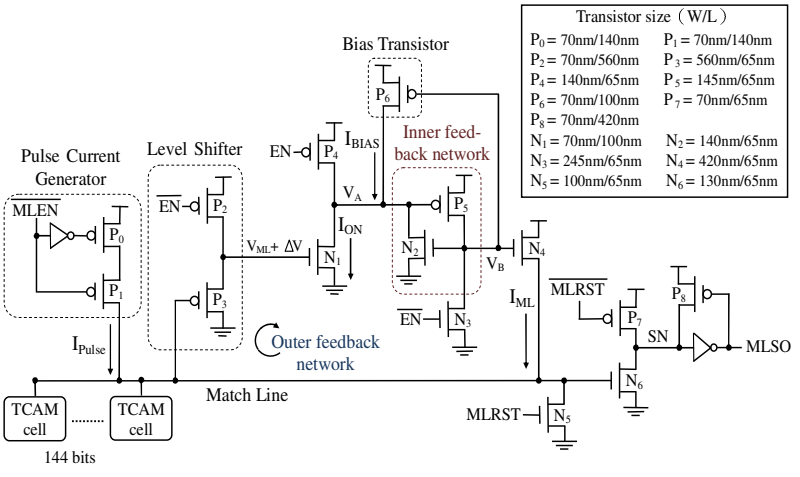


Fig. 3. Double-feedback match-line sense amplifier

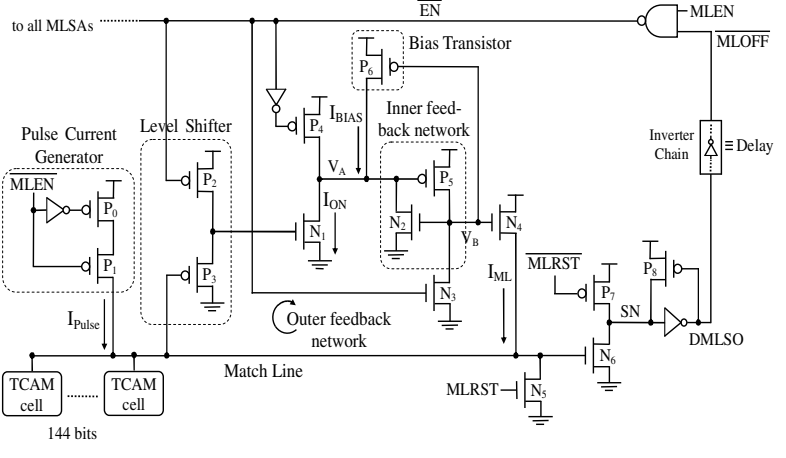


Fig. 4. Double-feedback match-line sense amplifier for the dummy row

Reset Phase. Prior to a search, MLEN (Match-Line Enable) is set to LOW, causing EN to become LOW (see Fig. 4), and MLRST (Match-Line Reset) is set to HIGH. From Fig. 3, EN being LOW turns PMOS P_4 on, charging node V_A to V_{DD}; $\overline{EN}$ being HIGH turns NMOS N_3 on, discharging node V_B to GND, and turns NMOS N_4 off, blocking the current I_{ML}. MLRST being HIGH turns NMOS N_5, discharging the match line to GND, and $\overline{MLRST}$ being LOW turns PMOS P_7 on, charging node SN to HIGH and setting the output of DF-MLSA to LOW.

Search Phase. To activate the search process, the search key is first applied to the SLs, MLEN is set to HIGH, causing EN to become HIGH and $\overline{EN}$ to become LOW, and MLRST is set to LOW. EN being HIGH turns PMOS P_4 off, and $\overline{EN}$ being LOW turns NMOS N_3 off and activates the level shifter composed of PMOS P_2 and P_3. The transition of $\overline{MLEN}$ from LOW to HIGH generates a pulse current I_{pulse} , which charges the match line for a short period of time. The voltage on a match-line develops differently depending on the number of mismatches.

Case 1 (case ML_0)**:** The data stored in the CAM word matches the search key. In this case, there is no conducting pulldown path between the match-line (ML_0) and ground. As the voltage on ML_0 ramps up, the voltage on the output of the level shift increases, turning the NMOS N_1 on and causing the current I_{ON} to flow. When I_{ON} becomes larger than current I_{BIAS}, voltage V_A (denoted by V_{A0} in Fig. 5) begins to ramp down. When V_A becomes smaller than V_{DD}-$|V_{TP}|$, PMOS P_5 turns on and voltage V_B (denoted by V_{B0} in Fig. 5) begins to ramp up. When V_B becomes larger than V_{TN}, NMOS N_4 and N_2 turn on. N_4 being on causes current I_{ML} to provide more charge to the match-line. Transistors N_4, P_3, N_1, P_5 form a outer positive feedback network and transistors P_5 and N_2 forms an inner positive feedback network, both of which amplifies the voltage on the match-line. When the voltage on the match-line exceeds V_{TN}, NMOS N_6 turns on, node SN discharges, and the output of the DF-MLSA, denoted by V_{MLSO_0} in Fig. 5, becomes HIGH, indicating the data stored in this CAM word is matched.

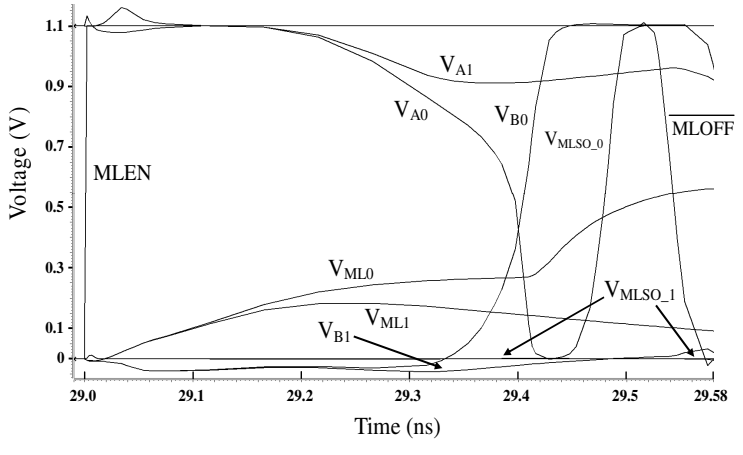


Fig. 5. Voltage waveforms for ML_0 (fully-matched) and ML_1 (one-bit mismatch)

Case 2 (case ML_n, n>1)**:** The data stored in the CAM word mismatches the search key. In this case, there is at least one conducting pulldown path between the match-line (ML_n) and ground. As a result, current I_{pulse} cannot charges ML_n to an enough voltage level to activate the inner and outer positive-feedback networks. The voltage on the match-line eventually decays to 0 V, and the output of the DF-MLSA keeps LOW, indicating the data stored in this CAM word is mismatched.

Dummy Row. The dummy TCAM memory word mimics a matched word and DMLSO, the output of the DF-MLSA for the dummy row, is used as a signal to shut off other DF-MLSAs. As shown in Fig. 4, when DMLSO becomes HIGH, $\overline{MLOFF}$ becomes LOW and $\overline{EN}$ becomes HIGH, causing current I_{BIAS} and I_{ML} to cease.

3 Simulation Results

The Hspice simulations for the TCAM with 64 words×144 bits using the proposed double-feedback match-line sense amplifier were performed using BPTM (Berkeley Predictive Transistor Model) 65 nm process technology with a V_{DD} of 1.1 V. Figure 5 shows the voltage waveforms of the match-line, V_A, V_B, and V_{MLSO} for the ML_0 and ML_1 cases. These notations are appended with a subscript to distinguish these two cases. For example, V_{A0} and V_{A1} denote the voltage waveform of V_A for cases ML_0 and ML_1, respectively. Figure 6 shows the voltage waveforms for match-line ML_n, $0 \le n \le 7$.

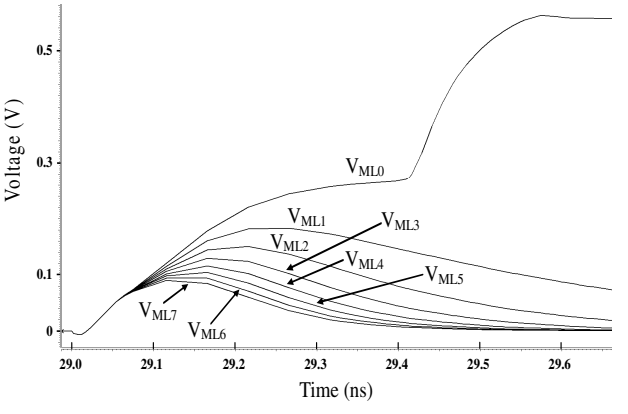


Fig. 6. Voltage waveforms for match-line ML_n, $0 \le n \le 7$

For comparison, the Hspice simulations for TCAMs using other match-line sensing schemes, including the traditional precharge scheme [1], the positive-feedback sense amplifier (PF-MLSA) [4,5], and the stability-based match-line sensing (SBMS) [6] were also performed using the same TCAM size and hspice parameters. Table 1 lists various comparisons for the TCAMs using four different match-line sensing schemes. The simulation results show that the proposed DF-MLSA can achieve a search time of

Table 1. Performance comparisons of four match-line sensing schemes

Match-line sensing scheme	PC-MLSA ([1])	PF-MLSA ([4, 5])	SB-MLSA ([6])	DF-MLSA (proposed)
Search time (ns)	1.1052	0.47518	1.0459	0.49699
Average power dissipation (mW)	11.752	4.4364	2.2880	3.7944
Power-delay product (fJ)	12989.0	2108.1	2393.0	1885.8
Energy（fJ/bit/search)	1.409397	0.225224	0.255662	0.201474
Energy-delay Product (yJS)	14355.44	1001.73	3073.90	937.22

0.49699 ns and an energy consumption of 0.201474 fJ/bit/search. The measured results show that the proposed DF-MLSA can reduce the energy consumption by 85.481%, 21.195%, and 10.545% compared to the conventional precharge MLSA, the stability-based MLSA, and the positive-feedback MLSA, respectively.

4 Conclusion

We have presented the design of the proposed match-line sense amplifier called double-feedback match-line sense amplifier (DF-MLSA), which achieves power savings by reducing ML voltage swing and SL switching activity in a TCAM and employs the double positive-feedback network to boost the search speed of TCAM. Simulation results have shown that the proposed DF-MLSA can outperform other match-line sensing schemes, such as the traditional precharge scheme, the positive feedback sense amplifier (PF-MLSA), and the stability-based match-line sensing (SBMS) in terms of power dissipation and power-delay product.

References

1. Pagiamtzis, K., Sheikholeslami, A.: Content-Addressable Memory (CAM) Circuits and Architectures: A Tutorial and Survey. IEEE J. Solid-State Circuits 41, 712–727 (2006)
2. Arsovski, I., Chandler, T., Sheikholeslami, A.: A Ternary Content-addressable Memory (TCAM) Based on 4T Static Storage and Including a Current-race Sensing Scheme. IEEE J. Solid-State Circuits 38, 155–158 (2003)
3. Arsovski, I., Sheikholeslami, A.: A Mismatch-dependent Power Allocation Technique for Match-line Sensing in Content-addressable Memories. IEEE J. Solid-State Circuits 38, 1958–1966 (2003)
4. Mohan, N., Fung, W., Wright, D., Sachdev, M.: Match Line Sense Amplifiers with Positive Feedback for Low-Power Content Addressable Memories. In: 2006 IEEE Custom Integrated Circuits Conference, pp. 297–300 (2006)
5. Mohan, N., Fung, W., Wright, D., Sachdev, M.: A Low-power Ternary CAM with Positive-feedback Match-line Sense Amplifiers. IEEE Trans. Circuits and Systems 56, 566–573 (2009)
6. Tyshchenko, O., Sheikholeslami, A.: Match Sensing Using Match-line Stability in Content-addressable Memories (CAM). IEEE J. Solid-State Circuits 43, 1972–1981 (2008)

High-Resolution Electronic Image Preprocessing of the Stored-Grain Insects

Hongtao Zhang[1], Yuxia Hu[2], and Sen Ju[1]

[1] Institute of Electric Power, North China University of Water Conservancy and Electric Power, Zhengzhou 450011, China
[2] College of Electric Engineering, Zhengzhou University, Zhengzhou 450001, China
zht1977@ncwu.edu.cn

Abstract. The stored-grain insects were of many species, small size and complex morphology. The high resolution electronic images of the stored-grain insects were acquired to be accurate identification. The image processing method based on multi-resolution analysis was proposed for the high resolution electronic images of the stored-grain insects. The algorithm effectively retained the head, elytra and other parts that had larger contribution to the recognition of the stored-grain insects. The object boundary of the insects was smooth after the pretreatment, and the efficiency of algorithm improved 15.12 times. The experiment showed that it was practical and feasible.

Keywords: stored-grain insects, multi-resolution analysis, image processing, feature extraction.

1 Introduction

It is especially important for our country to store grain safety. Accurate identification of stored-grain insects is the core content of the real-time detection of insects in grain mass. Methods to recognize the external adult insects include the artificial recognition, voice recognition, image recognition method [1-2]. The image recognition method has always been a research focus in the field of the detection for stored-grain insects for the past more than ten years. It is of high efficiency, and it is very convenient to connect the machine vision system with the current computer management system in grain depot [3-4].

Grain insects have many species, small size, and more complex morphology. The similarity among the some species is very high, such as *Cryptolestes ferrugineus, Cryptolestes pusillus and Cryptolestes turcicus*. The three species of grain insects were taken as a class for processing in the past research of image recognition [5]. The high-resolution electronic images of the stored-grain insects were used in order to achieve the local morphology features and the accurate classification of insects.

In the high-resolution image pre-processing of grain insects, it was very important to remove the body parts quickly and effectively that had the smaller contribution to classification of the grain insects, such as leg and antennae, and to retain the body parts that had larger contribution role. Then the binary electronic images that were

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 439–444.
springerlink.com

suitable for precise positioning, feature extracting and identifying of grain insects were formed.

To the high-resolution electronic images of grain insects, How to improve the running efficiency of the recognition software of insects was one of the difficulties in the software design of the detection system. In this paper, the high-resolution electronic images of grain insects were pretreated based on multi-resolution analyzer technology. And the experiment showed that it was practical and feasible.

2 High-Resolution Image Acquisition of Stored-Grain Insects

The automatic sieving and detection system was developed. The system included an automatic sieving subsystem of insects, an automatic transporting subsystem of insects, an illumination box and an image acquisition subsystem. The insects could be rapidly and efficiently separated from the grain sample, and removed dust automatically in the automatic sieving subsystem of insects. The automatic transporting subsystem of insects could accurately receive the sieve material, and transport the collection box to just below the image vision acquisition part for image acquisition. The even illumination chamber could provide even diffuse light for the sieve material in the collection box. Then the sieve material could be automatically removed after the image acquisition of insects. The system software analyzed the electronic images of sieve material, and determined the species of the grain insects based on computer vision and pattern recognition technologies.

The visible camera in the image acquisition subsystem (Model DH-SV2000FM, *Daheng* Co., Ltd.) was the line-scan CCD sensor. The resolution ratio of image was 1628×1236 pixels, the maximum frame rate was 14 frames per second. The exposure time was 20μs-3s, the rated power was 3W@12V. The camera was connected to the computer though IEEE-1394a acquisition card, and transferred captured electronic images to the computer.

A visible sieve material image of grain sample acquired by the detection system was shown in Figure 1. In the figure, the four dark dots near the boundaries of the image were characteristic markers. The markers were about the intersection of two

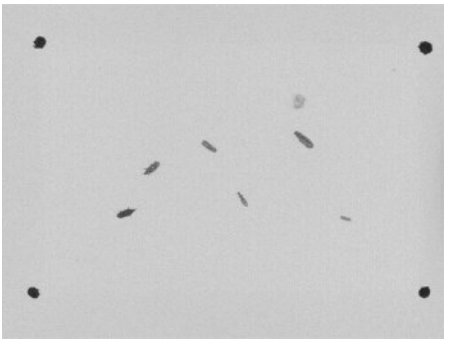


Fig. 1. The sieve material image of the wheat sample

lateral sides of the inverse trapezoid collection box. There was a certain height from the bottom of the collection box to the four markers. The markers were used to the precision location of grain insects. The light-colored circular object surrounded by the four markers was the impurity, and the remaining objects were the insects and the small impurities. The grain insects from left to right in the image were *Sitophilus oryzae, Sitophilus oryzae, Rhizopertha dominica, Oryzaephilus surinamensis, Tribolium castaneum* and *Cryptolestes ferrugineus.*

3 Multi-resolution Image Processing of Stored-Grain Insects

The small impurities and the body parts of the grain insects, such as leg and antennae, should be removed in image segmentation. The body parts that had larger contribution to the classification, like head and elytrum, should be kept unchanged. On the other hand, the resolution of the insects' electronic images acquired by the detection system was very high, and the operation of image preprocessing was relatively large. So we proposed the enhancement and segmentation method for the stored-grain insects based on multi-resolution image analysis.

The idea of this algorithm was to reduce the image resolution, and filter the low-resolution image. The logical "and" operation was carried out between the original image and the image restored to the original resolution. The specific treatment procedure was as the following steps:

S1: The grayscale image (Figure 1) was segmented into the binary image by the fixed thresholding (180).

S2: The size of the extended image was an integer multiple of the scaling factor to prevent the inconsistent size of the two electronic images after the image was reduced and enlarged. So we extended the image from 1628×1236 to 1632×1240. The inverse image was shown in Figure 2.

S3: The resolution of the original image was reduced to one fourth. Then the morphological open filter that the radius of the disc structure operator was one pixel was operated to remove small impurities and the grain insects' legs, antennae, etc. The results were shown in Figure 3.

Fig. 2. The extension and inverse of image

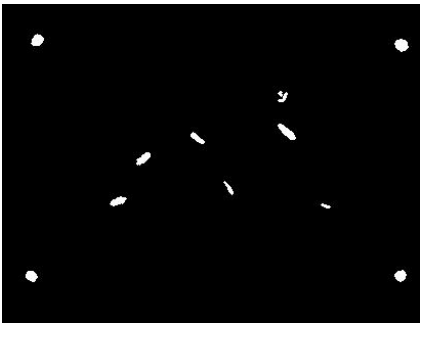


Fig. 3. The image by low-resolution filter

S4: In order to ensure the complete morphology of the grain insects after the subsequent operation between images, restoring the image to the original resolution(1632×1240), then operating the morphological dilation that the radius of the disc structure operator was two pixels. The results were shown in Figure 4.

S5: To remove the odontoid process of the object boundary and keep the object smooth, such as head and elytra of the insects, operating the logical "and" operation between the restored image (Fig.4) and the original image (Fig.5), then operating the "open" filter. The results were shown in Figure 5.

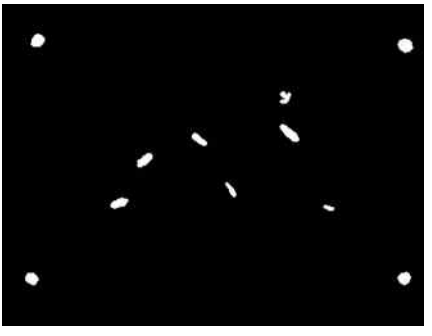


Fig. 4. The image by resolution restoration and filter

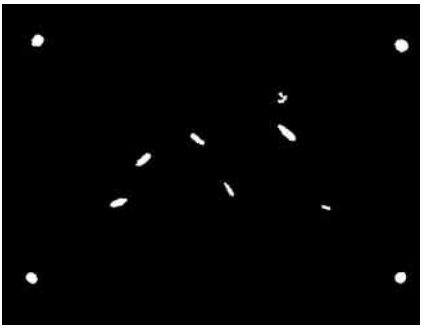


Fig. 5. The logical and operation of two electronic images and dilation

S6: The holes of the object in image needed to be filled. Because there might be holes in the object in the original binary image, and the holes might be still existence after the logical "and" operation. The results were shown in Figure 6.

S7: The markers were identified by area, roundness and other features. The centroids of markers (A, B, C, D) were (139.6, 143.2), (116.5, 1065.8), (1557, 1064.2) and (1560.8, 162) respectively. The markers taken as the background were removed. The impurities that were smaller than the smallest insects were removed according to the area threshold of the object (280). The results were shown in Figure 7.

Fig. 6. The image after removing holes

Fig. 7. The image by removing markers and some impurities

4 Results and Analysis

In the above step 3 and 4, if the equivalent filter was carried out directly in the original resolution, not in the low resolution, that was, firstly operating the "open" filter that the radius of the structure operator was four pixels, then carrying out the dilation operation that the radius of the structure operator was two pixels. The results showed that the smoothness of the object boundary was almost no change, and the speed of image processing was greatly reduced. The average time that the two steps run 10 times was from 0.135s to 2.041s, the operation efficiency reduced 15.12 times.

Although the resolution change of the image from decreasing to restoring the original condition lost the information of the object border and decreased the smoothness of the object border, the smoothness of the object boundary was made up by the subsequent extension operation of and the logical "and" operation between images. In addition, the resolution of the image was higher, and its computation that performed the same filtering operation was higher. If the resolution continued to decline, the efficiency had been greatly improved. On the other hand, the smoothness of the object boundary declined further and could not be made up well.

In addition, the large impurities in the image could not be accurately removed by the above algorithm. They could be identified and removed by the different morphological features between the stored-grain insects and impurities, such as complexity and elongation.

5 Conclusions

By the multi-resolution analysis of the stored-grain insects in image preprocessing, the body parts that had small contribution to identification of insects could be removed effectively, the parts that had large contribution could be retained effectively. The object boundary of insects was smooth after preprocessing, and the efficiency of algorithm improved 15.12 times. The subsequent precise location, feature extraction and classification could be met well.

Acknowledgments. Financial support was provided by the National Natural Science Foundation of China (No. 31101085 and 30871449), the Backdrop of Young Teachers Program, Universities of Henan Province (No. 2011GGJS-094) and the High-Level Personnel Science Research Start-up Project, North China University of Water Conservancy and Electric Power.

References

1. Neethirajan, S., Karunakaran, C., Jayas, D.S., et al.: Detection Techniques for Stored-product Insects in Grain. Food Control 18(2), 157–162 (2007)
2. Zhang, H.T., Mao, H.P.: Image Recognition and Classification of the Stored-grain Pests. Journal of Agricultural Mechanization Research 39(4), 157–162 (2008)

3. Singh, C.B., Jayas, D.S., Paliwal, et al.: Detection of Insect-damaged Wheat Kernels Using Near-infrared Hyperspectral Imaging. Journal of Stored Products Research 45(3), 151–158 (2009)
4. Zhang, H.T., Hu, Y.X., Mao, H.P.: Based on The Classification and Identification Research of Stored-grain Insect Image. Journal of Agricultural Mechanization Research 30(8), 36–38 (2008)
5. Zhang, H.T., Mao, H.P., Zhang, X.D.: Feature Selection of Stored-grain Insects Based on Artificial Fish Swarm Algorithm. Journal of Jiangsu University 31(5), 502–505 (2010)

Research of Fractional Calculus and Adaptive Algorithm Applied in Electronic Image Denoise

Zheng Wang

School of Electric and Control Engineering
Xi'an University of Science and Technology
Xi'an, China
wendy7830@126.com

Abstract. This paper applies fractional calculus to model for image processing firstly. Then the adaptive algorithm is used to define model parameters, and based on this, a fractional calculus and adaptive denoise algorithm is proposed. Numerical experiments showed that by means of the new algorithm, some good results can be achieved which can remove noise and keep the edges of the image and the texture detail. the textural detail enhancing capabilities of the new algorithm in this paper are better than that of Lee filter for rich-grained digital images.

Keywords: Fractional calculus, Adaptive algorithm, Image denoise.

1 Introduction

In the field of information science, especially in research of modern signal processing at home and abroad, fractional calculus are also emerging. Application of Fractional calculus in signal processing, especially in digital image processing, is a new discipline study branch.

In image processing, noise is one of the main reasons of the image quality decline. In image denoise, traditional methods such as Lee filters and wavelet denoise perform better, but they may fuzzy image details. In 2008, Aubert who obeys the speckle noise of Gamma distribution, proposed a denoise model called for Aubert-Aujol model. Aubert-Aujol model can be effective to remove speckle noise and keep the edge better, but be ineffective to remain the texture detail. In recent years, the method based on fractional models that can maintain and strengthen the image texture detail better, so this method become research hotspot of image denoise.

2 Fractional Calculus Modeling

Study on G-L of fractional calculus come from the classic definition of continuous function of integer order derivative, of which order number and dimensions are extended from integer to fraction.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 445–448.
springerlink.com

This article makes a fractional differential image model based on the Aubert-Aujol model .The following model is:

$$\min\left\{E(s)=\int_{\Omega}\left|D^{u}s\right|dxdy+\lambda\int_{\Omega}(\log s+f/s)dxdy\right\} \tag{1}$$

Where f is observed image, s is evaluated image. $\left|D^{u}s\right|=\sqrt{(D_{x}^{u}s)^{2}+(D_{y}^{u}s)^{2}}$, $D_{x}^{u}s$ and $D_{y}^{u}s$ are u -order partial differential of s about x and y , respectively.

Applying Calculus of Variation, formula (1) is: $s\cdot H^{u}(s)-\lambda\cdot\dfrac{f-s}{s}=0$, where $H^{u}(s)=\overline{(-1)^{u}}\left[\overline{D_{x}^{u}}(D_{x}^{u}s/\left|D^{u}s\right|)+\overline{D_{y}^{u}}(D_{y}^{u}s/\left|D^{u}s\right|)\right]$, $\overline{D^{u}}$ is conjugate operator of fractional differential operator D^{u} .

The size of the image is $N\times N$. $s=\left[s(i,j)\right]_{i,j=1}^{N}$, supposed $i,j<1$ or $i,j>N$, $s_{i,,j}=0$. Then the discrete operators $\Delta_{x}^{u}s$ and $\Delta_{y}^{u}s$ of $D_{x}^{u}s$ and $D_{y}^{u}s$ are respectively:

$$\left.\begin{aligned}(\Delta_{x}^{u}s)_{i,j}&=\sum_{k=0}^{K-1}(-1)^{k}\binom{u}{k}u(i-k,j)\\(\Delta_{y}^{u}s)_{i,j}&=\sum_{k=0}^{K-1}(-1)^{k}\binom{u}{k}u(i,j-k)\end{aligned}\right\} \tag{2}$$

Thus

$$\left[H^{u}(s)\right]_{i,j}=\sum_{k=0}^{K-1}(-1)^{k}\binom{u}{k}p_{1}(i+k,j)+\sum_{k=0}^{K-1}(-1)^{k}\binom{u}{k}p_{2}(i,j+k) \tag{3}$$

where

$$p_{1}(i,j)=\frac{(\Delta_{x}^{u}s)_{i,j}}{\sqrt{(\Delta_{x}^{u}s)_{i,j}^{2}+(\Delta_{y}^{u}s)_{i,j}^{2}+\varepsilon^{2}}},\quad p_{2}(i,j)=\frac{(\Delta_{y}^{u}s)_{i,j}}{\sqrt{(\Delta_{x}^{u}s)_{i,j}^{2}+(\Delta_{y}^{u}s)_{i,j}^{2}+\varepsilon^{2}}},$$

Where ε^{2} is a small constant which make the denominator is not zero. In practice, K need not a large value, here K=20.

So s is:

$$s^{(n+1)}=s^{(n)}+\Delta t\cdot\left[-s^{(n)}\cdot H^{u}(s^{(n)})+\sum_{j=0}^{L}2^{-2jm_{j}}\left(\lambda\cdot\frac{f-s^{(n)}}{s^{(n)}+\varepsilon^{2}}\right)_{j}\right] \tag{4}$$

In this iteration, parameters u , λ is updated adaptively.

In terms of computational complexity, for the size N*N of an image, when the model parameters are fixed, amount of calculation of the algorithm each iteration is O(N2). This paper takes into account the adaptive parameters approach, so the amount of calculation is larger than the fixed parameters.

3 Experiment Analyzing

Comparing the denoise effect of a artificial noise image by Lee filters with by the algorithm of this paper, Respectively. This paper algorithm parameters: Δt=0.1, L=3, $\varepsilon = 10^{-10}$. The comparison result is showed in table 1:

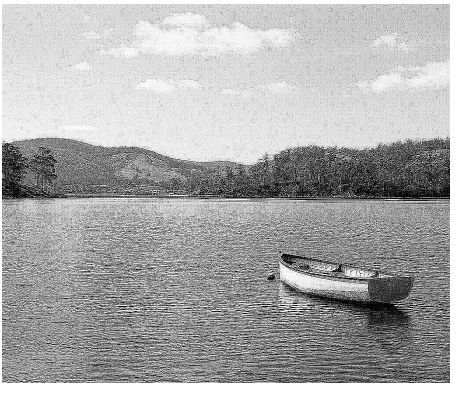


Fig. 1. Artificial noise original image

Fig. 2. The effect of Lee filter

Fig. 3. The effect of this paper algorithm

Table 1. Comparing the denoise effect of a artificial noise image by Lee filters with by the algorithm of this paper

	Iteration	PSNR	EPI
Original image		12.546	1.000
Lee filter	**300**	22.434	0.632
This paper algorithm	**300**	24.548	0.865

Notes: the peak signal-to-noise ratio (PSNR) serves as quantitative indicators measuring noise. Index of edge-preserving (EPI) maximum value is 1, the minimum value is 0, the value of EPI is more, then the ability of edge- preserving is better.

By comparing the results in table 1, after a number of iterations, the denoise effect of Lee filter is well but the edge and texture fuzzy seriously. This paper algorithm is good at the keeping of the edges and textures. From the comparison of peak signal-to-noise ratio, we can see that peak signal-to-noise ratio is the largest by the algorithm of this paper. The preserving effect of the image detail and edge is the best by the algorithm of this paper for 300 times iterations. From the comparison of quantitative index of edge-preserving (EPI), we can see that edge-preserving effect of this algorithm is also quite good.

4 Conclusion

Focusing on image denoise problems this paper comes up with a fractional differential model, and introduces adaptive algorithm applied on model parameters. Comparing with classic Lee filter, simulation experiments showed that adaptive algorithm based on fractional differential model in the denoise process, can acquire a better denoise effect and keep the image texture information. The algorithm is an effective method of image noise suppression.

References

1. Tatum, F.B.: The Relationship Between Fractional Calculus and Fractals. Fractals 3(I), 217–229 (2005)
2. Mandelbrot, B., Vainness, J.W.: Fractional Brownian Motions, Fractional Noise and Applications. SIAM Review 10(3), 422–437 (2008)
3. Yao, K., Zhang, X.: Research Announcements on the Fractional Calculus of A Type of Weiersrtass Functions 31(5) (October 2010)
4. Ver la referencia bibliogrdfica: Fractional differential equations, PODLUBNY. Igor (2009)
5. Samko, S.G., Kilbas, A.A., Marichev, O.I.: Fractional integrals and derivatives: theory and applications, Gordon & Breach (2007) (transl. from Russian (2003))
6. Audounet, J., Matignon, D., Montseny, G.: Diffusive representations of Fractional and pseudo-differential operators

The Application of Lifting DWT in Digital Image Processing

Azadeh Safari and Yinan Kong*

Department of Electronics, Faculty of science, Macquarie University
2109 Sydney, Australia
Azadeh.safari@mq.edu.au, yinan.kong@mq.edu.au

Abstract. Discrete wavelet transform (DWT) has shown great performance in digital image compression and denoising applications. It has been the transformation used for source encoding in JPEG2000 still image compression standard and FBI wavelet scalar quantization. This paper has adopted the lifting DWT which is the most computation-efficient scheme of wavelet analysis and outlines the multi-resolution features of the wavelet transform. Details of the lifting wavelet transform are analyzed to propose a high-speed, less-area and power-efficient digital image compression scheme. Maple 15 has been employed for design and simulation studies.

Keywords: Image compression, discrete wavelet transform, Lifting.

1 Introduction

The main objective of image compression is to decrease the storage space requirement, increase the transmission speed and minimize the chip overhead, which applies some constraints to specific applications. The lifting scheme helps optimize area and speed as well as achieve low power image coding.

In spite of high compression ratio and constructive image fidelity in the existent compression methods, they usually need long time for encoding. Using the lifting method is a promising new paradigm in digital image processing to handle area, power and timing restrictions [1], [2]. The aim of lifting compression in image processing is to achieve high throughput hardware/software architectures for real time applications [3]. This research aims to reduce the complexity of digital image processing by applying the lifting DWT which leads to performance of high speed and low consumption image processors. Furthermore, we have presented the details of the lifting wavelet followed by simulations in Maple 15.

The paper is organized in the following order: Details on image compression based on DWT are described in Section 2. Maple simulation is provided in Section 3. Lastly conclusions are given in Section 4.

* Corresponding author.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 449–453.
springerlink.com

2 Lifting DWT Image Encoding

Wavelet transform decomposes a signal *f* to its basis vectors using few coefficients. The basis vectors should match the signal features to represent the signal efficiently. In DWT, an image is decomposed by passing through an analysis filter bank which consists of high pass and low pass filter at each stage. Each stage of filtering results in four image sub-bands: LL, LH, HL and HH. The previous level's lower resolution reference signal LL becomes the next level sub-sampling input, and its related detail signal is obtained after the filtering at each level. The reconstruction procedure of the image is done by up sampling of 2 on all the image sub-bands at the coarsest scale followed by filtering the sub-bands in each dimension. After each level of decomposition, a reference signal with the resolution reduced by a factor of 2 and detail signals are obtained. These signals can be used for signal reconstruction when all the filtered sub-bands will be sum up to reach the LL sub-band. This procedure will be successively repeated until the image is fully reconstructed. Figure 1 shows the 2D wavelet transform [4], [5].

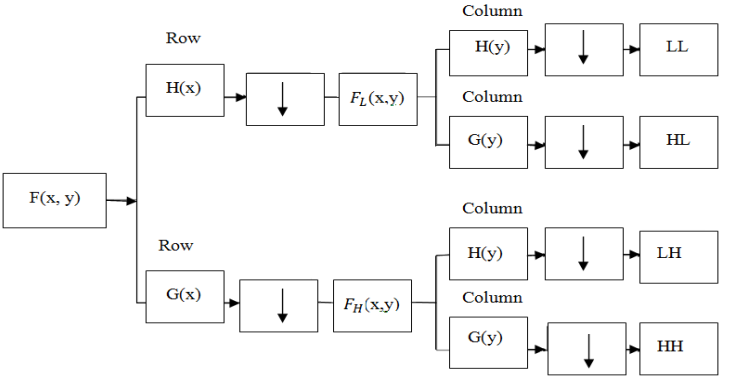


Fig. 1. The 2D wavelet transform

The (2, 2) Cohen-Daubechies-Feauveau (CDF) wavelet is widely used for image compression because of its good compression characteristics. The original filters have 5 + 3 = 8 filter coefficients, whereas an implementation with the lifting scheme has only 2 + 2 = 4 filter coefficients.

In Tables 1 and 2, forward and reverse filter coefficients are shown, where *x* represents the image pixel values, *s* stands for the summing or the low frequency coefficients and *d* stands for the difference or the high frequency coefficients.

The lifting scheme is the simple alternative in-place computation of wavelet coefficients [6].

Table 1. Forward transform coefficients of lifting scheme

	-()/2
	()/4

Table 2. Reverse Transform coefficients of lifting scheme

	/2
	+

Figure 2 shows block diagram of the lifting scheme. The first step only disjoints the odd and even samples. Then it computes the differences between the actual sample values and the prediction *(b-a)* and store it in place of *b*. An average value, *s*, is then computed in terms of *a* and the newly computed difference, *b*, as $s = a + b = 2$. The inverse transform can be performed by inverting the signs of polynomials and reversing the algorithm stages.

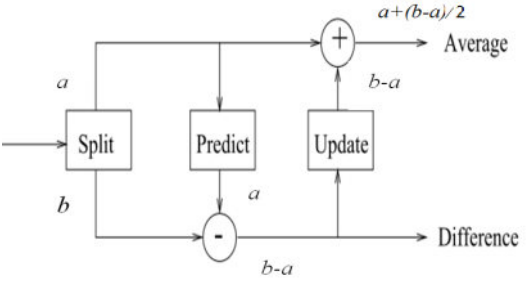


Fig. 2. Block diagram of the lifting scheme

If there are $2k$ elements in the row or column, then k low pass and k high pass wavelet transform coefficients will be produced after each level of filtering. The analysis filter equations are presented in (1) and (2).

High pass coefficients:

$$g(k)=2.x(2k+1)-x(2k)-x(2k+2) \tag{1}$$

Low pass coefficients:

$$h(k)=x(2k)+(g(k-1)+g(k)/8) \tag{2}$$

where $g(k)$ is the k^{th} high pass coefficient and $h(k)$ is the k^{th} low pass coefficient and $x(k)$ represents the input pixel value at the k^{th} position. The synthesis filter equations are presented in (3) and (4).

Even samples:

$$x(2k)=h(k)-(g(k-1)+g(k+1)/8) \tag{3}$$

Odd samples:

$$x(2k+1)=(g(k)+h(k)+h(k+1)/2) \tag{4}$$

The pseudo code for the lifting scheme is presented in Table 3.

Table 3. Pseudo code for the lifting scheme

```
Start
    Get input parameters:image, decomp level
    Set i:=0;
    while (the end condition is not met)
       Split the image pixels:
        s(i+1):=x(off1+(2*i*st)+1);
        d(i+1):=x(off1+(2*i*st)+st+1);
       Predict equations (5) and (6) were used to calculate
high pass and low pass coefficients, respectively:
       d(i+1):=d(i+1)+d(i+1)-s(i+1)-
       s(i+1+1);
       s(i+1):=s(i+1)+((d(i-
       1+1)+d(i+1))/8);
       Update Obtain the approximation signal:
       x(off1+(i*st)+1):=s(i+1);
       x(off1+((i+mid+1)*st)+1):=d(i+1);
    i := i + 1;
end
```

3 Maple Simulation

Maple is alternative software for digital image processing which has a huge amount of computations. Lifting DWT has been written in Maple worksheet for 256x256 phone.jpg image.

Image tools package performs various functions of image processing. By invoking this package, both grayscale and color images would be presented as arrays of 64-bit hardware floating point data. Using pseudo code in Table 3, the transform is as below:

```
for i from 0 by 256 to (nt-1)*256 do:
imgT:= T(A1,nt, ROW, i):
end do:
for i from 0 by 1 to nt-1 do:
imgT:= T(A1, nt, COL, i):
end do:
View(Create(imgT));
```

Figures 3 and 4, show the original image and decomposed image after 1 level of lifting DWT implementation, respectively.

Fig. 3. Original image 'phone.jpg' 256x256

Fig. 4. 2D Lifting wavelet implementation 'phone.jpg' 256x256

4 Conclusion and Future Work

Using lifting DWT for digital image processing reduces the complexity of operations and helps improve the simplicity of design. Applying lifting technique for image compression saves computation steps. In this scheme, the elementary algebraic operations are used instead of complicated convolution based operations. It allows for in-place computation which leads to faster image processors. The simplicity of lifting scheme was enhanced by applying symbolic computing software Maple for simulations. This experience is a gateway to enhance image processing features in the future. The foreseen of this paper is to design a high throughput, low power and less area image processor on field programmable gate array (FPGA). After FPGA prototyping we are aimed to implement our design on ASICs as well.

References

1. Grangetto, M., Martina, M., Masera, G., Piccinini, G., Vacca, F., Zamboni, M.: FPGA Power Efficient Inverse Lifting Wavelet IP. In: Proc. 35th IEEE Int. Asilomar Conf. Signals, Systems Computers, Pacif Groove, CA (November 2001)
2. Mccanny, P., Masud, S., McCanny, J.: An Efficient Architecture for the 2-D Biorthogonal Discrete Wavelet Transform. In: Proc. ICIP 2001, Thessaloniki, Greece (2001)
3. Andra, K., Chakrabarti, C., Acharya, T.: A VLSI Architecture for Lifting-Based Forward and Inverse Wavelet Transform. IEEE Transactions on Signal Processing 50(4) (April 2002)
4. Bhuyan, M.S., Amin, N., Hasni Madesa, M.A., Islam, M.S.: FPGA realization of lifting based forward discrete wavelet transform for JPEG 2000. International Journal of Circuits, Systems and Signal Processing 1(2), 124–129 (2007)
5. Jou, J.M., Shiau, Y.-H., Liu, C.-C.: Efficient VLSI architectures for the biorthogonal wavelet transform by filter bank and lifting scheme. In: IEEE Int. Symp. Circuits and Systems, Sydney, Australia, pp. 529–533 (May 2001)
6. Wagh, K.H., Dakhole, P.K., Adhau, V.G.: Design & implementation of JPEG2000 encoder using VHDL. In: Proceedings of the World Congress on Engineering, WCE, London, U.K, July 2-4 (2008)

Electronic Signal Processing Algorithm for Close-Talk System

Yi Jiang[1,2], Hong Zhou[2], Hao Zhang[2], Jun Qi[1], Yuanyuan Zu[2], Baoshuai Dong[1], and Wei Li[2]

[1] Department of Electronic Engineering, Tsinghua University, Beijing, P.R. China
[2] The Quartermaster Equipment Research Institute, CPLA, Beijing, P.R. China
jiangyi09@mails.tsinghua.edu.cn

Abstract. To portable electronic devices or in big noise environments, close-talk system was often introduced to collect more target electronic signal than common sound collect systems. But the performance of such system was not good enough in big noise applications. In this paper, within the framework of computational auditory scene analysis (CASA), an electronic signal processing algorithm of binary masking based on BelaSigna 300 development tool was proposed. Two microphones got the target sound and far noise at the same time, and then all in one IC calculated the inter-microphone intensity differences (IID) of the two microphones in time-frequency (T-F) units were used as cues to generate the binary masks for the near microphone. With the theory of ideal binary mask (IBM) and head-related transfer functions (HRTF), the threshold value of IID was set to 2. Experiments with one interfere noise in different positions were did to test this ideal.

Keywords: speech signal segregation, electronic engineering, Computational Auditory Scene Analysis (CASA), inter-microphone intensity differences (IID), system on chip (SOC).

1 Introduction

In daily life, how to obtain the clean target speech is every important to communication and robust ASR. The target speech is always mixed with the noise in common environments. So Close-talk system is often used to improve the quantity of the speech collection by near the target sound source, such as Mobil phone or head wear microphone. Even to do so, get the clean speech is also a hard task in complex auditory scenes, especially in big noise environments, such as rail station, airport, and subway station and so on. The collected speech would also be in low signal to noise ratio. With the widely use of portable device, the speech enhancement of close-talk system based on high integrated circuits or system on chip (SOC) was become important.

In recent years, great progresses have been made on the study of computational auditory scene analysis (CASA) algorithm for speech separation from complex audio scenes [1]. And, ideal binary mask (IBM) has become the goal of such systems [2] in the critical of signal to noise ratio. Within the framework of CASA, the key point is to find proper cues to generate binary masks, and make it approach to IBM. The main cues in monaural speech segregation system include pitch period[3], onset/offset[4]

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 455–461.
springerlink.com

Location based sound segregation method was used to estimate ideal binary mask for far sound sources [5], which used a two microphones system to generate the cues of inter-aural time differences(ITD) and inter-aural intensity differences(IID) as person's two ears do. For the position of two microphones and the sound sources location were lack of constraints, this algorithms would get multi spatial results, and segregation might meet some problem on certain positions. But it provided a method to segregation speech with double microphones.

Binaural cues extraction was the key point in location based CASA system. To close talk system, the target sound source was near one microphone, in most condition, the distances were within several centimeters. In this paper, another microphone was placed at right ear, little far from the mouth. Then the binary cues of IID could be used to segregate the target speech. The achievement of study in auditory localization [6-8] would be used to amend the approach. Due to the research in head-related transfer functions (HRTF) of the nearby source, both the theoretical calculations and the experiments indicate that the IID increase substantially for lateral sound source as distance decrease below 1m, even at low frequencies, which IID were small with distant sound sources [7, 9]. So IID was a robust cue to distinguish the near speech and noise from far sound source. BelaSigna 300 chip is the product of On Semiconductor Company, which integrate 4 ADCs and a HEAR Configurable Accelerator, is suitable for our applications.

The rest of the paper is organized as follows. Section 2describes the architecture of the proposed algorithm, and introduces a method to estimating the threshold of the IID cues. Section 3 presents systematic evaluation of the system for two sound sources with different location.

2 Electronic Speech Signal Segregate Processing

The close-talk system was shown in Fig.1. The proposed closed-talk speech segregate processing consists of two parallel parts: the same auditory Filter bank was used to decompose input mixture signal from two microphones into T-F representation units respectively. Then the IID cues of microphone A and B were extracted to generate the binary mask. Subsequently, the binary masks were affected on the output of the near microphone A as the better-ear to group the target speech, and then resynthesized to obtain the target speech

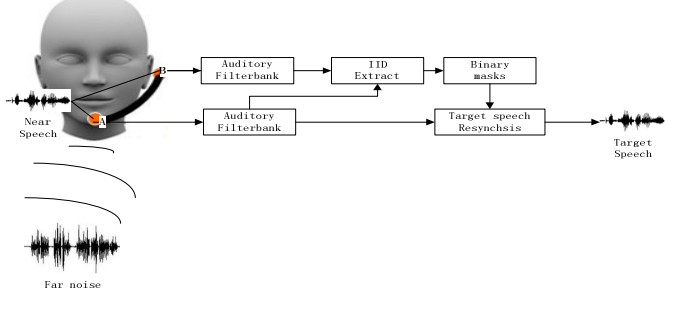


Fig. 1. Schematic diagram of the proposed algorithm. Microphone A was placed font of the mouth about several centimeters. Microphone B was placed near the right ear.

2.1 Inter-microphones Intensity Difference for CASA

Energy based CASA system used the energy of each time-frequency (T-F) units as the cues, have been widely used for its simple form and optimum in signal to noise ratio. The premier ideal of such system was to estimate the main energy source of T-F units, which can be describe as

$$m(t,f)=\begin{cases}1 & if\ |S(t,f)|^2 \geq |N(t,f)|^2 \\ 0 & otherwise\end{cases} \tag{1}$$

Where $|S(t,f)|^2$ indicated the energy of the target speech in the T-F units, and $|N(t,f)|^2$ indicated the energy of the mixed noise. $m(t,f)$ was the binary masks value in time t and frequency f, '1' indicated the target speech, '0' was not.In this paper, Background noises acoustically add to the clean speech and no correlation between them. So the nearby target sound source and far noise could be discussed separately. In this paper, we only discussed only one interfere noise condition. In the condition of two sound sources, the target one was near and the interferer noise was far away. And there were no correlations between the two sounds. The energy of microphone A and B in each channel could be calculated as

$$|X_A(t,f)|^2=|S_A(t,f)|^2+|N_A(t,f)|^2 \tag{2}$$

$$|X_B(t,f)|^2=|S_B(t,f)|^2+|N_B(t,f)|^2 \tag{3}$$

In close-talk system as fig.1, $|X_A(t,f)|^2$ were the T-F units energy of the microphone A, $|S_A(t,f)|^2$ were the energy of the target speech in microphone A, and $|N_A(t,f)|^2$ were the energy of noise in microphone A. while $|X_B(t,f)|^2$ could get form the peripherally analysis of microphone B. then the inter-microphone intensity differences of microphone A and B, IID_{AB} can be calculated in

$$IID_{AB}(t,f)=\frac{|X_A(t,f)|^2}{|X_B(t,f)|^2}=\frac{|S_A(t,f)|^2+|N_A(t,f)|^2}{|S_B(t,f)|^2+|N_B(t,f)|^2}=\frac{\frac{|S_A(t,f)|^2}{|N_A(t,f)|^2}+1}{\frac{|S_B(t,f)|^2}{|N_A(t,f)|^2}+\frac{|N_B(t,f)|^2}{|N_A(t,f)|^2}} \tag{4}$$

Then we defined the

$$IID_S(t,f)=\frac{|S_A(t,f)|^2}{|S_B(t,f)|^2}, \tag{5}$$

$$IID_N(t,f)=\frac{|N_A(t,f)|^2}{|N_B(t,f)|^2} \tag{6}$$

To describe the inter-microphone intensity differences of microphone A and B from near sound and far noise.

As the theory of IBM, the units was belong to the target unites, which can be described as $|S_A(t,f)|^2 \geq |N_A(t,f)|$, So the local signal to noise ratio (LC) of the microphone A, which is $\frac{|S_A(t,f)|^2}{|N_A(t,f)|^2}$ can be calculated from IID_{AB}.

To the target units, LC should equal or large than1. From the study in [5], $IID_{AB}(t,f)$ would bigger than $IID_N(t,f)$, and $IID_{AB}(t,f)$ less than $IID_S(t,f)$ 1 then in the close system:

$$LC = \frac{|S_A(t,f)|^2}{|N_A(t,f)|^2} = \frac{\frac{IID_{AB}(t,f)}{IID_N(t,f)} - 1}{(1 - \frac{IID_{AB}(t,f)}{IID_S(t,f)})} \tag{7}$$

$$IID_{AB}(t,f) \geq \frac{2}{\frac{1}{IID_N(t,f)} + \frac{1}{IID_S(t,f)}} \tag{8}$$

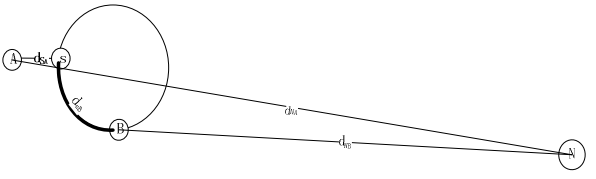


Fig. 2. IID estimate for a rigid spherical head

$$IID_S(t,f) = \frac{|S_A(t,f)|^2}{|S_B(t,f)|^2} = \frac{d_{SB}^2}{d_{SA}^2} \tag{9}$$

$$IID_N(t,f) = \frac{|N_A(t,f)|^2}{|N_B(t,f)|^2} = \frac{d_{NB}^2}{d_{NA}^2} \tag{10}$$

2.2 Inter-microphones Level Differences Estimate for a Rigid, Spherical Head

The cues of ILD found in the HRTF for nearby sources were important to auditory distance perception in the proximal region. For the complex of the auditory scene may generate in the proximal region, a simple sphere-model was used to explained the HRTFs intuitively [6]. In this paper, we also used a crude illustration of this phenomenon in Fig. 2.

Without thinking of the influence of head, the IID value of near target speech and far noise could be calculate as:

Where d_{SA} was the distance between mouth and microphone A, d_{SB} was the distance to microphone B. Where $d_{NA,}d_{NB}$ was the distance form noise to microphone A and B.

A series experiment of rabbits ears, the corresponding levels computed with the rigid sphere model for the sound source positioned at 90° azimuth and distances between 10 and 160 cm shown that the level increased with decreasing distance, and was relatively flat across frequency[9]. And in the condition of 160 cm, the ILD was under 5dB, which mean the biggest value of IID of the two ears in such distance was less than 3.1. In low frequency condition the value almost 0dB. In the close-talk system, the two microphones were positioned closer to each other than two ears did, and the noise in far position, so the value of d_{NB} and d_{NA} closed to each other, the value of IID_N was almost one. To the nearby sound sources, such as in the condition of 10cm distance, the IID was bigger than 10dB and almost 15dB in frequency higher than 10 kHz. In our close-talk system, the microphone was close to mouth no more than 5 centimeters, so the value of IID_S would bigger than the value in such experiment, and also increased with the frequency increase. So we can set the value of $\frac{1}{IID_S}$ to zero roughly. So the threshold value of IID_{AB} was equal to 2.

At last, the target speech was resynthesized from the output of microphone A with the binary masking based on proposed algorithm as mentioned forward to gain the target speech with little noise.

3 Evaluations and Comparison

In the one chip electronic evaluation system, a geography head model based on Chinese was used, and development tool of BelaSigna 300 as shown in fig 3. Two microphones and ADCs were used got the electronic signal form the position of mouth and right ear separately as the inputs. The target sound was position in front of the mouth about two centimeters so simulate the close talk. Another speaker was placed in front of the mouse one meter to generate the interfere noise on 0 degree or 45 degrees. The near target speech was a male speech; one female speech was used as the noise.

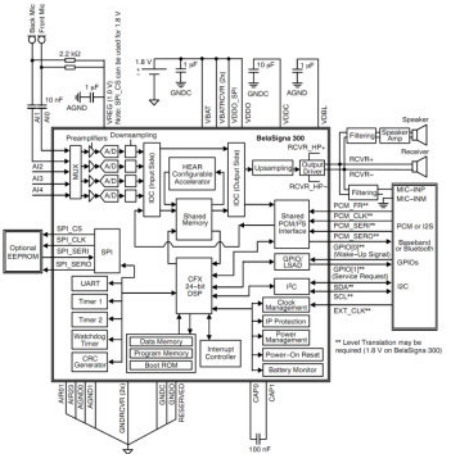


Fig. 3. The functional diagram of development tool BelaSigna 300

The system modeled auditory filtering by decomposing the input mixture signal into the time frequency domain using a bank of 128 gammatone filters or WOLA filters, with their center frequencies equally distributed on the equivalent rectangular band-width rate scale from50 to 8000Hz. In each filter channel, the output is divided into 20-ms time frames with 10-ms overlap between consecutive frames.

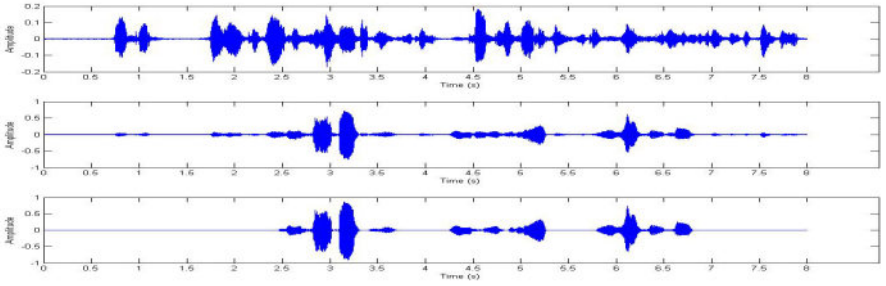


Fig. 4. Segregation result with 0 degree noise the top panel was the mixture signal of microphone B. there were enormous noise, and hard to find the target speech. the middle panel was the mixture signal of microphone A, in which the target speech was the mainly components. The target speaker was near to microphone A, and collected more target speech, which shape would also reduce the energy arrival of far noise. The bottom panel was the signal we segregated form microphone A with the algorithm we proposed.

In Fig.4, It was obviously that the noises from 0 to 2 seconds were reduced.

In Fig.5, when noise and target sound was in different orientation, the result of segregation was obviously good, especially compare to the mixture signal of microphone B. For the influence of speaker was reduced, the mixture signal of microphone A has bigger value than values in fig.4. Which mean the value should be increased to get better performance.

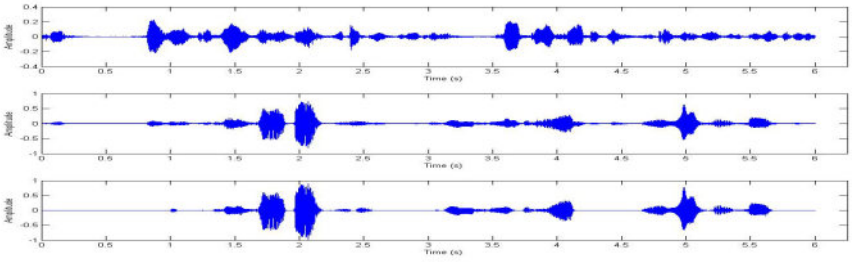


Fig. 5. Segregation result with 45 degree noise the signal got from microphone B and A was placed on up and middle panel. The bottom panel was the result of segregation.

5 Conclusions

We have presented an electronic signal processing Algorithm for close-talk system based on SOC development tool BelaSigna 300. With the integrated two ADCs and electronic hearing accelerator, the cues of IID extracted from the two microphones, Binary masking algorithm was proposed with certain threshold value of IID. The result of speech segregate experiment indicated it had good performance in one interfere noise on far position. As the founding of HRTF and electronic technology, IID was raised with the frequency raise. So the threshold value of the IID should be

various with the different frequency. But as all known, the accurate values of IID were hard to obtain, not only it depended on test equipment and processor, but also various with persons. So more experiments should be done to get mend this algorithm. At the same time, the performance of the algorithm we proposed should be evaluated with different noise types, various noise numbers and in different signal to noise ratio. The circuits of development tool were also need to be careful designed to reduce the noise of electronic circuits.

References

1. Brown, G.J., Cooke, M.: Computational auditory scene analysis. Computer Speech and Language 8(4), 297–336 (1994)
2. Wang, D.: On Ideal Binary Mask As the Computational Goal of Auditory Scene Analysis. In: Divenyi, P. (ed.) Speech Separation by Humans and Machines, pp. 181–197. Kluwer, Norwell (2005)
3. Brungart, D.S., Chang, P.S., Simpson, B.D., Wang, D.: Isolating the energetic com ponent of speech-on-speech masking with ideal time-frequency segregation. Journal of The Acoustical Society of America 120(6), 4007–4018 (2006)
4. Hu, G., Wang, D.: A Tandem Algorithm for Pitch Estimation and Voiced Speech Segregation. IEEE Trans. Audio, Speech, and Language Processing 18(8), 2067–2079 (2010)
5. Chao-Ling, H., Jang, J.S.R.: On the Improvement of Singing Voice Separation for Monaural Recordings Using the MIR-1K Dataset. IEEE Trans. Audio, Speech, and Language Processing 18(2), 310–319 (2010)
6. Yang, S., Srinivasan, S., Zhaozhang, J., Jin, Z., Wang, D.: A computational auditory scene analysis system for speech segregation and robust speech recognition. Computer Speech and Language 24(1), 77–93 (2010)
7. Hu, G., Wang, D.: Monaural speech segregation based on pitch tracking and amplitude modulation. IEEE Trans. Neural Networks 15(5), 1135–1150 (2004)
8. Boll, S.F.: Suppression of acoustic noise in speech using spectral subtraction. IEEE Trans. Acoustics Speech and Signal Processing 27(2), 113–120 (1979)
9. Wang, D.L., Brown, G.J. (eds.): Computational Auditory Scene Analysis: Principles, Algorithms, and Applications. Wiley and IEEE Press, Hoboken, NJ (2006)

Research and Optimization of Virtual Scene Roaming Implementation in Electronic Image Processing

Shunxin Li and Yufan Mo

Wuhan University of Science and Technology, Wuhan, China, 430065
lishunxin72@163.com, vincemo@126.com

Abstract. Briefly introduce the concept and the principle of virtual scene roaming. Both photorealistic drawing and virtual environment drawing are the key technologies of virtual scene roaming in electronic image processing. Research is especially into the environment drawing method. Based on the study of the two methods in electronic image processing, there are some problems of virtual scene real-time roaming. After the analys, necessary selection and optimization should be done. This paper takes a new way to take the advantage of both methods and optimize the design**.**

Keywords: scene roaming, modeling, electronic image processing.

1 Virtual Scene Roaming

Virtual tour makes use of necessary input and output tools to produce a specific realistic virtual environment with visual auditory and tactile in, which puts the users like just in the reality. Meanwhile, the users can observe from different prospective of the scene and operate objects in the roaming space as they were just in it. The combination of virtual scene tour system is just as the follow figure 1.

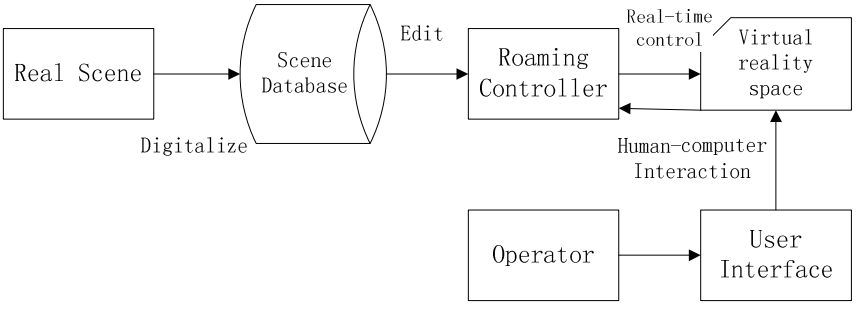


Fig. 1. The combination of virtual scene roaming system

2 The Key Technology of Virtual Scene Roaming

Virtual scene roaming technology is the ideal tool to exhibit architecture and planning design. The technology integrates the latest developing technical results such as

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 463–468.
springerlink.com

computer graphics, multimedia, artificial intelligence, multi-sensor, networks. It also makes user into the environment by lots different kinds of sensor devices, so user could naturally do some direct interaction with the environment. The main problems of virtual scene roaming are modeling and real-time drawing. As each of the two problems has its own features and it is a real obstacle to optimize both the methods in one time, so we need a balance between real-time display speed and model granularity. There are two different kinds of methods to produce 3-demensional realistic graphics: One is based on modeling, the specific description provided by computers could transfer the graphic with the point of observers' view; the other is based on graphics, a series of clipped scene kept by computers could generate the picture of other viewpoints. It is the large power of 3-demensional space, the man-machine interactive operation environment and the strong sense of immersive that makes one who is not so professional could take part in the design process.

2.1 Photorealistic Rendering Method

There are two common photorealistic rendering methods in electronic image processing: light tracing algorithm and radiometric algorithm.

Light tracing algorithm is a practical graphics drawing technology with a high sense of reality. It can generate overall lighting effects which cannot simulate by Z-buffer algorithm. It has a high speed of computation and is easier than radiation method which involved in recursive intersection and anti-aliasing technology.

In the process of light tracing, every ray in fact intersects with few surfaces. Because building shape surface distribute inhomogeneous in the whole scene area, octree data structure is applicable to the algorithm. Taking the importance of computing speed into account, Phong model is chosen to be the partial illumination model. The overall light is generated by Wllited model. The formula is just as equation (1).

$$I = Ic + ksIs + ktIt$$

$$Ic = kaIpa + [KdIpd(No \cdot Lo) + KsIps(No \cdot Ho)n] \tag{1}$$

Ic is the partial brightness produced by Phong model; Is, It are environment brightness from the direction of specula reflection and regular transmission; Ipa is the brightness of reflected light under the perpendicular incident light. Ips is the brightness of specular reflection from the direction of specular reflection; Ka, Kd, Ks, Kt are the ratio coefficient of environment reflection, diffuse reflection, specualr reflection and regular perspective reflection; No is normal vector; L is light vector; Ho is the unit vector along the angle bisector of light vector and view vector; n is the convergence index of specualr reflection.

Radiometric algorithm takes the curved surfaces which combined the environment as a closed system, and presume that the curved surface of the closed system is diffuse reflection curved surface. Then calculate the energy of every curved surface according to the law of conservation of energy principle. As all this steps be done, the brightness of the observed point is clearly seen.

2.2 Virtual Environment Drawing Method

There are two ways: the first is Geometry-Based Modeling and Rendering—GBMR; the second is Image-Based Modeling and Rendering—IBMR.

Geometry-Based Modeling and Rendering is based on computer graphics. At the beginning, texture mapping and control parameter setting should be done. Then, abstract the real scene, constructing 3-dimentiomal model of virtual landscape (including landform, buildings, trees, etc) by polygons, generating the light and scene-rendering picture of virtual environment. Finally output real-time scene-rendering pictures on the output devices under the control of special software from the fact of illumination, vanishing information, position of observer, thus finish roaming of the whole scene.

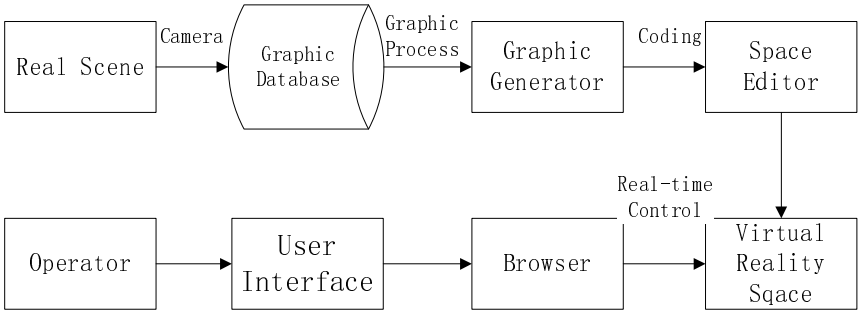


Fig. 2. The virtual drawing based on geometric modeling

Image-based modeling and rendering is based on the theory and method of computer vision and image processing, using the real picture to instead the traditional geometric modeling, taking the place of the traditional rendering process by image space transform operation. Compared with traditional computer graphics, rendering technology based on image has more advantages in generating realistic scene and higher speed of rendering. So it gradually becomes a hotspot in the development of computer graphics, especially in constructing virtual scenes. This kind of technology takes photos of real scenery as the processing object, so systems based on this technology are generally roaming in the scene which is actually existing and static, such as QuickTimevR, Hvs, etc. Virtual roaming based on images is not dependent on 3-dimentional geometric modeling, but use the discrete images or continuous video which collected by camera as the basic data, generating real whole-scene images by image processing. Then it organizes a number of whole-scene images into virtual real-scenery space via proper space model. Users could do some operations such as moving forward, moving backward, looking around, looking down, looking up, seeing near, seeing far, thus achieving the effect of all-round 3-dimentional scene roaming.

Compared with the two methods of virtual environment drawing, this article chose the hybrid modeling technology. The basic idea of hybrid modeling technology at first makes use of IBMR to generate virtual scene environment to get the vivid visual effects, and then take GBMR to construct the entity used to interact with user in the virtual

environment. This not only increased scene realism, but also guaranteed the real-time interactivity and enhanced user's immersed sense. In the practical application, for example, military or battlefield simulation, this kind virtual reality application of complicate scene involves simulation for lots of natural scenery such as landform, the wild jungle, lakes and rivers, etc and the virtual process of military simulation is always carried out on networks but not local computer, so it is obviously impossible to finish with only geometric scene modeling. Moreover, the large amount of data and calculation made it difficult to transfer, update and display the data in real-time. While hybrid modeling of virtual scene can keep the image data set which could generate virtual combat environment in local database, and the objects in geometric model which need to be interacted such as battle devices, weapons and so on are also stored in local database. Otherwise the algorithm to simulate the warring process is going on the server. This made the scene feels more verisimilar, and the network data transmission quantity is lower, so the simulation of the whole battle process could be successfully done.

3 The Real-Time Roaming of Virtual Scene

It is a big problem for real-time roaming in the enormous virtual scene. In the large computer graphic and image process system, it is also difficult to guarantee producing the geometric model in time.

3.1 Level Image Quick Caching Algorithm

This algorithm was published by Shade and so on. Its basic idea is the projection position is merely changing between the two pictures front and behind this frame when the object is far away from the view point. If there are lots of figures in the scene, this algorithm will generate a number of texture images which is great burden for hardware. In order to solve this problem, binary division technology was taken out to dissect scenery level. This algorithm use three steps to dissect level: the first step is to reduce the separation objects and virtual products; the second step is to keep the length ratio of node insulation box close to 1; the third step is to balance system structure. The algorithm geometrically operated the whole scene level but not the single object.

3.2 Roaming Technology Based on Texture Improved Level Presentation

This technology, which suggested by Maciel and so on, is to show the level details. It has two main features as follow: the first is to calculate the benefit value from the viewing angle, for example, automatically generating pictures and level structure by graphic hardware; the second feature is it selected database structure which has only one level, including the pictures of objects and object sets.

4 Optimization Design

Practice shows that user will lose their patience when the screen refresh rate is lower than 10 frames per sec in the scene roaming, and that will directly lead them to lose their

interests in watching it. In this article, optimization is taken in the 3DS MAX modeling procedure via the control of scene model complexity. In doing this, the time to refresh will decrease a lot. The optimization is executed in the following two main steps.

① Make a distinction between the important and the lesser one. Distinguish between the primary and secondary objects in the scene with different resolutions, so that to reduce the amount of files.

② Optimize in the file coding procedure. Use the ".gif" or ".jpg" files to instead the object which is not necessary to emphasize the precision.

It also can be optimized in the procedure of file editing. The own grammar features of VRML could reduce calculation cost of the web browser. At first import the ".wrl" scene files in VRMLPAD, checking the grammar and value type mistakes, mismatch of route and domain name, repeated node, and the value domain is out of the limit or not. After solving all of these problems, then debug or preview the selected node by the VRML browser in VRMLPAD is more convenient.

It is the file size which directly affects the browsing speed of the files in RAM. So it is essential to optimize the VRML files.

(1) Code reuse. There are often same objects in the virtual area; all these objects can be copied by DEF and USE example cited method, such as windows, desks, lights. There are always the same copies of these objects in people's daily life; they are just the same, so use the prototype method to do copies can avoid a lot of repeated coding.

(2) Use compressing tools. When make virtual room model with 3DS MAX, a large amount of redundant data left with saving ".wrl" files. It is necessary to compress the ".wrl" files in the end. The compressed files are one third of the original files in common. We can also use the special compressing tool provided by VRMLPAD or the gzip compressing tool to realize the compression of VRML files. It would produce a mess of data during the users linking to the internet, and in which way to record the user internet behaviors and the other data has great influence in system efficiency. The browser behavior monitor stores the data in different types for different information, including pictures, XML documents and database.

5 Conclusion

Virtual scene roaming is a new application in electronic image processing. With the rapid development of the society and the computer technology, the virtual scene roaming is about to get giant influence in the field of manufacture, art, electronic business, education, communication engineering, etc. In a certain sense, virtual reality technology changes people's way of thinking and convert the view of the world itself space and time. In a word, the implement of virtual scene roaming will turn machine code space nowadays into a colorful 3-dimentional world.

Compared with the advantages of GBMR and IBMR, this article provided the mixed rendering technology based on graphics and images which combined GBMR and IBMR. And then study and research the basic thought of real-time roaming. Find out the existing problems and analyze them. After all of these, optimization proposal is finally obtained.

References

1. Ramos Jr, J.G., Maeta, S.M., Bergerman, M., Bueno, S.S.: K.A Development of a VRML. Java unmanned airship simulating environment, pp. 1354–1359. IEEE (1999)
2. Alex, J., Vikramaditya, B., Nelson, B.J.: Teleoperated micromanipulation within a VRML environment using Java, pp. 1747–1752. IEEE (1998)
3. Aliaga, D.G., Lastra, A.A.: Smooth Transitions in Texture-Based Simplification. Computer & Graphics 22, 71–81 (1998)
4. Ibrahim, N., Balbed, M.A.M., Yusof, A.M., Hani, F., Salleh, M., Singh, J., Shahidan, M.S.: Virtual reality approach in acrophobia treatment. In: Proceedings 7th WSEAS International Conference on Applied Computer & Applied Computational Science, Hangzhou, China, pp. 194–197 (2008)
5. Wang, Z.: Virtual Reality Technology in Interior Design. Pioneering with Science & Technology Monthly 20(10), 181–182 (2007)

Study of Hybrid Genetic Algorithm Based on Simulated Annealing in Electronic Image Processing

Li Kang and Tao Yang

Department of Electrical Engineering and Automation, Luoyang Institute of Science and Technology, Luoyang, 471023, Peoples R China
{kangli20112012,yangtao2012}@163.com

Abstract. For setting up and maintaining a steady image conversion, hybrid genetic algorithm plays a key role in electronic image processing. This paper has researched the simulated annealing on the basis of introduction of the basic genetic algorithm, and analyzed the methods of the simulated annealing algorithm combined with genetic algorithm. Simulation of the simulated annealing hybrid genetic algorithm by a multi-peak functions in electronic image processing, and analyzed its performance. The hybrid genetic algorithm is capable of avoiding the premature convergence, improving the global search ability of the algorithm and the accuracy of the minimum function value in electronic engineering. The results show that the convergence speed can greatly accelerate.

Keywords: Convergence Speed, Simulated Annealing, Hybrid Genetic Algorithm, Electronic Image Processing.

1 Introduction

Genetic algorithm plays an important role in areas such as electronic engineering, mechanical engineering, electronic image processing. But basic genetic algorithm still produces many shortcomings and defects in the theory and application of technology [1].

Genetic algorithm is evolutionary algorithm of the earliest and most influence from more extensive application a research direction and field [2], it can achieve the whole solution space distributed information collection and search, although genetic algorithm and the traditional optimization algorithm has many advantages and distinct characteristics, but the genetic algorithm in application to basically have two big defects: fist, the convergence speed of genetic algorithm is too slow; Second, the genetic algorithm is easy to produce the precocious phenomena, and is easy to fall in local optima. Because it is difficulty to settle two large defects of basic genetic algorithm, so people began to study the method with improving the performance of hybrid genetic algorithm based on simulated annealing [3].

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 469–473.
springerlink.com

Simulated Annealing algorithm belongs to a general random search algorithm. Theory has proved that it can achieve the global minimum values, which attract attentions of many scholars [4] [5].Mixed genetic algorithm is one to make up for each other's deficiencies policy, it can make some kinds of performance, it is good to the improvement of the genetic algorithm appropriate mix to produce more superior performance of genetic algorithm [6].

In this paper, image conversion is introduced. The proposed hybrid genetic algorithm is based on simulated annealing, In this paper the basic genetic algorithm is insufficient, the simulated annealing algorithm and genetic algorithm combining method, puts forward a new hybrid genetic algorithm, the simulated annealing simulation mixed new genetic algorithm.

2 Simulated Annealing Algorithm

The basic ideas of simulated annealing mean from the process of five solids annealing [7]. Core teachings and thermodynamics are very similar to the mechanics principle, similar to the flow and crystal and cooling and the manner of annealing the metal. Simulation of annealing the algorithm in the search tactical and traditional random search method, it is not only import a proper random, but also introduced a system of annealing the process of the mechanism of introducing an algorithm, the object is used to control the parameter values, it is good at search for a complex area and to find out who have high expectations of the region. Simulation of annealing the algorithm theory proves that it can achieve the minimum.

At present, the simulated annealing algorithm research a lot, such as rapid simulated annealing algorithm[8], it can make the convergence speed of the algorithm improved greatly, adaptation of the simulated annealing algorithm, and makes the algorithm is intelligent.

3 The Hybrid Genetic Algorithm in Electronic Image Processing

Hybrid genetic algorithm based on simulated annealing and genetic algorithm are heuristic random genetic algorithm search, and the local search capability of genetic algorithm is low, but the capability of sure overall process is strong; And the local search capability of hybrid genetic algorithm based on simulated annealing is strong,wich can search through into the local and the best solution, if carry genetic algorithm with anneal algorithms are combined, learn from other's strong points and close the gap , there may be a fine performance of the new global search algorithm.

3.1 The Introduction of New Algorithm

Hybrid genetic algorithm based on simulated annealing combines the simulated annealing algorithm with basic genetic algorithms. And the basic genetic algorithm of the overall operation process is similar, genetic simulated annealing algorithm is also a

group of randomly generated from the initial solution began to the global optimal solution of the search process, it first through the selection, crossover and mutation genetic operation to create a new group of individuals, and then to the independent produced by each individual, it is simulated annealing process, with its results as the next generation of group of individuals. The operation process repeatedly until meet a termination conditions so far. Simulated annealing algorithm structure of the hybrid genetic algorithm is shown in figure 1.

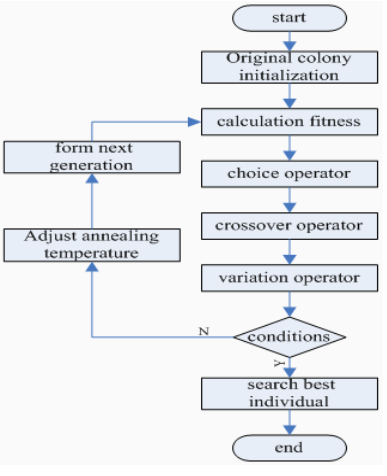


Fig. 1. The Flowchart of hybrid genetic algorithm

Hybrid genetic algorithm based on simulated annealing is two layers of parallel search structure. In the space level, GA provides parallel search structure, make SA into parallel SA algorithm, therefore, simulated annealing hybrid genetic algorithm to optimize group always parallel. In the process level, annealing genetic algorithm in the temperature were serially GA and SA search, is a kind of two layers of serial structure. Among them the SA initial solution from the evolution of the GA result, SA Metropolis sampling process solution through again become GA further evolution of the initial population.

Hybrid genetic algorithm based on simulated annealing using different fields search structure. The whole optimization process includes GA copy; cross, mutation and SA state of different fields such as produce function search structure. Copy operation optimization process to produce good model of redundant information for future generations inherit crossover operation father generation of fine mode, high temperature of SA operation help to optimize process of the local state small scale chemo taxis mobile, thus improve the algorithm in the solution space search ability and search efficiency.

The search behavior of hybrid genetic algorithm based on simulated annealing is controlled. The whole optimization process of the search behavior can be through the

annealing process try to control. Initial temperature control, can control algorithm of the initial search behavior: control of high and low temperature, can control algorithm sudden jump ability strong and the weak, high temperature sudden jump to the strong sexual avoid getting into the local optimum temperature chemo taxis of optimization to improve local search ability; Declining rate of temperature control, can control the declining ability of the sudden jump range, the influence of the search process smooth; Control sampling frequency, can control the temperature search ability, the influence of the search process corresponding homogeneous markov chain's smooth a probability distribution.

3.2 The Realization and Simulation of the Hybrid Genetic Algorithm in Electronic Image Processing

Here we come to study how to realize the maximum problem of the function:

$$f(t) = \exp[-2\ln 2(\frac{t-0.1}{0.8})^2]\sin^6[5\pi(t^{3/4}-0.05)] \quad (1)$$

This function ($t \in [0, 1]$) exists within the scope of the five great value points, and five great value point uniform with different sizes. This function can be a very good at test genetic algorithm, it is proposed to global search capability. The maximum function $f_{\max}$=0.989, maximum position t^*= 0.079.

The key control parameters of basic genetic algorithm are: population size N as 20, p_d =0.59 p_n=0.08, select the operation scale operator is used to anneal. The key control parameters of simulated annealing are: set t_0=10, the same as the ratio of methods of the temperature fall, thenα=0.8, per the number of iterations S=6,a five and a chain length markov L_k is 50, the total circle, the temperature down a total of T is20 .

The flowchart of hybrid genetic algorithm shown as in figure.1, as shown in figure 1 of mixed blood flow of annealing the algorithm for the functions and for the largest value of the operation, the experimental results show well based on the best preserved of the basic algorithm and based on the optimal conservation genetic algorithm and base on simulate annealing, each algorithms are both for 200 emulation, the result is two hundred times the average of simulation.

When the maximum value error $\varepsilon < 10^4$, this function is convergence, the convergence algebra of the basic genetic algorithm is 156.66, and the convergence algebra of hybrid genetic algorithm based on simulated annealing is 11.24. From the experiment result for the purpose of annealing, hybrid genetic algorithm based on simulated annealing is the best use of genetic algorithms faster and mixed genetic algorithm which greatly improved the convergence speed.

4 Conclusion

In the field of electronic engineering, this paper describes hybrid genetic algorithm based on simulated annealing combines the simulated annealing algorithm with basic genetic algorithms to make itself with more advantages of global optimization

problems. In the basic of basic genetic algorithms finish preliminary search, and then simulation annealing process, enhanced the basic the local search ability of genetic algorithm, accelerated greatly the convergence rate of the basic genetic algorithm in electronic image processing.

References

1. Huang, Z.-C., Hu, X.-L., Chen, S.-D.: Dynamic Traveling Salesman Problem based on Evolutionary Computation. In: Proceeding of the 2001 IEEE Congress on Evolutionary Computation, Seoul, Korea, vol. 3(1), pp. 32–33 (2001)
2. Prins, C.: Competitive genetic algorithms for the open-shop scheduling problem (2000)
3. Dethloff, J.: Vehicle routing and reverse logistics: The vehicle routing problem with simultaneous delivery and pick-up (2001)
4. Davies, T., Jnifene, A.: Multiple waypoint path planning for a mobile robot using genetic algorithms. In: Proceedings of 2006 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications, vol. 7(1), pp. 21–26 (2006)
5. Chen, C.-L., Vempati, V.S., Aljaber, N.: Anapplication of genetic algorithms for flow shop problems (1995)
6. Suer, G.A., Czajkiewicz, Z.: A heuristic procedure to minimize number of tardy jobs and total tardiness in single machine scheduling (1992)
7. Lee, K.Y., Mohamed, P.S.: A real-code genetic algorithm involving a hybrid cross over method for power plant control system design. In: Proceedings of 2002 Congress on Evolutionary Computation, vol. 2(12-17), pp. 1069–1074 (2002)
8. Suer, G.A., Czajkiewicz, Z., Baez, E.: Minimizing the number of tardy jobs in identical machine scheduling, March 8-10 (1993)

Electronic Trace Fossil Image Retrieval Based on Extended Direction Local Binary Pattern

Xiaosheng Wu[1] and Junding Sun[1,2]

[1] School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454003, China
[2] Image Processing & Image Communication Lab, Nanjing University of Post & Telecommunication, Nanjing, 210003, China
{wuxs,sunjd}@hpu.edu.cn

Abstract. Based on the direction local binary pattern (dLBP) and the center-symmetric local binary pattern (CS-LBP), an improved direction local binary pattern (idLBP) was presented. The directionality, coarseness and contrast of a texture region were considered together in the new definition. LBP, dLBP and idLBP were firstly tested on two commonly used texture image sets. Finally, idLBP was used for electronic trace fossil image retrieval. The results show that idLBP gives the best retrieval performance in the three descriptors.

Keywords: texture spectrum, dLBP, idLBP, electronic trace fossil image retrieval.

1 Introduction

Electronic trace fossil image has been used widely in the analysis trace fossil. As an effective method for electronic image description, Local binary pattern (LBP) [1] has successful used in texture classification, face recognition and image retrieval. Various extensions and modifications have been proposed during the past decades [2-8]. However, only the absolute gray-level differences of two pixels, the center pixel and one of its neighbors, are considered when defining the binary pattern in both LBP and most of its extensions. Consequently, it is hard to denote the directionality, coarseness and contrast of a texture region in a full way.

Heikkilä introduced center-symmetric local binary pattern (CS-LBP) for interest region description [7], which reduce the dimension of LBP from 256 to 16 for an 8-neighbourhood. In [8], direction coded local binary pattern (dLBP) was presented which adopted the gray-scale intensity invariance and the intensity inversion invariance and it produced the same number bits as LBP.

In this work, the power of dLBP was enhanced by introduced a new scheme for generating binary codes. The new operator is denoted as improved direction local binary pattern (idLBP). Comparisons were given on two commonly used texture databases. After that, idLBP was used for electric trace fossil image Retrieval.

The rest of the paper is organized as follows. In Section 2, we first briefly reviewed the conventional LBP descriptor and its extensions, and then introduced the proposed descriptor in details. The experimental results and discussions are given in Section 3. Finally, conclusions are drawn in Section 4.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 475–480.
springerlink.com

2 Local Binary Patterns

2.1 LBP

Ojala et al. [1] represent the texture information by defining texture T in a local neighborhood of a gray level image as the joint distribution of the gray levels of P image pixels as follows,

$$T = t(g_c, g_0, \cdots, g_{P-1}) \approx t(s(g_0 - g_c), \cdots, s(g_{P-1} - g_c)) \tag{1}$$

where $s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$, g_c corresponds to the gray value of the central pixel, and g_p (p=0,1,…,P-1) denote the gray values of the neighbors. The LBP value of a pixel is calculated as,

$$LBP = \sum_{p=0}^{P-1} s(g_p - g_c) \times 2^p \tag{2}$$

2.2 CS-LBP

CS-LBP does not compare the gray-level of each pixel with the center pixel, but the gray-level difference of the pairs of opposite pixels in a region [7]. A threshold is also set to obtain the robustness on flat regions. CS-LBP is given as follows,

$$CS\_LBP = \sum_{p=0}^{P/2-1} s(g_p - g_{p+P/2}) 2^p, s_1(x) = \begin{cases} 1, & x \geq threshold \\ 0, & x < threshold \end{cases} \tag{3}$$

It is clear that the dimension of CS-LBP is greatly reduced to the original LBP.

2.3 dLBP

Trefný [8] defined four base directions through the center pixel in a region. Then, they encode intensity variation along the directions into two bits, thus the binary word has the same length as the original LBP. The dLBP value can be written as,

$$dLBP = \sum_{p=0}^{P/2-1} (s(g_p - g_c)(g_{p+P/2} - g_c) 2^{2p} + s(|g_p - g_c| - |g_{p+P/2} - g_c|) 2^{2p+1}) \tag{4}$$

2.4 idLBP

Though the center pixel of a region is considered in dLBP, it has been proved that it has lower performance than LBP in the experiments because only four base directions were considered. In this paper, more directions are adopted in the definition of the improved direction local binary pattern. The new operator idLBP is given as follows,

$$idLBP = \sum_{p=0}^{P/2-1} (s(g_p - g_c) \odot s(g_c - g_{p+P/2})) 2^{2p} + s((g_p - g_c) - (g_c - g_{p+P/2})) 2^{2p+1}) \tag{5}$$

3 Experimental Results

Two commonly used texture databases are employed as test beds and χ^2 distance is chosen as a measurement criterion.

$$\chi^2(\boldsymbol{H}_1, \boldsymbol{H}_2) = \sum_{i=1}^{K} \frac{(h_{1i} - h_{2i})^2}{h_{1i} + h_{2i}} \tag{6}$$

where $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ are the texture spectrum histograms of two images. For simplicity, the operators are computed in an 8-neighborhood and the retrieval performance is assessed in terms of the commonly used precision and recall.

3.1 The Performance on Image Set 1

Image Set 1 was chosen form CUReT database, which includes 45 classes and each class has 20 texture images. Fig. 1 shows the 45 example textures. This experiment employs each image in the first 30 classes as query images, and 600 retrieval results are obtained. Fig. 2 presents the comparision graph of the average results. The retrieval results show that idLBP produces the best performance in the three operators.

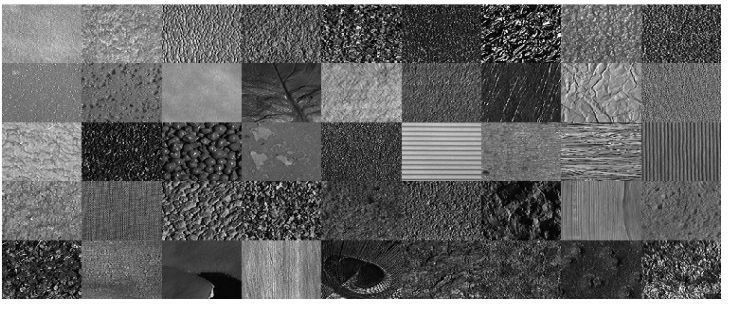


Fig. 1. Example images in set 1

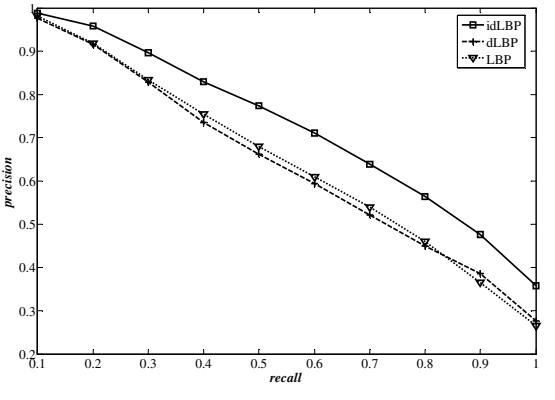


Fig. 2. Comparison on the first image set

3.2 The Performance on Image Set 2

Image Set 2 includes 109 Brodatz textures (Fig. 3), and each image is partitioned into 16 non-overlapping parts.Then, we can get 1744 images. This experiment employs each image in the first 60 classes as queries, and 960 retrieval results are obtained. Fig. 4 presents the comparision graph of the average results. The retrieval results also show that idLBP gets better average recall and precision than dLBP and LBP.

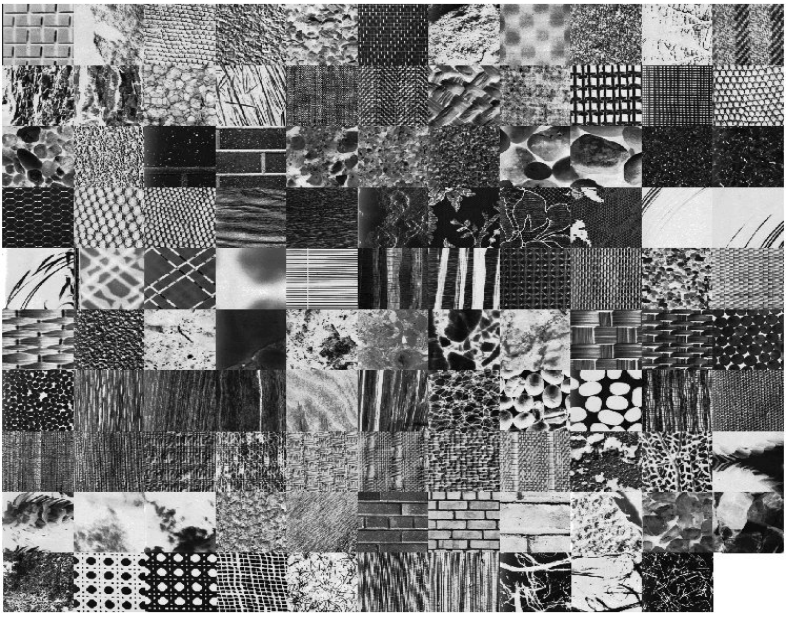


Fig. 3. Example images in set 2

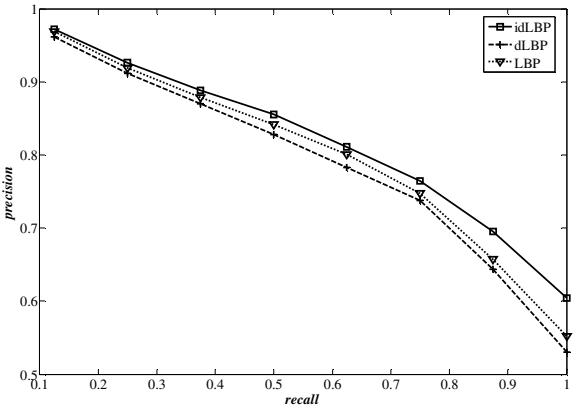


Fig. 4. Comparison on the second image set

3.3 Electric Trace Fossil Image Retrieval

In this experiment, we chose 521 electronic trace fossil images. Some examples are given in Fig. 5. Ten classes are chosen as query images in this experiment, and 160 retrieval results were obtained. Fig. 6 presents the precision-recall graph of the average results. The retrieval results show that idLBP is the best descriptor for electric trace fossil image retrieval.

Fig. 5. Example images in set 3

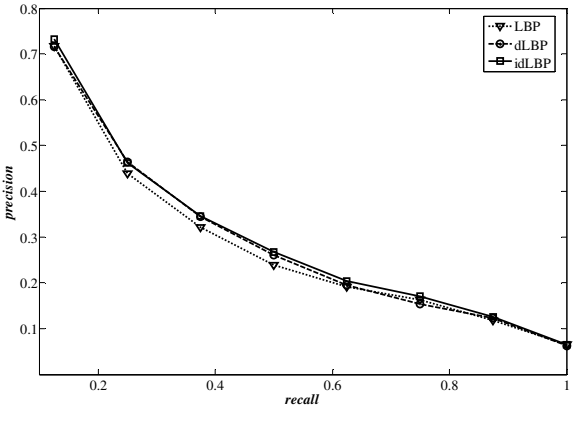


Fig. 6. Comparison on the third image set

4 Summary

This paper points out the shortcoming of dLBP, and gave an extension idLBP. Because of the considering of more directions in a local region, more information is fused in the new descriptor. The three methods, LBP, dLBP and idLBP were compared in image retrieval experiments, and three commonly used image sets were chosen as test beds. Experimental results show that idLBP is very robust and can

achieve significant performance improvement over the LBP and dLBP. The methods mentioned in the paper are also effective for electric trace fossil image.

Acknowledgements. This work was supported by the Key Project of Chinese Ministry of Education (210128), The Backbone Teacher grant of Henan Provience (2010GGJS-059), Image Processing & Image Communication Lab of Jiangsu Provience (LBEK2011002).

References

1. Ojala, T., Pietikäinen, M., Hardwood, D.: A Comparative Study of Texture Measures with Classification based on Feature Distribution. Pattern Recognition 1, 51–59 (1996)
2. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution Gray-scale and Rotation Invariant Texture Classification with Local Binary Patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence 7, 971–987 (2002)
3. Zhou, H., Wang, R.S., Wang, C.: A Novel Extended Local-Binary-Pattern Operator for Texture Analysis. Information Sciences 22, 4314–4325 (2008)
4. Guo, Z., Zhang, L., Zhang, D.: Rotation Invariant Texture Classification using LBP Variance (LBPV) with Global Matching 3, 706–719 (2010)
5. Tan, X., Triggs, B.: Enhanced Local Texture Feature Sets for Face Recognition under Difficult Lighting Conditions. IEEE Transactions on Image Processing 6, 1635–1650 (2010)
6. Sun, J.D., Wu, X.S.: Content-based Image Retrieval based on Texture Spectrum Descriptors. Journal of Computer-Aided Design & Computer Graphics 3, 516–520 (2010) (in Chinese)
7. Heikkilä, M., Pietikäinen, M., Schmid, C.: Description of Interest Regions with Local Binary Patterns. Pattern Recognition 3, 425–436 (2009)
8. Trefný, J., Matas, J.: Extended set of local binary patterns for rapid object detection. In: Computer Vision Winter Workshop 2010, Libor Spacek, Vojtech Franc, pp. 1–7 (2010)

Research on the Synesthesia-Based Conversion for Electronic Digital Image and Digital Audio

Yu Zhang[1,2]

[1] School of Electronic and Electrical Engineering
Shanghai University of Engineering Science
Shanghai 201620, China
zy@sues.edu.cn
[2] Department of Electrical and Computer Engineering
The University of Michigan-Dearborn
Dearborn, MI 48128-2406, U.S.A.
yuzhng@umd.umich.edu

Abstract. Synesthesia is a psychological phenomena that the sense caused by other stimulated sense organs. The most common synesthesia is "colored - listen" synesthesia, the sense of color can lead to the appropriate the sense of hearing. On the contrary, the sense of sound can cause a corresponding vision. In this paper, it has been studied the synaesthesia phenomenon on audio and images, then the software on electronic digital audio and digital image format conversion software has been implemented. The further work will be on doing ahead work for the designing the software conversing the voices of different people into different personalize patterns in electronic image and voices processing area.

Keywords: synesthesia, digital audio, digital image.

1 Introduction

The audio and art with hearing and vision are determined art features and effects, while the vision is a show of hearing and the hearing makes the vision more rich and vivid. Although the hearing and the vision can be sensed the mental activity and emotional activity that difficult to tell with language, as well as the abstract thinking and meaning, the sisters relationship between the hearing and the vision let more people experience the strange realm of art. There is a basic properties with non-visual and non-semantic in audio, but while people are listening to the audio, they can also be evoked the feeling of a rich vision by the feeling of appreciation the beauty to audio. The stimulation of a sensory organ can cause the feeling of the other senses, the psychological phenomenon is known as "synesthesia". "Synesthesia" is occurred on the sensory organs' level of the feeling, so it is a common psychological reaction for everyone.

As a means of modification, since ancient times, from the West to the East, many people have been used Synaesthesia. Synaesthesia was used as a creative means by

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 481–486.
springerlink.com

many poets. "Apricot branches spring trouble", "Numerous Willow had passed shadowless". In many countries, the "synesthesia" has had been researched beyond a "modified" field. It has been researched its nature in psychology, sociology and neuro-linguistics. In recent years, people confirmed through PET (Positron Emission Tomography) that each area of the human brain is responsible to a sensory component, that makes the voice belongs to auditory area, rather than visual or olfactory area. However, sometimes it is fused in two different areas, so you can see the note in color and hear the taste of potatoes, and also you can immediately identify the same set of numbers in many numbers, since these numbers can be told out from the color. This is the phenomenon of synesthesia.

As far as you know, it is not that common but the most common form of synesthesia, researchers believe, is colored hearing sounds, music or voices seen as colors.

In the human experience, there is potential relationship between hearing with high frequency sounds and vision with bright colors, which gives a "rising feeling". In the other hand, there is "sinking feeling" for hearing with low-frequency sound and vision with dark colors. The more tone is "low"', the more color experience is "deep". Otherwise, the more "high" tone will cause 'light' bright feeling. In this paper, it is studied corresponding relationship between image and audio with certain science and art rules, proposed the corresponding algorithm, using C # as the development software, the digital audio and digital image format conversion software has been implemented.

2 Conversion of Electronic Digital Image into Digital Audio

2.1 Methods and Implementation of Conversion

First, the image file were converted into 24bit color BMP file format, analyzed the BMP file, and saved in an array, then according to the correspondence between these data, changed them into WAV format audio file and saved it. Finally, the audio can be played on its broadcast interface.

The Bitmap file format is used for bitmap graphics on the Windows platform only. Unlike other file formats, the BMP format is saved image data from bottom to top and pixels in blue/green/red order. Compression of BMP files is not supported, so they are usually very large. For example, in 24-bit mode, three bytes are represented one pixel. The first byte represents the red part, the second the green and the third the blue respectively. So every pixel is contained a literal RGB-value, Here, Bitmap file format is converted to HSL file format. HSL file format is tried to describe more accurately than RGB color, which in expression of color perception and remain computationally simple. HSL means the hue (hue), saturation (saturation) and lightness (brightness), The formula conversed the RGB into HSL is as follow:

$$h = \begin{cases} 0° & \text{if max} = \text{min} \\ 60° \times \dfrac{g-b}{\max - \min} + 0° & \text{if max} = \text{r and g} \geq \text{b} \\ 60° \times \dfrac{g-b}{\max - \min} + 360° & \text{if max} = \text{r and g} < \text{b} \\ 60° \times \dfrac{b-r}{\max - \min} + 120° & \text{if max} = \text{g} \\ 60° \times \dfrac{r-g}{\max - \min} + 240° & \text{if max} = \text{b} \end{cases}$$

$$S = \begin{cases} 0 & if\ l = 0\ or\ \max = \min \\ \dfrac{\max - \min}{\max + \min} = \dfrac{\max - \min}{2l} & if\ 0 < l < \dfrac{1}{2} \\ \dfrac{\max - \min}{2 - (\max + \min)} = \dfrac{\max - \min}{2 - 2l} & if\ l > \dfrac{1}{2} \end{cases}$$

$$l = \tfrac{1}{2}(max + min)$$

Let (r, g, b) is the coordinates of the color red, green and blue respectively, their values are a real number between 0-1. Let max is equivalent to the biggest one among r, g and b, while min is the minimum of these values. To find the value in the HSL space (h, s, l) of h ∈ [0, 360) is the hue, s and l ∈ [0,1] is the saturation and brightness.

The value of h is usually standardized to between 0 to 360°, and h = 0 means max = min (that is gray) ,rather than let h is undefined.

Because HSL and WAV are all composed of bits, according to the experience of person, There are potential relationship in hearing with high-frequency sound and visual with bright colors, which rising feeling are given to person, and hearing with low-frequency sound and vision with the dark hue, which sinking feeling are given. Found out the conversion relation between the HSL and WAV file, The implementation coding is as following:

```
public static int[] Change(byte[,] rgb)
      {
          int[]  result  =  new  int[rgb.GetLength(0)  *
       (rgb.GetLength(1) / 3)];
          int location = 0;
          for (int i = 0; i < rgb.GetLength(0); i++)
          {
            for (int j = 0; j < rgb.GetLength(1); j += 3)
            {
                int[] tmp = rgbToHsl(rgb[i, j], rgb[i, j
       + 1], rgb[i, j + 2]);
                result[location++] = tmp[0] + 128;
             }
          }
            return result;
        }
```

2.2 The Design of Audio Player Module

The design of the audio module is for the play of the converted audio. The main functions are included six modules: Open Audio Files, Play, Pause, Stop, Time Display and Volume. playing sound waveform is shown in Fig.

Fig. 1. Sound waveform

3 Conversion of Electronic Digital Audio into Digital Image

The focus of this study is the conversion of the digital audio in WAVE format into the digital images in BMP format.

Using HSL color model is rather easy to find out relationship between digital audio in WAVE format with synesthesia and digital images in HSL color model, where H represents hue, S represents the saturation, L represents the brightness respectively. Since the people have a Synesthesia of "color - listen", it can be easily thought of soft audio as a touch of yellowish color, muffled audio as a gray or dim colors. Whereas the same purpose with the HSL theory. Therefore, through the WAVE format was converted into HSL format, then into the RGB format. WAV file can be converted into the final effect, there is no trace can be found between WAVE format and HSL format . In order to find some linkages which may exist, a large number of experiments have been undergone. WAVE information was obtained from DATACHUNK block, which is consisted of a 8bit binary. HSL is also Consisted of a 8b converted binary. As H is the hue, it is a key to decide on a color, the first step is WAVE format digital audio be converted to digital image with HSL color model. The coding conversion of each DataChunk into the hue H are as follows:

```
/// <summary>
/// WAVE'DATA convert into H
/// </summary>
/// <param name="w">Wave Instance</param>
/// <returns>returns H</returns>
 private static int[] Waveto H(BigBang.Wave w)
 {
    int[] H = new int[w.Data.Data.Count];
   for (int i = 0; i < w.Data.Data.Count; i++)
   {
```

```
            H[i] = w.Data.Data[i];
        }
        return H;
    }
```

To be converted HSL into RGB format, the value of the definition of a color is in HSL space (h, s, l) , while h(hue angle) is directed in the range [0, 360), s(saturation) and l(brightness) in the range [0, 1] respectively. Corresponding with the three primary colors in the RGB space (r, g, and b) respectively. Corresponding to the red, green and blue, r, g and b are also in the range [0, 1]. The algorithm is as follows:

$$q = \begin{cases} l \times (1 + s), & \text{if } l < \frac{1}{2} \\ l + s - (l \times s), & \text{if } l \geq \frac{1}{2} \end{cases}$$

$$p = 2 \times l - q$$

$$h_k = \frac{h}{360} \text{ (h normalized to the range [0,1))}$$

$$t_R = h_k + \frac{1}{3}$$

$$t_G = h_k$$

$$t_B = h_k - \frac{1}{3}$$

$$\text{if } t_C < 0 \rightarrow t_C = t_C + 1.0 \quad \text{for each } C \in \{R, G, B\}$$

$$\text{if } t_C > 1 \rightarrow t_C = t_C - 1.0 \quad \text{for each } C \in \{R, G, B\}$$

For each color vector, Color = (ColorR, ColorG, ColorB) = (r, g, b),

$$Color_C = \begin{cases} p + ((q - p) \times 6 \times t_C), & \text{if } t_C < \frac{1}{6} \\ q, & \text{if } \frac{1}{6} \leq t_C < \frac{1}{2} \\ p + \left((q - p) \times 6 \times (\frac{2}{3} - t_C)\right), & \text{if } \frac{1}{2} \leq t_C < \frac{2}{3} \\ p, & \text{otherwise} \end{cases}$$

$$\text{for each } C \in \{R, G, B\}$$

If s = 0, the result is the color of non-colored or gray. In this particular case, r, g and b are equal to l. Pay attention to the value of h in this case is undefined.

So far, it has been completed the conversion of a synaesthesia-based electronic digital audio to digital images. Fig.2 shows converted electronic images from the QQ on-line voice.

4 Conclusions

In this paper, it had been made a try on the realization of the synesthesia-based conversion software for electronic digital image and digital audio format. There are certain practical significance for such a design to help the blind to realize self life inspired. This will allow more things which usually can not be seen by the blind to be converted into the voices they can hear so they can feel the changes in the surrounding environment and other things. Of course, the algorithm is also a little rough in this paper. There is still a long way to go to complete the corresponding relationship among all colors and all kinds of sound. The further work is on trying to

find a relationship of synesthesia with a variety of colors corresponding to the specific sounds, so it can be resolved fundamentally. Through a lot of trial and application, breakthrough development will be made of the awareness and research on synesthesia-based, which will play an important role in the field of computer science, information science, electronic image processing, statistics, biology, medicine, entertainment. It has benefit for both social and mankind.

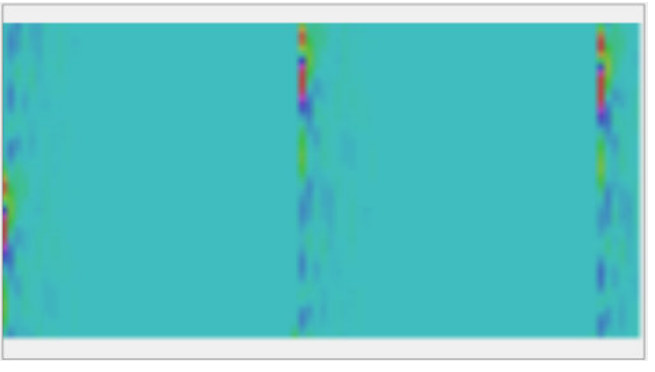


Fig. 2. Converted electronic images from QQ on-line voice

Acknowledgments. This work is supported by Scientific and Technical Development Foundation of SUES (Grant No.2011XY19).

References

1. Zhang, Y., Shi, Y., Wu, J.: The Research and Implementation on the Conversion of Image File Format. Electrical Automatics (12) (2008) (China)
2. Zhang, Y.: The Research and Implementation on the Conversion of Audio File Format. Microcomputer Application (12) (2008) (China)
3. Wikipedia. Synesthesia (OL), http://en.wikipedia.org/wiki/Synesthesia
4. Graphics and Mutimedia (OL) (2004), http://msdn.microsotf.com/library/
5. Vincent, L.: From Images to Sounds, a Dual Representation. Computer Audio Journal 20(3), 59–69
6. Pfeiffer, S., Fischer, S., Effelsberg, W.: Automatic Audio Content Analysis. In: Proceedings of the ACM International Multimedia Conference & Exhibition (1996)

Design of Embedded Electronic Video Processing System Based on DM6467

Chunlian Yao, Keyou Guo, and Huming Jiang

Beijing Technology and Business University, Beijing, China
demons1217@sina.com

Abstract. In view of real-time electronic image or video processing system and current situation of video compression algorithm, an embedded video processing system based on TMS320DM6467 is provided. TMS320DM6467 is a dual-core processor, ARM core and DSP core, and it uses Davinci technology. DSP core is used to encode video with H.264 standard, ARM core is used for video transmission and store. By optimization of software framework and compression algorithm, the system can real-time process 720×480 HD (High Definition) video and store compression data into USB-Disk. Especially, this system can process multi-channel video for some other application.

Keywords: Video processing DaVinci DM6467 H.264 Motion estimation.

1 Introduction

With the technology development of network and multimedia technology, video compression, analysis, and transmission has been gradually widely used. The performance of small size, low power and high adaptability is commonly required by some application. With the advantage of flexibility, expansibility, portability, upgradeable and controllable, General DSP is very suitable for video processing.

For a video system processing, it is very important to support high performance video coding standard, such as H.264 standard. And when we design a video processing, we should think about how to reduce product cost, increase video channel, reduce the developer's difficulty, and make intelligent system.

This paper put forward an embedded real-time video processing system based on DM6467 DSP. This system can capture and process HD video. With the two A/D chips, this hardware platform can capture 8-channel SD(Standard Definition) video or two HD channel video. And it can implement co-operation of DSP core and ARM core. High-speed SDRAM is used to transfer video data. USB-Disk driver is used to store compression data. With several kinds of optimizing technique, such as software framework, encoder algorithm, the system can compress one HD channel video data (the resolution is 720×480), compress time is 20ms per frame.

2 Hardware Architecture

TMS320 DM6467 is a dual-core processor, including ARM926EJ-S subsystem and TMS32064x+ DSP subsystem. DM6467 is to realize system control and video

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 487–492.
springerlink.com

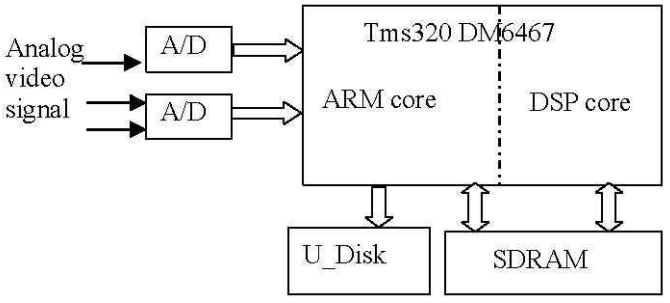


Fig. 1. Hardware Framework of video processing system

compression algorithm. As is shown in figure 1, this system has three parts, capture, coding, and store.

1) Capture part: which is consists of ARM core of DM6467 and two A/D chips. A/D chip convert analog video into digital video, and two A/D chips can capture 8-channel SD(Standard Definition) video or 2-channel HD video. ARM is used to store video data captured by A/D chip into SDRAM chip. 2) Coding part: which is consists of DSP core of DM6467 and SDRAM. DSP core is used to video compression; SDRAM chip is used to data store for reference frame, construction frame, and compress bit-rate. 3) Store part: Store part is consists of ARM core, SDRAM, U-Disk, Which store system initialization code, video processing code and video data.

3 System Software Design

First, the video processing system realizes the capture of video through V4L2(video for liunx) interface. Then, system realizes H.264 compression encoding based on DaVinci technology architecture, which is a key technology for the system implementation. DSP algorithm needs to be in accordance with xDM standard, and to satisfy the realtime requirement,we need to optimize the H.264 coding algorithm.

3.1 Capture Processing

TMS320 DM6467 chip can support Linux and WinCE embedded operating system. For features and prices considerations, system chooses MontaVista Linux v4.0 as embedded operating system, which with the advantage of high real-time, multi-task concurrency etc. MontaVista Linux can make full use of system resource, and fulfill multi-task scheduling.Video capture module acquires original video data through V4L2 interface. The process of video capture is shown as figure 3.

3.2 H.264 Encoding Server

We realizes H.264 encoding server based on the TI DaVinci technology, and has finished the integration about Codec server and Codec Engine. The application program uses this encoding server through the calling of VISA APIs. The process of H.264 video compression encoding using Codec Engine is shown in figure 4.

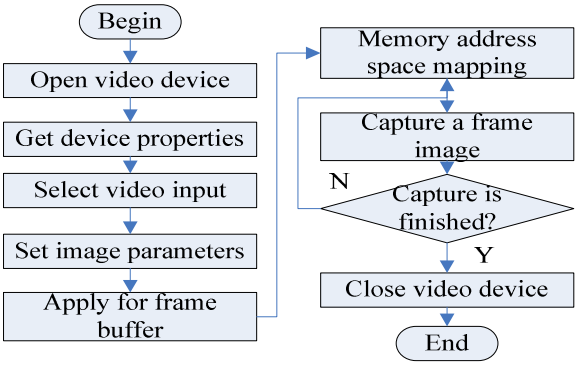


Fig. 3. Video capture process

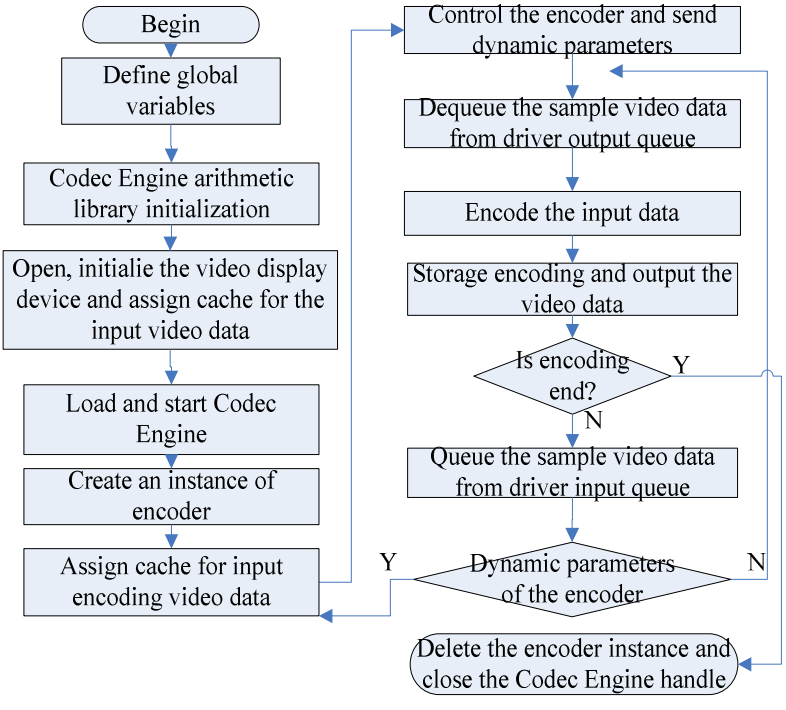


Fig. 4. The process of H.264 video compression encoding using Codec Engine

3.3 Optimization of Algorithm in H.264

H.264 coding standard uses new encoding technology, such as, intra prediction, multi-frame reference prediction, accuracy motion estimation etc, and has improved the efficiency of data encoding.In order to reduce computation complexy,we optimize the motion estimation(ME) algorithm of H.264. ME is efficient to eliminate redundancy between different frames, yet on the other hand, it is the most computationally

intensive part. ME is made up of two stages, inter-pixel accuracy motion estimation (IME) and sub-pixel accuracy (SME). IME takes coded frames as reference frames and searches the best match block in them directly. IME process is shown in Fig.5. Based on the result of IME, SME gets more accurate reference frames by interpolating samples at sub-pixel positions and tries to find a better match for the current block by searching the interpolated samples around IME result, which improves performance of compression, but increases computation and memory cost sharply. The $1/2^K$ sample accuracy SME, adopted by international standards widely, generally needs to check $8\times K$ blocks at sub-pixel positions in addition to IME. Together with the interpolation computation cost, SME takes largish weight among the whole process of ME. Besides, due to interpolation image expanding to $2^{2K}-1$ times larger than input image, additional memory has to increase $2^{2K}-1$ times. In order to solve the store and computation problem, A fast SME is proposed, which uses the result of IME as input value, and then calculate the best match point of sub-pixel, thus the algorithm can avoid search computation and store data.

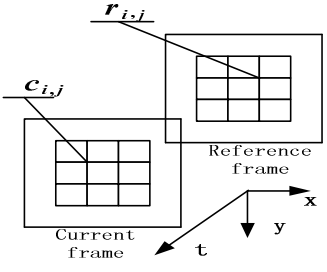


Fig. 5. Integer pixel motion estimation

As shown in Fig.5, the point of the current block $c_{i,j}$, $i\in(0,M)$ $j\in(0,N)$, M,N is the horizontal and vertical direction pixel number of the block, the best match point of IME is $r_{i,j}$, the half pixel of $r_{i,j}$, $r_{i+1,j}$ is $h_{i,j}$,as shown in Fig.6. MSE is the match criteria, the MSE of two point $d_{x,y}$ and $d'_{x,y}$ is $D_{d_{x,y},d'_{x,y}}$.

With Bi-linear interpolation, we can draw:

$$h_{i,j}=\frac{1}{2}\left(r_{i,j}+r_{i+1,j}\right) \tag{1}$$

From IME, we can know that:

$$D_{c_{i,j}\ r_{i,j}}=\frac{1}{MN}\sum_{i=0}^{M}\sum_{j=0}^{N}(c_{i,j}-r_{i,j})^2 \tag{2}$$

$$D_{c_{i,j}\ r_{i+1,j}}=\frac{1}{MN}\sum_{i=0}^{M}\sum_{j=0}^{N}(c_{i,j}-r_{i+1,j})^2 \tag{3}$$

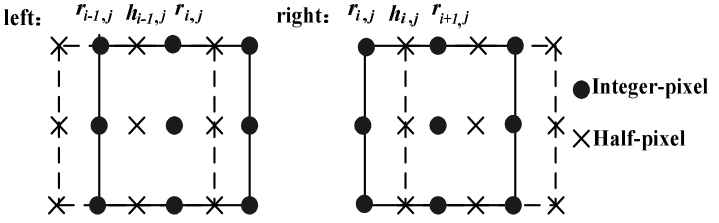


Fig. 6. Horizontal motion of half pixel in reference frame

The MSE value of SME is:

$$D_{c_{i,j}\ h_{i,j}} = \frac{1}{MN}\sum_{i=0}^{M}\sum_{j=0}^{N}(c_{i,j} - h_{i,j})^2$$

With formula (1),(2),(3) we can get that:

$$D_{c_{i,j}\ h_{i,j}} = \frac{1}{2}D_{c_{i,j}\ r_{i,j}} + \frac{1}{2}D_{c_{i,j}\ r_{i+1,j}} - \frac{1}{4}D_{r_{i,j}\ r_{i+1,j}} \tag{4}$$

Then we can draw that:

$$D_{c_{i,j}\ h_{i-1,j}} = \frac{1}{2}D_{c_{i,j}\ r_{i-1,j}} + \frac{1}{2}D_{c_{i,j}\ r_{i,j}} - \frac{1}{4}D_{r_{i-1,j}\ r_{i,j}} \tag{5}$$

From (4) and (5), we can draw that:

$$\begin{aligned} &D_{c_{i,j}\ h_{i,j}} - D_{c_{i,j}\ h_{i-1,j}} \\ &= \frac{1}{2}D_{c_{i,j}\ r_{i+1,j}} - \frac{1}{2}D_{c_{i,j}\ r_{i-1,j}} - \\ &\frac{1}{4MN}\left(\sum_{i=M}^{M}\sum_{j=0}^{N}(r_{i,j} - r_{i+1,j})^2 - \sum_{i=0}^{0}\sum_{j=0}^{N}(r_{i,j} - r_{i-1,j})^2\right) \end{aligned} \tag{6}$$

In (6), $D_{c_{i,j}r_{i+1,j}}$、$D_{c_{i,j}r_{i-1,j}}$ is the medium result of IME.

Then the one-dimension best match point x is:

$$x = \begin{cases} -\frac{1}{2} & if(D_{c_{i,j}\ h_{i,j}} \geq D_{c_{i,j}\ h_{i-1,j}}) \\ \frac{1}{2} & f(D_{c_{i,j}\ h_{i,j}} < D_{c_{i,j}\ h_{i-1,j}}) \end{cases} \tag{7}$$

With (6) and (7), we can calculate the half pixel point position which has the least MSE between current block and horizontal half pixel block.

The PSNR and coding time of SME and the method of the paper are shown in table 1. We select three typical sequences provided by MPEG standard organization, basket, Flower, Coastguard. CIF format (the resolution is 352×288).We encode 100 frames for each sequence, and the structure of GOP is I, PPP.

Table 1. Test result

sequence	SME PSNR(dB)	Algorithm proposed PSNR(dB)	Reduce time (%)
Basket	27.65	27.62	50.48
Flower	25.88	25.84	59.06
Coastguard	30.21	30.20	43.04

As shown by the experimental results, the PSNR value of the two method is same, but the method of the paper reduce times about 40~60％ and avoid restore consummation of half pixel interpolation.

4 Conclusion

In this paper, An embedded electronic image and video processing system is provided based on dual-core DM6467 processor. The electronic system has the three main functions,capture and coding 720×480p video, and store compress data into U-disk.We configure and program of the whole software at application layer, that avoids the tedious implementation details. With the sub-pixel motion estimation proposed in the paper, the resource utilization and program code optimization, the video quality and real-time is satisfied.

Acknowlegement. This work is supported by National Nature Sciense Fund (No.61103124).

References

[1] Lappalainen, V., et al.: Performance of H.26L video encoder on general-purpose processor. Journal of VLSI Signal Processing Systems 34(3), 239–249 (2003)
[2] Veigh, J.M., et al.: A software-based real-time MPEG-2 video encoder. IEEE Transactions on Circuits and Systems for Video Technology 10(7), 1178–1184 (2000)
[3] Benkrid, K., et al.: High level programming for FPGA based image and video processing using hardware Skeletons. In: IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 3227–3231. IEEE Press, Istanbul (2000)
[4] Zhao, B.-J., et al.: Implementation of real-time 2D-DCT with FPGA and DSP. Acta Electronica Sinica 31(9), 1317–1319 (2003) (in Chinese)
[5] Peng, Q.: DaVinci technology-new digital image/video signal processing platform, pp. 1–7, 36–60, 126–140. Electronic Industry Press, Beijing (2008)
[6] Texas Instruments. TMS320DM6467 Digital Media System-on-Chip (2007)
[7] Li, B., et al.: A fast block matching algorithm using smooth motion vector filed adaptive search technique. Journal of Computer Science and Technology 18(1), 14–21 (2003)

Joint Time-Frequency Analysis of Magnetic Storms during February 1999

Yihong Li, Wenliang Guan, and Chao Niu

Xi'an Research Inst. Of Hi-Tech, Xi'an, China
gyihongli@126.com

Abstract. Temporal characteristic and joint time-frequency characteristic of two magnetic storms during February 1999 was analyzed in this paper based on signal processing method. The rate of change of the geomagnetic signal time-domain waveform and its auto-correlation and auto-covariance were analyzed in the time domain. Empirical Mode Decomposition (EMD) and Short Time Fourier Transform (STFT) were introduced in time-frequency domain to study the morphology of the magnetic storm characteristics and development process. The results show that more information of magnetic storm development and evolution can be provides through time-frequency analysis. It is a good idea for the magnetic storm research from the view of signal processing and electronic technology.

Keywords: time-frequency analysis, magnetic storms, EMD, electronic technology.

1 Introduction

Magnetic storm is a kind of severe global geomagnetic disturbance with characteristic of large scale variation of amplitude, complex morphology, long duration and global synchronization [1]. Magnetic storm, a kind of disastrous space weather, is mainly due to the outbreak activity of the sun. When magnetic storm happened, navigation, communications, satellites and power systems, oil pipelines and other technology systems would be harmed.

According to evolution in morphology, magnetic storms can be divided into several stages, such as initial phase, main phase and recovery phase. According to increase speed of the geomagnetic horizontal component, magnetic storms can be divided into sudden commencement (SC) and gradual commencement (GC).

Magnetic storms began almost simultaneously in the global. The feature of SC is that the positive pulse of horizontal component appears suddenly. The variational rate of the impulse is generally 10~20nT and the maximum can be up to more than 50nT. Sometimes, there is a small negative pulse before the positive pulse and the sudden commencement is denoted by the SC*. The feature of GC is relatively gentle. Magnetic storm is divided into medium common magnetic storm (MC), medium strong magnetic storm (MM), strong magnetic storm (MS). Magnetic storm is a hot research field of geomagnetism, space weather, etc. The studies include storms physical causes, morphology analysis, analysis of typical magnetic storm and ionosphere disturbances during magnetic storms.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 493–498.
springerlink.com

2 Data and Methods

Ten minutes mean of horizontal component H of Geomagnetic three components recorded at Beijing Geomagnetic Station (40°02′22″N, 116°10′30″E) in 1999 is used in this paper. The data used is original H minus by 29000nT. In addition, the time of this paper is world time. According to the magnetic storm records, two magnetic storms were broken out continuously during February 17 to 18, 1999, it recovered to calm gradually on the 20th. Both magnetic storms are sudden commencement. The first storm started at 7:09 on the 17th, the disturbance intensity is MS, and the maximum activity K index is 5. The second storm started at 2:45 on the 18th, the disturbance intensity is M, and the maximum activity K index is 7. The temporal wave of H component of Magnetic storms on 1999.2.17 and 1999.2.18 recorded at Beijing Geomagnetic Station (BJI) is shown in Fig.1.

The morphology of magnetic storm is complex and changeable and H component time-varying sequence has the features of nonlinear and nonstationary from the perspective of signal processing. The variation characteristics can not be represented clearly from the time domain analysis only. It can obtain more information combined with frequency domain analysis. Yingyan Wu [3] used Natural Orthogonal Components (NOC) Analysis, Correlation Analysis and Fourier analysis to separate storm-time variation (Dst), solar daily variation (Sq) and disturbance daily variation (SD) through three steps respectively. The rate of variation of the geomagnetic signal time-domain waveform and its auto-correlation and auto-covariance were analyzed in the time domain in this paper. EMD and STFT were adopted to analyze the H component during the magnetic storm in time-frequency domain to comprehend geomagnetic horizontal intensity varying with time and frequency within magnetic storm evolution process more carefully.

The definition of the variational rate of the geomagnetic horizontal component H is $rateH_n = [h(n+k) - h(n)]/k$, where k=10. The variational rate reflects the speed of H changes with time. Auto-correlation and auto-covariance describe the statistical characteristics of the signal in the time domain. When the mean is zero, auto-correlation is equal to auto-covariance function.

EMD was proposed by N. E. Huang [2], it is an self-adaptive signal processing method. Since the EMD is introduced, it has been used in geophysics [2] [3], such as the seismic signal analysis [4], oceanic waves [5], climate [6] etc. Signal can be into decomposed into multiple Intrinsic Mode Functions (IMFs) components by EMD method.

STFT is a traditional time-frequency analysis method. The basic principle is to get the local frequency spectrum of a part of the signal by moving a very narrow window, and find its Fourier transform. STFT of each IMF is the joint time-frequency distribution of the IMF. It is more conducive to understand the characteristics of each band component of the geomagnetic signal changing with time during Magnetic Storm.

3 Analysis of Results

3.1 Time-Domain Analysis

Time domain waveform of H component from1999-02-16-00-10 to 1999-02-20-23-50 is shown in Fig.1. We can see that the geomagnetic data in 17th February appeared minor disturbance, but appeared violent disturbances in 18th February. The disturbance in 17th February began at 7:09, and stopped at 22:00 according to the storms report. The total duration is short, but the geomagnetic disturbance showed characteristics of the storms obviously. The development process contained the sudden commencement, the initial phase, main phase and the recovery phase and had obvious time limit. The variation amplitude of sudden commencement was up to 17nT, the start time of main phase is 8:30, and the biggest drop is 97nT. The disturbance in 18th February began at 2:45. The variation amplitude of sudden commencement pulse was up to 60nT. After the sudden commencement, the disturbance declined rapidly. There was a 15nT positive phase pulse after the disturbance declined approximately 110nT at 5:30 on the 18th. And then, the disturbance continued to decline to a minimum on the 18th 10:20. A total decrease was up to 260nT in the main phase, during which it oscillated repeatedly. After main phase, the recover phase began and it restored to calm on the 20th. The entire disturbance lasted 4 days from the 17th to the 20th, with general characteristics of long duration and repeated oscillations. The medium common magnetic storm could be considered as a prelude before a medium strong magnetic storm.

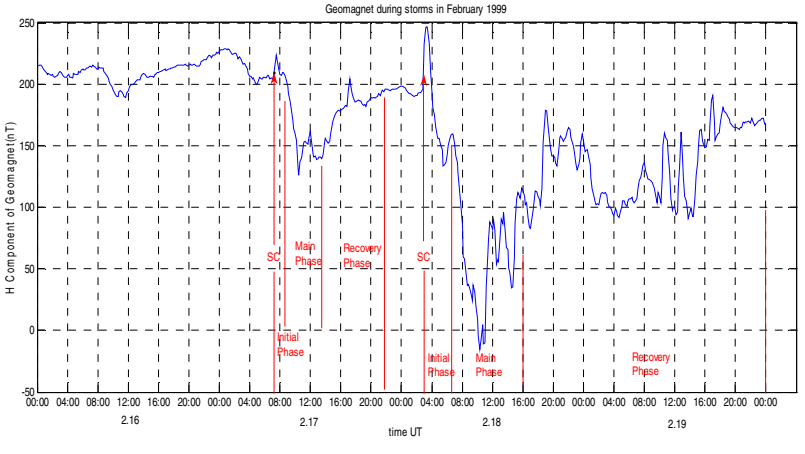


Fig. 1. Temporal wave of H component of Magnetic storms on 1999.2.17 and 1999.2.18

The variational rate of 10 minutes mean the H component, the auto-correlation function and the auto-covariance function of two storms are shown in Fig.2 and Fig.3 respectively. We can see that the variational rate corresponded well at the time of the sudden commencement and main phase, the auto-correlation sequence at the two magnetic storm sudden commencement time and main phase appeared jitters and the covariance reached its minimum during the the main phase.

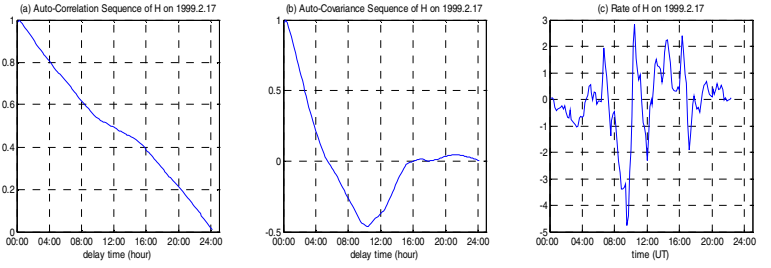


Fig. 2. Auto-correlation sequence, auto-covariance sequence and rate of H component on 1999.2.17

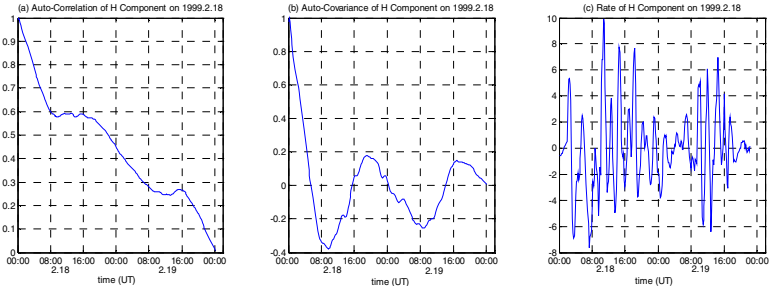


Fig. 3. Auto-correlation sequence, auto-covariance sequence and rate of H component on 1999.2.18

3.2 Frequency-Domain Analysis

The 10 minutes mean data of geomagnetic H component of the two magnetic storms were decomposed into, five IMFs and one remainder term base on EMD method, which were shown in Fig.4. The morphology of each IMF was significantly different in time domain waveform. IMFs contained different frequency components, which ranked from high to low according to the order of the decomposition. The joint time-frequency spectrum of the first four IMFs was obtained based on STFT, which were shown in Fig.5. The frequency of IMFs ranked from high to low. Each IMF showed a different time-frequency characteristics. It can be seen that the spectrum distribution of IMFs of the two storms are very different. The contrast analysis was as follows:

(1) IMFs time-frequency analysis of Magnetic storm broke out on 17th February.

Spectral energy of $imf_1 \sim imf_4$ concentrated in the main phase stage, which corresponding normalized frequency center was 0.145, 0.06, 0.03 and 0.01.

(2) IMFs time-frequency analysis of Magnetic storm broke out on 18th February.

Spectral energy of imf_1 was concentrated in the normalized frequency center of the 0.05 in sudden commencement phase and of 0.1 in recovery phase. Time-frequency distribution of imf_2 was more complex. The spectral energy was concentrated in the normalized frequency of 0.065 in the sudden commencement, of 0.09 in main phase

lowest point, and of 0.056 and 0.059 in recovery phase severe variational points. Spectrum energy of imf_3 was concentrated in normalized frequency of 0.033 in main phase. Spectrum energy of imf_4 was concentrated in normalized frequency of 0.01 in the sudden commencement and main phase lowest point. The sudden commencement was reflected in imf_1, imf_3 and imf_4, but the frequency was different. The lowest point of the main phase was reflected in imf_2, imf_3 and imf_4, and the frequency was different either. The whole main phase is reflected in imf_3. The dramatic disturbance of recovery phase was reflected in imf_1 and imf_2.

Overall, the magnetic storm occurred on 17th February had shorter duration time, less disturbance amplitude and simpler variation morphology than on 18th February.

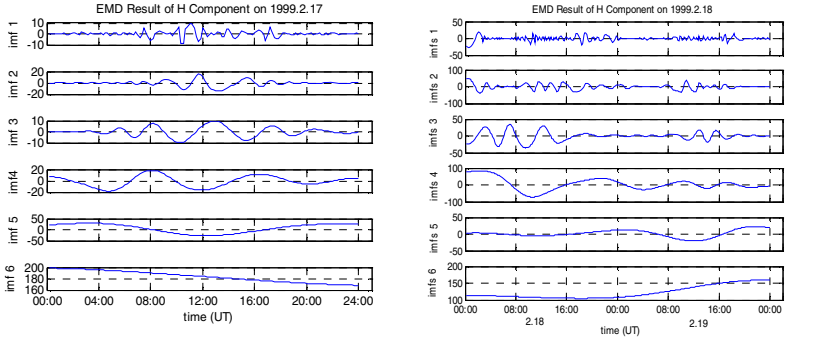


Fig. 4. EMD results of H component for magnetic storm on 1999.2.17 and 1999.2.18

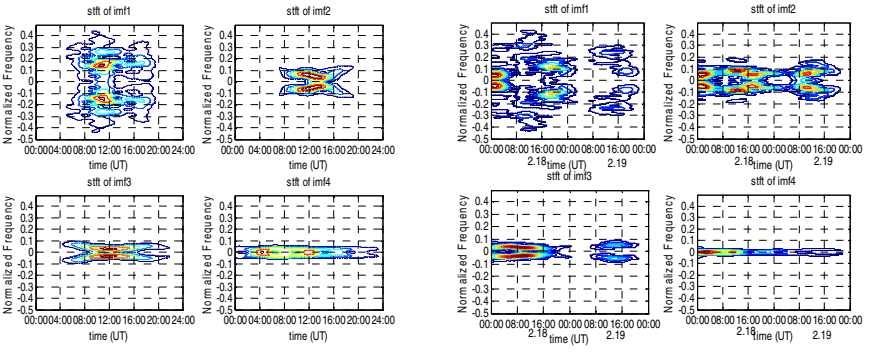


Fig. 5. STFT of the first four IMFs of H component for magnetic storm on 1999.2.17 and 1999.2.18

4 Conclusion

Two sudden commencement magnetic storms were broken out continuously during February 1999, but their disturbance intensity and variation morphology were quite different. Statistical analysis and time-frequency analysis methods which are signal processing methods and electronic technology were used to do comparative analysis of time domain and frequency domain characteristics of two magnetic storms during February 1999. The magnetic storm on 17th had shorter duration time and less the disturbance amplitude than on 18th.from the time-domain. From the view of time-frequency domain, the first four IMFs on 17th were mainly reflected different frequency components of the main phase. On the 18th, the first four IMFs frequency component is relatively complex and the frequency components of various stages were reflected.

References

1. Xu, W.: Earth physics of electromagnetic phenomena. University of Science and Technology of China Press (2009)
2. Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C., Liu, H.H.: The Empirical Mode Decomposition and The Hilbert Spectrum for Nonlinear and Non-stationary Time Series Analysis. Proc. Roy. Soc. London A 454, 903–995 (1998)
3. Huang, N.E., Zheng, S., Steven, R.: A New View of Nonlinear Water Waves: The Hilbert Spectrum. Annu. Rev. Fluid Mech. 31, 417–457 (1999)
4. Wu, C., Zhou, R.Z.: Application of Hilbert-Huang transform in extracting dynamic properties of seismic signals. Earthquake Engineering and Engineering Vibration 26(5), 23–28 (2006)
5. Huang, N.E., Shin, H.H., Long, S.R.: The Ages of Large Amplitude Coastal Seiches on The Caribbean Coast of Puerto Rice. Physical Oceanography 30(8), 2001–2012 (2000)
6. Wan, S.Q., Feng, G.L., Dong, W.J., et al.: On The Climate Prediction of Nonlinear and Nonstationary Time Series With the EMD Method. Chinese Physical 14(3), 628–633 (2005)

Research on 1-D Integer SDCT Algorithm and the Application in Electronic Image Processing

Wenbang Sun[1], Haiyan Tang[2], Yan Chang[3], and Fei Guo[1]

[1] Aviation Information Department, Aviation University of Air Force
130022 Changchun, China
[2] Aviation theory Department, Aviation University of Air Force
130022 Changchun, China
[3] Unit 65042, PLA
110001, Shenyang, China
{chwenbang,feer_2403}@163.com, chhairong@126.com,
chang_y040@sina.com

Abstract. With the development of electronic image processing technology, DCT has been used more and more widely and the research on DCT algorithm is very much too. But the transform matrix of DCT is expressed with floating-point numbers, so the computational complexity is high and more system resources are occupied. To overcome this drawback, 1-D integer SDCT (Submatrix Discrete Cosine Transform) was proposed. First, several matrix operation methods were defined. Then, the basic principle of 1-D integer SDCT was deduced in detail. The theoretical analysis and experimental results show that 1-D integer SDCT is convenient to electronic image processing.

Keywords: 1-D Integer SDCT, 1-D DCT, Integer Transform Basic Matrix.

1 Introduction

Nowadays, the DCT is applied widely in the field of digital signal processing, and especially the DCT has become the core part of motion or static image compression standard, such as JPEG, MPEG, H.26x etc. Because of the widespread application, the research on DCT algorithm is very much. Such as, Chen [1] firstly proposed a fast DCT algorithm in 1977. And after that, many better DCT algorithms has occurred successively [2-4].

But the physical performance of DCT is unclear. In addition, the transform matrix of DCT is expressed with floating-point numbers, so the computational complexity is high and more system resources are occupied. To overcome such drawbacks, a novel operation method, 1-D Integer SDCT, is proposed in the paper.

2 Matrix Expression and Operation

In order to describe conveniently, 1-D matrix expression form and some 1-D matrix operation methods need to define firstly.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 499–505.
springerlink.com

2.1 Matrix Expression

The bold letter, $\boldsymbol{A}$ or the abbreviation, $[a_r]$, is adopted to express 1-D matrix, where r represents position of an element in 1-D matrix. When the size of the 1-D matrix is needed to express, $\boldsymbol{A}_{1\times R}$ or $[a_r]_{1\times R}$ is adopted. If the R of two matrices is equal, the two matrices are called the same type matrix.

In addition, a larger matrix can be divided into sub-matrices (or sub-blocks). The matrix with sub-matrix for elements is called partitioned matrix. The partitioned matrix can be expressed as $\boldsymbol{A}_{1\times R,1\times C}$, where, $1\times R$ indicates sub-matrix size, $1\times C$ indicates sub-matrix amount. The sub-matrix can be expressed as $\boldsymbol{A}_{c,1\times R}$, where, $1\times R$ indicates sub-matrix size too and c indicates sub-matrix position in partitioned matrix. A certain element in sub-matrix be expressed as $a_{r,c}$, where, r indicates element position in sub-matrix, and c indicates sub-matrix position in partitioned matrix.

For example, if a 1-D matrix size is 1×4, the matrix can be described as $\boldsymbol{A}_{1\times 4}$. If the matrix is divided into 2 same type sub-matrices, the 1-D matrix will became a partitioned matrix shown as Eq.1.

$$\boldsymbol{A}_{1\times 4} = [a_{1,1} \quad a_{1,2} \,|\, a_{2,1} \quad a_{2,2}] = [\boldsymbol{A}_{1,1\times 2} \quad \boldsymbol{A}_{2,1\times 2}] \tag{1}$$

2.2 Matrix Operation Method

Definition 1. For two matrixes, $\boldsymbol{A}=(a_r)$ and $\boldsymbol{B}=(b_r)$, where $1\le r\le R$, the "point product" operation of the two matrices is marked as $\boldsymbol{D}=\boldsymbol{A}\otimes\boldsymbol{B}$.

$$\boldsymbol{D} = (d_r) = (a_r \times b_r) \tag{2}$$

Definition 2. For two matrices, $\boldsymbol{A}=(a_r)$ and $\boldsymbol{B}=(b_r)$, where $1\le r\le R$, the "matrix element product" operation of the two matrices is marked as $\boldsymbol{A}\nabla\boldsymbol{B}$, and the "matrix element product" operation is defined as Eq.3. The operation result is a 1×1 matrix, that is, the result is a constant.

$$\boldsymbol{A}\nabla\boldsymbol{B} = \sum_{r=1}^{R} a_r b_r \tag{3}$$

Definition 3. For two matrices, $\boldsymbol{A}=(a_r)$ and $\boldsymbol{B}=(b_{r,c})$, where $1\le r\le R$ and $1\le c\le C$, the "sub-matrix product" operation of the two matrices is marked as $\boldsymbol{A}\triangleright\boldsymbol{B}$, and which result is a $1\times C$ matrix. The "sub-matrix product" operation is defined as Eq.4.

$$\begin{aligned}\boldsymbol{A}_{1\times R}\triangleright\boldsymbol{B}_{1\times R,1\times C} &= \boldsymbol{A}_{1\times R}\triangleright[\boldsymbol{B}_{1,1\times R} \quad \boldsymbol{B}_{2,1\times R} \quad \cdots \quad \boldsymbol{B}_{C,1\times R}] \\ &= [\boldsymbol{A}_{1\times R}\nabla\boldsymbol{B}_{1,1\times R} \quad \boldsymbol{A}_{1\times R}\nabla\boldsymbol{B}_{2,1\times R} \quad \cdots \quad \boldsymbol{A}_{1\times R}\nabla\boldsymbol{B}_{C,1\times R}]_{1\times C}\end{aligned} \tag{4}$$

Definition 4. For two matrices, $\boldsymbol{A}=(a_r)$ and $\boldsymbol{B}=(b_{c,r})$, where $1\le r\le R$ and $1\le c\le C$, the "matrix superimposition product" operation of the two matrices is marked as $\boldsymbol{A}\triangleleft\boldsymbol{B}$, and which is defined as Eq.5. The operation result is a $1\times C$ matrix.

$$\boldsymbol{A}_{1\times R}\triangleleft\boldsymbol{B}_{1\times C,1\times R} = [a_1 \quad a_2 \quad \cdots \quad a_R]\triangleleft[\boldsymbol{B}_{1,1\times C} \quad \boldsymbol{B}_{1,1\times C} \quad \cdots \quad \boldsymbol{B}_{R,1\times C}] = \left[\sum_{r=1}^{R} a_r \boldsymbol{B}_{r,1\times C}\right]_{1\times C} \tag{5}$$

3 The Principle of 1-D Integer SDCT

Set 1-D signal sequence be $f(r)$, where r=0,1,…,7, The 1-D DCT are defined as follows:

$$F(u)=\sum_{r=0}^{7} f(r)Q(u,r) \qquad f(r)=\sum_{u=0}^{7} F(u)Q(u,r) \tag{6}$$

Where, $Q(u,r)=\sqrt{\frac{1}{4}}K(u)\cos\left(\pi u\frac{2r-1}{16}\right)$ and $K(u)=\begin{cases}1/\sqrt{2} & u=0\\ 1 & \text{other}\end{cases}$.

The Eq.6 can be expressed in the matrix form as Eq.7.

$$\boldsymbol{Y}=\boldsymbol{P}_0\boldsymbol{X} \quad \boldsymbol{X}=\boldsymbol{P}_0^T\boldsymbol{Y} \tag{7}$$

Where, the $\boldsymbol{P}_0$ is called 1-D DCT transform matrix. According to the symmetry and periodicity of trigonometric function, $\boldsymbol{P}_0$ can be expressed as Eq.8.

$$\boldsymbol{P}_0=\begin{bmatrix} a & a & a & a & a & a & a & a\\ b & d & e & g & -g & -e & -d & -b\\ c & f & -f & -c & -c & -f & f & c\\ d & -g & -b & -e & e & b & g & -d\\ a & -a & -a & a & a & -a & -a & a\\ e & -b & g & d & -d & -g & b & -e\\ f & -c & c & -f & -f & c & -c & f\\ g & -e & d & -b & b & -d & e & -g \end{bmatrix} \tag{8}$$

Where, $a=\frac{1}{2\sqrt{2}}$, $b=\frac{1}{2}\cos\frac{\pi}{16}$, $c=\frac{1}{2}\cos\frac{2\pi}{16}$, $d=\frac{1}{2}\cos\frac{3\pi}{16}$, $e=\frac{1}{2}\cos\frac{5\pi}{16}$, $f=\frac{1}{2}\cos\frac{6\pi}{16}$, $g=\frac{1}{2}\cos\frac{7\pi}{16}$.

A public coefficient can be extracted from each line of $\boldsymbol{P}_0$, these public coefficients form an vector $\boldsymbol{V}$, $\boldsymbol{V}$=[a, k, f, k, a, k, f, k]. The transform matrix can be turned into Eq.9.

$$\boldsymbol{P}_0=\boldsymbol{P}\otimes\boldsymbol{G}_0 \tag{9}$$

Where,

$$\boldsymbol{P}=\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ i_1 & i_2 & i_3 & i_4 & -i_4 & -i_3 & -i_2 & -i_1 \\ i_5 & 1 & -1 & -i_5 & -i_5 & -1 & 1 & i_5 \\ i_2 & -i_4 & -i_1 & -i_3 & i_3 & i_1 & i_4 & -i_2 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ i_3 & -i_1 & i_4 & i_2 & -i_2 & -i_4 & i_1 & -i_3 \\ 1 & -i_5 & i_5 & -1 & -1 & i_5 & -i_5 & 1 \\ i_4 & -i_3 & i_2 & -i_1 & i_1 & -i_2 & i_3 & -i_4 \end{bmatrix} \text{ and } \boldsymbol{G}_0=\begin{bmatrix} a & a & a & a & a & a & a & a \\ k & k & k & k & k & k & k & k \\ f & f & f & f & f & f & f & f \\ k & k & k & k & k & k & k & k \\ a & a & a & a & a & a & a & a \\ k & k & k & k & k & k & k & k \\ f & f & f & f & f & f & f & f \\ k & k & k & k & k & k & k & k \end{bmatrix}$$

The $\boldsymbol{P}$ is called 1-D integer DCT transform matrix. Where, $i_1=c_1/k$, $i_2=c_2/k$, $i_3=c_3/k$, $i_4=c_4/k$ and $i_5=c_5/f$. The $\boldsymbol{G}_0$ is called coefficient matrix of 1-D integer DCT. Where, a, k and f are determined as follows equations according to the normalized properties of transform matrix $\boldsymbol{P}_0$.

$$a=\sqrt{\frac{1}{8}} \quad k=\sqrt{\frac{1}{2(i_1^2+i_2^2+i_3^2+i_4^2)}} \quad f=\sqrt{\frac{1}{4(i_5^2+1)}}$$

From the matrix $\boldsymbol{P}$, we can see that the DCT operation will turn into integer transform if the i_1, i_2, i_3, i_4 and i_5 all are integer. So the values of i_1, i_2, i_3, i_4 and i_5 are the key to integer DCT. When i_1, i_2, i_3 and i_4 are selected already, the transform performance will be the best if i_5 equals to 2 [5]. So i_1, i_2, i_3 and i_4 are called transform basis of integer DCT. Reference [4] gives several integer bases, (10,9,6,2), (5,6,4,1), (6,6,3,2), (6,7,5,1) and (4,5,3,1), which are suitable for video sequence, and their computational complexity are compared. The comparison results are shown as table 1.

Table 1. Operation Contrast Table

No.	Transform basis	Add times	Shift times
1	10, 9, 6, 2	36	10
2	5, 6, 4, 1	32	6
3	6, 6, 3, 2	36	10
4	6, 7, 5, 1	32	10
5	4, 5, 3, 1	28	6

According to the Eq.9, the $Q(u,r)$ can be rewritten to Eq.10. Where, $p(u,r)$ and $g(u,r)$ are the element of matrix $\boldsymbol{P}$ and matrix $\boldsymbol{G}_0$ respectively.

$$Q(u,r)=p(u,r)g(u,r) \tag{10}$$

The Eq.6 can be rewritten as Eq.11 by substituting the Eq.10.

$$Y(u)=\sum_{x=0}^{7} f(x)p(u,x)g(u,x) \quad f(x)=\sum_{u=0}^{7} Y(u)p(u,x)g(u,x) \tag{11}$$

Because all the elements of each row in matrix $\boldsymbol{G}_0$ are same, the Eq.11 can be turned further into Eq.12.

$$Y(u) = g(u)\left[\sum_{x=0}^{7} f(x)p(u,x)\right] \quad f(x) = \sum_{u=0}^{7}\left[Y(u)g(u)\right]p(u,x) \tag{12}$$

Set $\boldsymbol{T}_{1\times 8,1\times 8}=[\boldsymbol{Q}_{1,1\times 8}\ \boldsymbol{Q}_{1,1\times 8}\ \ldots\boldsymbol{Q}_{1,1\times 8}]$, where, $\boldsymbol{Q}_{(u+1),1\times 8}=[p(u,0)\ p(u,1)\ \ldots\ p(u,7)]$. The $\boldsymbol{Q}_{(u+1),1\times R}$ is called integer basis matrix, and matrix $\boldsymbol{T}$ is called integer transform basic matrix. When the transform basis of integer DCT is selected as (10,9,6,2), the integer transform basic matrix is show as Fig.1.(Note: The range of the integer transform basic matrix is adjusted to [0,255] in order to improve visual effect.)

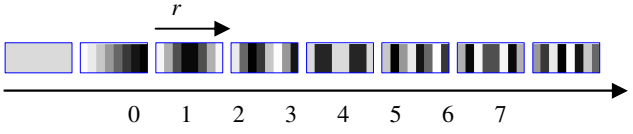


Fig. 1. Transform basic matrix of 1×8 signal

According to the definition of integer transform basic matrix, Eq.12 can be abbreviated as the follows.

$$\boldsymbol{F} = \left(\boldsymbol{f} \triangleright \boldsymbol{T}\right) \otimes \boldsymbol{V} \quad \boldsymbol{f} = \left(\boldsymbol{F} \otimes \boldsymbol{V}\right) \triangleleft \boldsymbol{T} \tag{13}$$

The Eq.13 is SDCT expression form of 1-D DCT. The 1-D integer SDCT has concise mathematical expression, and the transform process of 1-D integer SDCT can be shown as Fig.2.

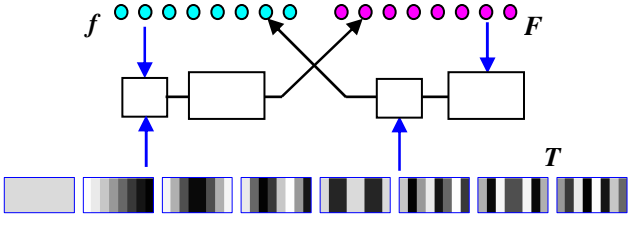


Fig. 2. The sketch map of 1-D integer SDCT operation

From the Fig.2, we can see that the "matrix element product" operation of 1-D signal $\boldsymbol{f}$ and a certain sub-matrix of matrix $\boldsymbol{T}$ can obtain an integer DCT result corresponding to the sub-matrix position. If the 1-D signal $\boldsymbol{f}$ and each sub-matrix of matrix $\boldsymbol{T}$ are performed one time "matrix element product" operation respectively, 1×8 integer DCT results are obtained. Then the "point product" operation of the integer DCT results and matrix $\boldsymbol{V}$ can obtain the final transform result $\boldsymbol{F}$. In addition, the "point product" operation of the $\boldsymbol{F}$ and matrix $\boldsymbol{V}$ can obtain a version matrix. Then the "matrix superimposition product" operation of the version matrix and matrix $\boldsymbol{T}$ can accomplish 1-D IDCT. So the 1-D integer SDCT is easy to comprehend and

convenient to operate. In addition, the computation complexity is simplified, so the computation speed will be faster than the 1-D DCT.

4 The Application in Electronic Image Processing

At present, the DCT is used more and more widely, especially in the field of image coding and video coding. And the DCT algorithm has been made special chipset. If the integer SDCT algorithm can be made special chipset as the DCT algorithm, the performance of the special chipset will be improve well.

In order to validate the performance of integer SDCT, a comparative experiment is performed between DCT and integer SDCT.

Because the 3-D DCT can be accomplished by three times of 1-D DCT successively, the 3-D DCT video coding is experimented by the method proposed in the Reference [6]. The experimental data used in the paper are standard CIF data, "news", shown as Fig.3 (the first frame image). Only the luminance components of the front 16 frames are taken into account in the paper, the chrominance components can be processed similarly.

Fig. 3. The experimental data

The peak signal to noise ratio (PSNR) was used to evaluate the quality of reconstructed video. The PSNR is expressed as follows.

$$PSNR = 10\log\frac{255^2}{\frac{1}{MN}\sum_{r=0}^{M}\sum_{c=0}^{N}[f(r,c)-f'(r,c)]^2}(\text{db}) \tag{14}$$

Where, M is the frame width of video data, N is the frame height, $f(r,c)$ is the pixel value of original frame, and $f_0(r,c)$ is the pixel value of reconstructed frame. The experiment results are shown as Fig.4.

From Fig.4, we can see that the video compression effect of integer SDCT and DCT are very similar, which enough prove that the integer SDCT has similar performance as DCT. But the 1-D integer DCT is convenient to operate.

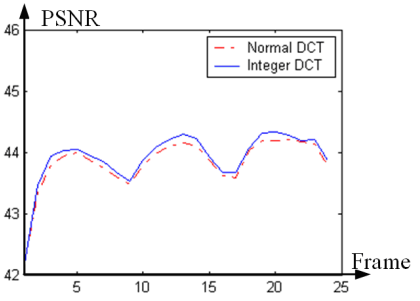


Fig. 4. The experimental results

5 Conclusion

The 1-D integer SDCT proposed in the paper not only has concise mathematical expression, easy to comprehend, convenient to operate, but also simplifies the calculation of 1-D DCT. And the integer SDCT has similar performance as DCT. So the 1-D integer SDCT has much potential to be applied in the field of electronic image processing.

References

1. Chen, W.A., Harrison, C., Fralick, S.C.: A Fast Computational Algorithm for the Discrete Cosine Transform. IEEE Transactions on Communications 25(9), 1004–1011 (1977)
2. Lee, B.G.: A New Algorithm to Compute the Discrete Cosine Transform. IEEE Transactions on Acoustics, Speech, and Signal Processing ASSP-32(6), 1243–1245 (1984)
3. Suehiro, N.: Fast algorithms for the DFT and other sinusoidal transform. IEEE Transactions on Acoustics Speech and Signal Processing 34(3), 642–644 (1986)
4. Giurcaneanu, C.D., Tabus, I.: Low-complexity transform coding with integer-to-integer transforms. In: ICASSP, Salt Lake City, Utah, USA, vol. IV, pp. 2601–2604. Utah State University (May 2001)
5. Cham, W.K.: Development of integer cosine transform by the principle of Dyadic Symmetry. IEE Proceedings 136, 276–288 (1989)
6. Tang, H., Sun, W., Gao, B., Zhang, F.: Research on Quantization and Scanning Order for 3-D DCT Video Coding. In: ICCSEE, Hangzhou, China, vol. I, pp. 200–204 (March 2012)

Apply Neutron Radiation and X-Ray Images Fusion for Defect Detection in Manufacturing

Juan Wang[1], Leting Tan[1], and Siyu Lai[2]

[1] Department of Computer Science, China West Normal University, NanChong, China
[2] Department of Medical Image, North Sichuan Medical College, NanChong, China
{wjuan0712,LeT_Tan,lsy_791211}@126.com

Abstract. Electronic products defect detection is an important branch of electronic image processing, in which, image fusion can be applied to deal with the task. In view of contourlet is more suitable to describe anisotropic lines, curves and boundary characteristics of image than wavelet, and can effectively capture the geometry structure. Comparing with contourlet transform, non-subsampled contourlet transform (NSCT) not only enjoys features of multi-scale, multi-direction, but also possesses translational invariance. In this paper, a new digital neutron and X-ray images fusion method based on contourlet and IHS transformation is proposed. The results show that our method preserves more image detail and hence provides more accurate information, which is superior to IHS and wavelet-IHS based fusion approaches. It is proved to be an effective approach to industrial production.

Keywords: image fusion, wavelet transform, contourlet transformation, electronic image processing.

1 Introduction

Multi-sources sensing data are of redundant, complementary and collaborative [1]. Image fusion is an effective approach that integrates data from varied sources and reduces uncertainty, which is beneficial to scientific observation and social production. It has been adopted in many fields including remote sensing, robotics, medical image processing, electronic technology, and military purposes. Digital neutron radiography [2] is the imaging technology that similar to X-ray radiography [3], however, the mechanism by which the former interact with material is quite different from that of the latter; therefore, they have their own advantages in dealing with varying materials, respectively. Then the fusion of digital neutron radiography and X-ray radiography for the same object will provide more internal and detailed information, so we can analyze components and structures for specific objects better. And so many skills have been investigated to improve the quality of fusion image. As a result, it has been widely investigated to improve the accuracy of image visual interpretation and automatic identification, and Intensity Hue Saturation (HIS), Principal Component Analysis (PCA), Brovey, and wavelet transform etc. are proposed.

As a traditional multi-scale tool for image analysis, wavelet transform is with good locality in spatial and frequency domains and is extensively applied in image fusion

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 507–512.
springerlink.com

fields [4-5]. Wavelet is regarded as the best basis when processing objective function contains point-like singularity, but it is not the case in representing lines or curves that constitute image structure, and is prone to bring about image details blurring when perform image fusion. Ridgelet, Curvelet, Contourlet etc. multi-scale geometric analysis theories have recently been proposed to mend the disadvantage of wavelet in expressing high-dimensional singularity [6] and are adopted to solve in image processing. In which, contourlet transform [7-8] is flexible in setting number of direction and choosing the type of filter, is also better direction sensitive and with good anisotropy. Hence, it is capable of accurately capturing edge information of image and filling it into sub-bands with different scales and frequency. But the subsampled action in transform progress make contourlet does not possess shift invariance, which leads to the introduction of pseudo-Gibbs phenomenon when processing images. Non-Subsampled Contourlet Transform (NSCT) [8] do no have sampling step, but is composed of non-sampled pyramid decomposition and multi-level directional filter instead. It is not only inherits multi-scale, multi-dimensional features but also has property of shift invariance.

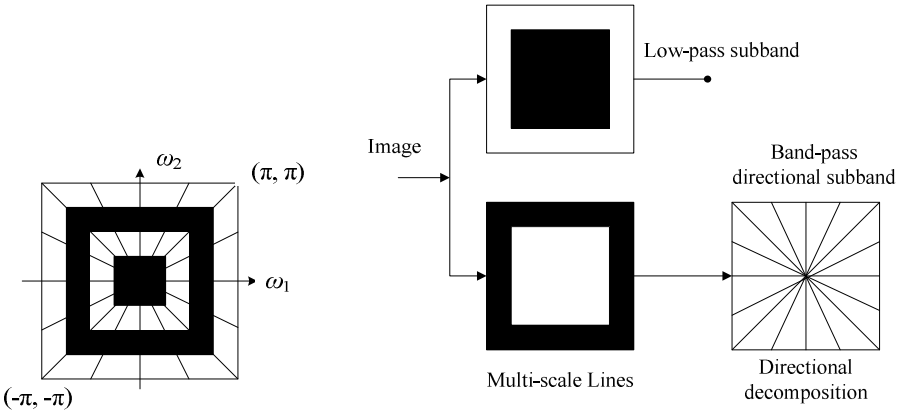
IHS transformation is defined as transform the image from RGB space into relative independent, can be controlled separately, IHS space that is capable of quantificationally describe the characteristics of color precisely. IHS transform fusion method greatly improves the spatial resolution of source images, and that is why it has been widely applied. In this paper, subsampled and non-subsampled contourlet transform are combined with IHS transform to fuse panchromatic and multi-spectral image. Because contourlet shows excellent performance in expressing geometric feature, especially the edge information, so it extracts characteristics of the source image and provides more abundant and accurate information. Experimental results show that the contourlet - IHS fusion method is superior to IHS, àtrous - IHS fusion method [5].

2 Contourlet and Non-subsampled Contourlet Transfromation

Contourlet transformation [7] is two-dimensional image representation proposed by Minh N Do and Martin Veteterli in 2002, which is composed of the Laplician Pyramid (LP) decomposition and Directional Filter Bank (DFB). contourlet method first performs multi-scale decomposition on source image to capture singular point by utilizing LP transformation, and each LP operation will produce a low-frequency sub-band image whose resolution is half of the original one and a high-frequency sub-band image. By using iterative LP method on sub-band image, the multi-scale decomposition is realized, then divides high-frequency sub-image employing DFB. And at last, integrate the singular points in the same direction into profile lines according to the directional information, namely, synthesize a coefficient and thus capture image contour. contourlet processes high-frequency part using directional filter and divide high-frequency information into combination of several directional components, which make contourlet have great performance of non-linear approximation.

Contourlet transformation does not have shift invariance property in contourlet construction process due to the existence of subsampling operation in LP and DFB,

which will inevitably cause Gibbs phenomenon. For the purpose of eliminating the flaw, A. L. Cunha etc. proposed NSCT transformation in 2005 [9], which is structured by Non-Subsampled Pyramid (NSP) and Non-Subsampled DFB decomposition (NSDBF). NSCT contains a series of filters, Fig.1 (a) shows the NSCT frame, from which the two-dimensional frequency domain is divided into the wedge-shaped directional sub-bands as demonstrated in Fig.1 (b). NSCT removes steps of signal extraction after filtering analysis and interpolation before synthesized filtering using the equivalent shift attribute of *Z*-transform, and relies instead on up-sampling the relative filters first and then conducts the filtering and synthesized analysis. NSCT not only inherits the multi-scale and multi-directional property of contourlet but possesses the translation invariant feature at the same time, which can get rid of the Gibbs phenomenon effectively.



(a) Ideal NSCT frequency division (b) Sturcture of NSCT transformation

Fig. 1. Experimental data and different fusion images

3 Contourlet Based IHS Transform Fusion Method

The neutron (NEU) image and X-ray (X) one are both required to be accurately registered before the fusion is carried out. The SCT and NSCT are introduced to HIS fusion method respectively in our paper, and the same fusion rule is implemented to realize the neutron and X-ray images. The detailed Contourlet-IHS based neutron an X-ray images fusion entries are listed below:

(1) Perform IHS transformation on neutron image and get its intensity component I, hue component H and saturation component S;

(2) Run histogram matching program to neutron image with component I and conduct contourlet decomposition to neutron image and component I respectively, and then achieve separate low-frequency approximated and high-frequency detailed coefficients;

(3) Low-frequency and high frequency coefficients are processed accordingly in the light of different fusion rules and generate fused coefficients; The fusion rule

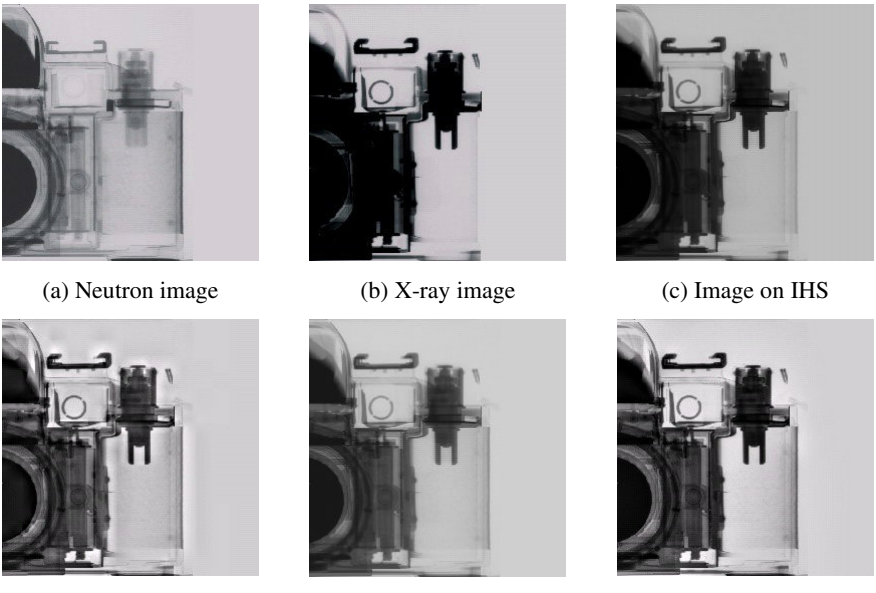
meets the following two points , the low-frequency approximated coefficients come from X-ray image and the high-frequency detailed coefficients of I and neutron in different scales and directions, with larger absolute magnitude, is selected;

(4) Conduct inverse contourlet transformation to coefficients resulting from step (3), that is, image reconstruction;

(5) The reconstructed image from step (4) are chosen as new component I, with which the inverse IHS transformation are performed plus components H and S derive from step (1), and finally the ultimate fusion image is obtained.

4 Experiments and Results Analysis

Two images at the same object are selected as the source samples, Fig.2 (a) is digital neutron radiography image of a camera and (b) is X-ray radiography image of the same camera, both are the size of 256 by 256 and have been registered. Fig.2 (c), (d), (e), (f) are fused images based on IHS, wavelet-IHS, subsampled contourlet-IHS and NSCT-HIS fusion rules.



(a) Neutron image (b) X-ray image (c) Image on IHS

(d) Image on wavelet-IHS (e) Image on SCT-IHS (f) Image on NSCT-IHS

Fig. 2. Experimental data and different fusion images

From the visual point of view, four fusion methods are significantly improved in aspect of spatial-resolution, IHS fusion image (c) is entirely dark comparing with the original neutron (a) image, the wavelet based IHS fusion image (d) is lighter than (a), and images on our method (e) and (f) are close to (a) in intensity, in particular, the NSCT-IHS based image (f) is more consistent with the original neutron grayscale.

Consequently, our method has better details preserving function than any other ways at the time when promoting spatial resolution.

To verify the validity of this method further, this paper evaluates it objectively with five statistics of Information Entropy (IE), Average Gradient (AG), Correlation Coefficient (CC), Universal Image Quality Index (UIQI) and Mutual Information (MI). IE reflects the amount of information carried by image, bigger IE means better fusion effect and stronger detail expression performance. AG is used to indicate the definition of the image, the greater the IE is, and the clearer is the image. CC depicts how correlative the two images are, larger CC of two images represents more fine fusion. The criterion MI of two images measures the mutual dependence of two images, the bigger the MI is, it means, the more information inherited by fused image from source one and better effect is achieved. Index UIQI is calculated as follows:

$$Q_0 = \frac{4\sigma_{xy}\overline{x}\overline{y}}{(\overline{x}^2 + \overline{y}^2)(\sigma_x^2 + \sigma_y^2)} \tag{1}$$

where x and y represent source image and fused image; $\overline{x}$ and $\overline{y}$, σ_x and σ_y denote for mean, standard deviation of source and fused images respectively; σ_{xy} are covariance of source and fused images. Above equation can be divided into three parts by which it can be multiplied:

$$Q_0 = \frac{\sigma_{xy}}{\sigma_x\sigma_y} \times \frac{2\overline{x}\overline{y}}{(\overline{x}^2 + \overline{y}^2)} \times \frac{2\sigma_x\sigma_y}{(\sigma_x^2 + \sigma_y^2)} \tag{2}$$

Equation (2) measures the difference of two images from points of relevant loss, intensity and contrast distortion, if Q_0 close to 1 means the fusion performance approach to the best. Given that A and B are original images, F is fused image, the corresponding gray level are [0, a], [0, b] and [0, f] respectively, the mutual information of F to A, F to B can be relatively written as:

$$I_{FA} = \sum_{i=0}^{a}\sum_{j=0}^{f} P_{FA} \log_2 \frac{P_{FA}}{P_F P_A} \tag{3}$$

$$I_{FB} = \sum_{i=0}^{b}\sum_{j=0}^{f} P_{FB} \log_2 \frac{P_{FB}}{P_F P_B} \tag{4}$$

where P_A, P_B and P_F are the probability densities of images A, B and F; P_{FA}, P_{FB} represent the joint probability density of two images.

Table 1. The comparison of three sets of data

Fusion Method	IE	AD	UIQI	CC		MI	
				NEU	X	NEU	X
IHS	6.892	34.523	0.802	0.818	0.874	4.621	5.726
Wavelet-IHS	6.947	39.496	0.867	0.837	0.803	4.792	4.165
Contourlet-IHS	7.159	38.754	0.936	0.925	0.825	5.235	4.477
NSCT-IHS	7.132	36.413	0.943	0.934	0.842	5.421	4.557

Table 1 shows statistics of indexes above. Seen from Table 1, we find that the fused image on the basis of subsampled or non-subsampled Contourlet IHS transformation is superior to HIS and wavelet-IHS methods in criteria of IE, UIQI, CC and MI except index AG is slightly inferior to wavelet-HIS approach. These statistics indicates that our method merges more spatial information and is more complementary. And except the NSCT-IHS based fusion method owns a little bit low values in IE and AG than the subsampled Contourlet-HIS, the rest of data are all larger. Take a summarized considering into account, the conclusion that the NSCT based fusion method is the best is drawn.

5 Conclusions

This paper presents a new digital neutron radiation image and X-ray image fusion method based on Contourlet-IHS transformation. As expected, the fusion image on the ground of Contourlet-IHS outperforms IHS and wavelet-IHS transformation both in spatial resolution enhancement and details preserving. According to IE and AG, the subsampled contourlet method is better than NSCT fusion while the NSCT based fused image is more correlative with the original neutron image, synthesizing MI and UIQI, we deem that the NSCT-IHS fusion method is optimal. So, the proposed method is promising to be used in industry, especially the defect detection related field in electronic technology. Combine the advantages of wavelet and contourlet will be the next step to our research work.

Acknowledgments. This work is supported by Sichuan Provincial Department of Education (11ZB034), China. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

References

1. Gustavo, C.V., Lorenzo, B.: Kernel Methods for Remote Sensing Data Analysis. Wiley publishing house (2009)
2. Halmshaw, R.: Neutron Radiography Techniques and Applications. New York Press (1999)
3. Mo, D.W.: Neutron Radiography. Atomic Energy Press, Beijing (1996)
4. Data Fusion for Multi-sources Remotely Sensed Images based on IHS Transformation and Wavelet Analysis. Environmental Science and Information Application Technology, 1 (2009)
5. Liu, J.J., Guan, L., Li, L.L.: Image Fusion of Landsat 7 ETM+ Data based àtrous Wavelet Transform. Remote Sensing for Land & Resources 23, 11–15 (2007)
6. Duan, G.D., Li, J.P., Huang, T.X.: Study on Image Multi-scale Geometric Analysis. Application Research of Computers 43, 46–50 (2007)
7. Do, M.N., Vetterli, M.: Contourlets. Academic Press (2002)
8. Do, M.N., Vetterli, M.: The Contourlet Transform: An Efficient Directional Multi-resolution Image Representation. IEEE Trans. on Image Processing 14, 2096–2106 (2005)
9. Cunha, A.L., Zhou, J., Do, M.N.: The nonsbsampled contourlet transform: Theory, Design, and Applications. IEEE Trans. on Image Processing 15, 3089–3101 (2006)

Design of an Electronic Image Processing Waveguide with Defects Based on Triangular Lattice Photonic Crystal

Xiao-ling Chen and Wei Li*

Department of Science, Nangchang Institute of Technology, Nanchang, 330099, China
liweidting@163.com

Abstract. In this paper, we proposed an electronic image processing waveguide with a special structure at the intersection based on triangular lattice photonic crystal. The effect of the radius of the innermost air holes and the electronic image processing waveguide are investigated. With the change of the radius of the electronic image waveguide, we can get when r_{wg} equals to 0.25a, the structure presented can be applied as a time recorder in précised calculation.

Keywords: photonic crystal, waveguide, transmission, electronic image processing.

1 Introduction

Although a large number of work has done to demonstrate the fabrication feasibility of PhC (Photonic Crystal) [1] in different dimensionalities and with different materials, owing to the ability of the PhCs to manipulate photons in the wavelength dimension, various PhC devices such as mircoresonators, waveguides, channel-drop filters, and optical switches have been developed as building blocks for integrated photonic systems [2-3]. In analogy to impure states of semiconductors, defect modes also can be realized in perfect periodic structures. Moreover, PhC waveguides generated by line defects have been studied intensely due to their unique ability to confine and control the propagation of electromagnetic waves in optical circuits. They are different from conventional dielectric electronic image waveguides. PhC waveguides strongly depend on photonic band gap configuration while conventional dielectric waveguides depend on total internal refraction [4-5].

By virtue of advanced numerical methods, different optical devices based on PhC coupled waveguides have been designed. For instance, a multiplexer-demultiplexer based on PhC waveguide coupler has been proposed and the wavelength demultiplexing properties have been theoretically investigated. And a compace polarizing beam splitter and a novel mode converter for a wide frequency range have been realized in PhC coupled waveguides[6-7]. These various structures play an important role in modern optical devices. At present, the use of triangular lattice photonic crystal is the most popular because the wider range of bandgap than in the square lattice.

* Corresponding author.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 513–519.
springerlink.com

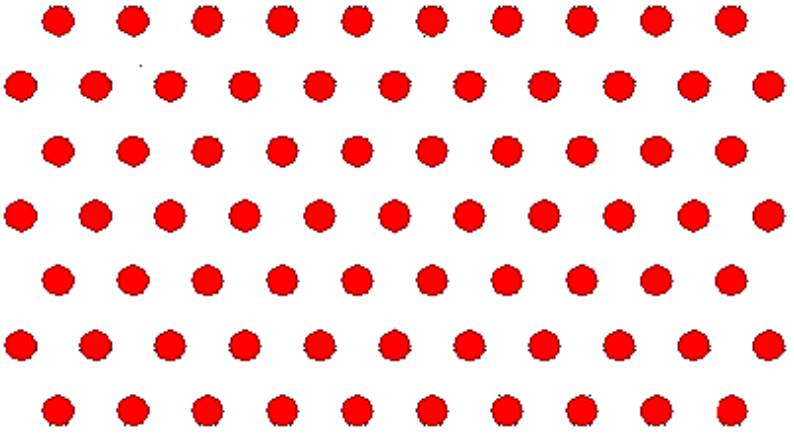


Fig. 1. The structure of the triangular lattice photonic crystal

As shown in Fig.1, the structural parameters are as follows: the lattice constant is a, the radius of the medium rods $r = 0.18a$, the slab is made of AlGaAs with a dielectric constant $\varepsilon = 11.56$ and a nonlinear-index coefficient $n_2 = 1.5 \times 10^{-17} m^2 / W$ at a wavelength of $1.55\,um$, and the air holes are air with $\varepsilon = 1$. AS this kind of waveguide only exits in the TM modes band gaps, we choose lightwave TM as the incident source. The TM band gap of the structure is in the range of $0.303 \le a/\lambda \le 0.4883$, where λ denotes the optical wavelength in free space.

2 Description of the Structure

We use a PhC cavity at the intersection for wavelength selectivity. Here a single cell cavity is modified to have high factor (Q), as shown in Fig 2. Six nearest neighbor holes are reduced to r_n=0.18a and pushed away from the cavity centers. Here a is the lattice constant. The distance (c_n) from the cavity center to the nearest holes center is 1.1a. The radius of the air holes (r_{pc}) constituting the PhC is 0.25a. The refractive index used for PhC slab in our simulation is 3.4.

It was demonstrated that symmetry mismatching between a waveguide mode and a cavity mode could prevent photons propagating from the cavity to the waveguide. In addition, spectral filters, which can eliminate transmission to the unwanted waveguides, were realized by using the mode-gap effect resulting from modulation of the width, the lattice constant, the air hole radius of the waveguides[8]. To improve the efficiency of transport, the modes in cavity are divided into two janes which odd and symmetry. Thus, the interferences of two modes in waveguide are damaged, and the transmission light will not return to the port. Figure 3 gave the two mode charts of the cavity.

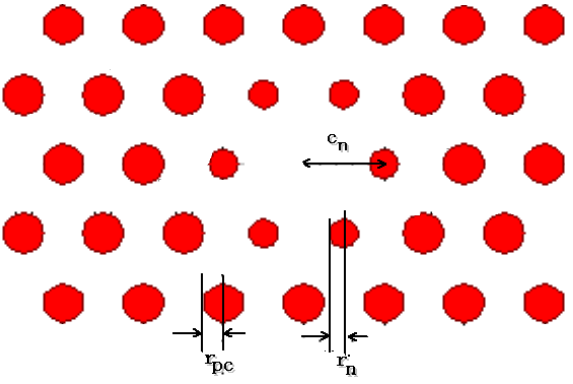


Fig. 2. Modified single cell cavity structure. r_{pc}, r_n, and c_n are 0.25*a*, 0.18*a*, and 1.1*a*, respectively.

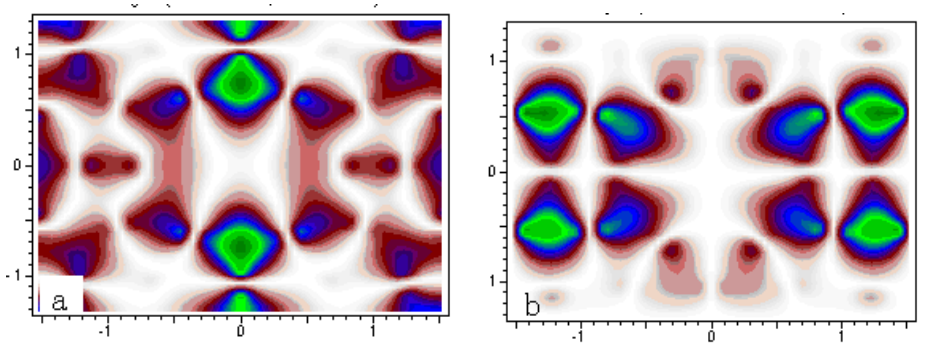


Fig. 3. Two mode charts of the cavity (a) symmetry mode (b) odd mode

In this paper, we propose a new design of an intersection for crossing waveguides without cross-talk in triangular lattice PhC by combing both: the effects of the symmetry mismatch and the mode gap, which can be generally applied to any waveguide-cavity coupled system, shows in figure 4 below. We also illustrate the transmission behavior in the intersections by two-dimensional (2D) finite difference time-domain (FDTD) simulations.

To eliminate the cross-talk among the waveguides, we can propose a special C_{6v} structure at the cross [5]. In this paper, we use a single line defect waveguide formed by filling in one row of air holes and enlarging the innermost holes (r_{wg}), as shown in Fig. 4. Here the enlarged innermost holes shift the frequency of the dispersion curve of the waveguide mode. Since the mode has a relatively broad single mode region, we focus on the fundamental waveguide mode. The magnetic field pattern of the fundamental mode has even mirror symmetry to the axis along the waveguide, as shown in Fig. 3.

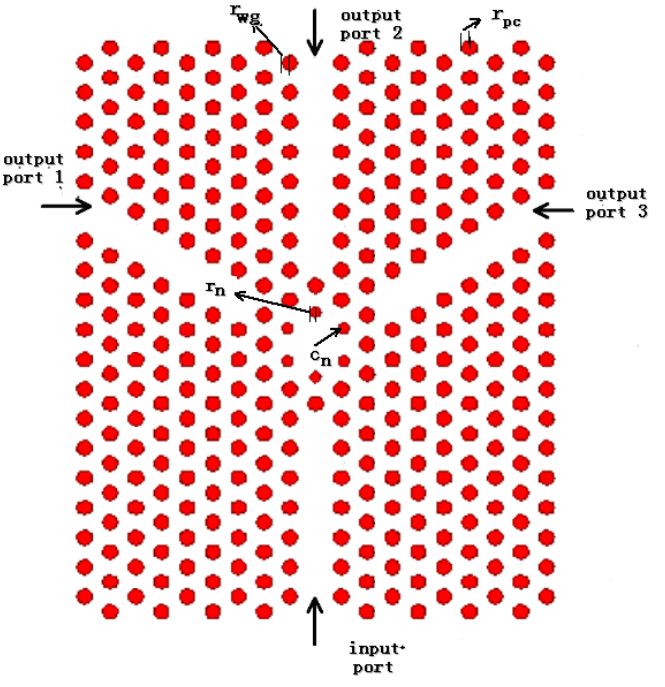


Fig. 4. The structure of the waveguides. It includes one input port and three output ports, and here r_{wg} is the radius of the air holes along the waveguides.

In this paper, we use a single line defect waveguide formed by filling in one row of air holes and enlarging the innermost holes (r_{wg}). Here the enlarged innermost holes shift the frequency of the dispersion curve of the waveguide mode. Since the mode has a relatively broad single mode region, we focus on the fundamental waveguide mode. The magnetic field pattern of the fundamental mode has even mirror symmetry to the axis along the waveguide

3 Simulation

To the structure showed in Fig. 3, the background is AlGaAs with a dielectric constant $\varepsilon = 11.56$, and the dielectric cylinders are air with $\varepsilon = 1$. AS this kind of waveguide only exits in the TM modes band gaps, we choose lightwave TM as the incident source. And the t wavelength of the incident lightwave is the communication wavelength 1.31. And the simulate method here is 2-D FDTD with the pulsed excites.

To the structure showed in Fig. 4, we can change the radius of r_{wg} , we can find out that the transmission at the two output ports are different and variable. As shown in Fig. 4, the maximum output power ratio are achieved when r_{wg} at the different value. When in each condition of r_{wg}, the display of the transmission seems to have no different, but we can find out r_{wg} ~ 0.25a is the best choice.

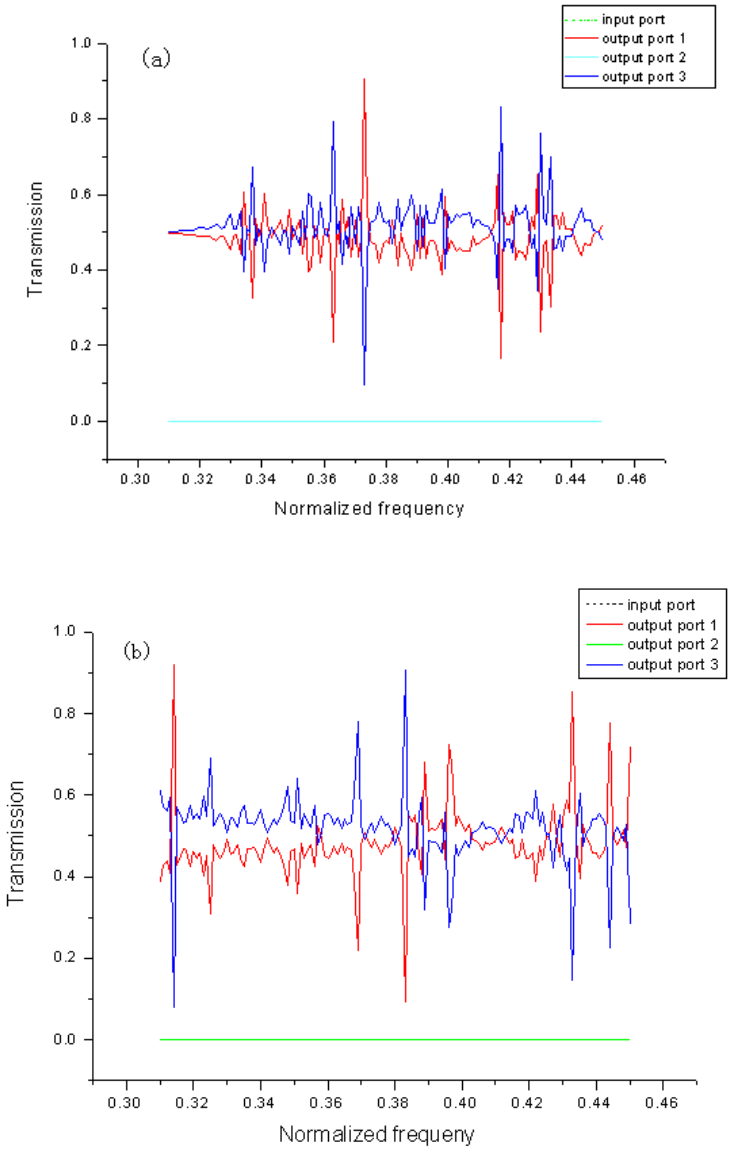


Fig. 5. The transmission diagrams of different r_{wg}. (a) r_{wg}~0.18a, (b) r_{wg} ~ 0.25a, (c) r_{wg}~0.35a

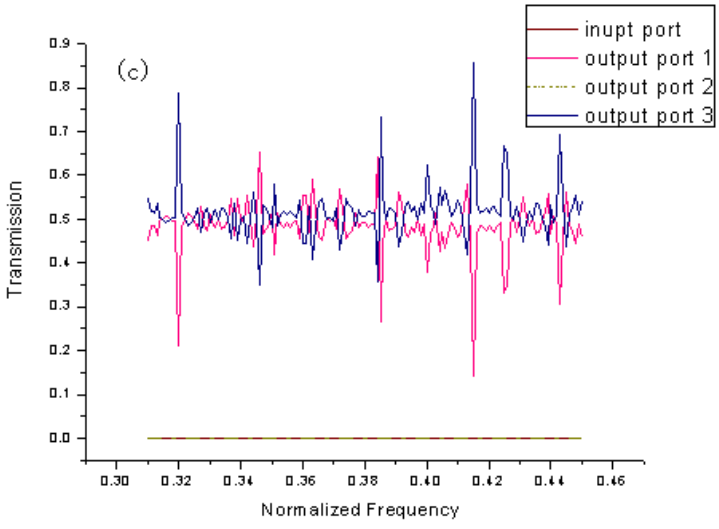


Fig. 5. (*continued*)

Given the bandgap that between 0.303 and 0.4883 provided at the beginning, in Fig. 5(a) and 5(c), the transmission at both output ports varies up and down at 0.5, and neither have a relative stability. However, in Fig. 5(b), the transmission at output port 1 increases to a high value at the normalized frequency 0.314, as high as 0.919. Since that, the transmission at port 1 less than that at port 2 until the frequency is 0.383, the frequency deviation of the two points is relatively large in precision communication system. And this structure presented can be also applied as a time recorder in calculation.

4 Conclusion

In this paper, we present the principle of the operation of a electronic image processing waveguides-crossing with negligible cross-talk in a triangular lattice PhC by 2D FDTD simulation. The signal from the input port mostly passes to the two bend branches with 60^0 output ports through the resonant tunneling of one of two modes in a single cell cavity. The cross-talk to the longitudinal ports can be strongly suppressed by the symmetry mismatch between the cavity mode and the waveguide mode by changing the radius of the electronic image waveguides and the mode-gap resulting from the radius modulation of the innermost air holes of the waveguide, as a result the reflectivity and output port that longitudinal the input port are both near to zero. The proposed device can be used to construct complex photonic crystal systems. The demonstrated symmetries and the mode-gap ideas can be also applied to analyze the coupling behaviors between the cavity and the electronic image processing processing waveguide modes in any PhC based systems, and also can be applied as a time recorder.

Acknowledgment. Project funds: Natural Science Foundation of Jiangxi Province (2008GZW0007), Natural Science Foundation of Nanchang Institute of Technology (2010KJ020).

References

1. Yablonovitch, E.: Inhibited spontaneous emission in solid-state physics and electronics. Phys. Rev. Lett. 58(20), 2059–2062 (1987)
2. Painter, O., Lee, R.K., Scherer, A., Yariv, A., O'Brien, J.D., Dapkus, P.D., Kim, I.: Two-dimensional photonic band-gap defect mode laser. Science 284, 1819 (1999)
3. Kwon, S.-H., Kamp, M., Forchel, A., et al.: Elimination of cross-talk in waveguide intersections of triangular lattice photonic crystal. Optics Express 16(15), 11399 (2008)
4. Park, H.G., Kim, S.H., Kwon, S.H., Ju, Y.G., Yang, J.K., Baek, J.H., Kim, S.B., Lee, Y.H.: Electrically driven Single-Cell Photonic Crystal Laser. Science 305, 1444–1447 (2004)
5. Song, B.S., Noda, S., Asano, T., Akahane, Y.: Ultra-high-Q photonic double hetero - structure nanocavity. Nat. Mater. 4, 207–210 (2005)
6. Tokushima, M., Yamada, H.: Light propagation in a photonic-crystal-slab line-defect waveguide. IEEE J. Quantum Electron 38, 753–759 (2002)
7. John, S.: Strong localization of photons in certain disordered dielectric superlattices. Phys. Rev. Lett. 58(23), 2486–2489 (1987)
8. Noda, S., Chu Tinan, A., Imada, M.: Trapping and emission of photons by a single defect in a photonic bandgap structure. Nature 407(5), 608–610 (2000)

Automated Land Resource Classification of Electronic Photograph Based on Satellite CMOS Detector

Yifu Huang, Haoyan Chen, Haibin Cai, Chao Peng, and Linhua Jiang*

Shanghai Key Laboratory of Trustworthy Computing,
East China Normal University, Shanghai, China
lhjiang@sei.ecnu.edu.cn

Abstract. The research of automated classifying and recognizing electronic photograph based on satellite detectors with various machine learning methods has been in the limelight of current research. In our experiment, we applied K-means clustering algorithm to segment urban electronic photograph of CMOS detector, and interpreted segmentation results as different object classes such as greenbelt and sandy desertification, serving for machine intelligence application of remote electronic sensor equipments. In the paper, first we discussed the difference between traditional segmentation methods and K-means segmentation methods, and then compared gray image segmentation with color based image segmentation; furthermore we did proper semantic analysis research on the object classes of segmentation results and made relevant statistics. Experiment results show that, K-means clustering can get preferable results when electronic photograph has high discrimination in color scales L*a*b* values. Statistical analysis of segmentation results has great scientific significance and practical value for both city planning and automated resource monitoring.

Keywords: electronic photograph, K-means clustering, semantic understanding, land resource classification, CMOS detector.

1 Introduction

As one of the typical electronic engineering applications, the processing of remote electronic photograph shows great significant in the research area of satellite remote sensing, which allows the application of machine intelligence on electronic sensor equipments. The intelligent application of satellite detectors is of great benefits to the society for dozens of years. The main goals of remote sensing are identifying materials on the land cover and analyzing the spectral signatures acquired by satellite or airborne electronic sensors [1]. Mostly digital CCD and CMOS cameras are employed for industrial imagery. The system uses a CMOS image sensor e.g. STAR250 to get image in satellite becomes more and more widely used.

Due to the great significance of remote electronic photograph processing and application, the classification and recognition of landscape matters as the key processes,

* Corresponding author.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 521–527.
springerlink.com

are important as well. How to achieve precise classification and understanding the content on remote electronic photograph is thus one of the hot topics in the state-of-art research.

In the early stage, the research of image classification and recognition focused on processing image edge and region. Paik et al [2] used a neural network to detect edge, Zhou [3] proposed a circular smoothing edge detector, and Chehdi et al [4] proposed a segmentation method in order to localize and identify principal homogeneous regions. However at present, researchers focus on adopting unsupervised, supervised, semi-supervised machine learning methods to classify image. (1)Unsupervised method refers to the problem of trying to find hidden structure in unlabeled data. Nsaibi et al [5] proposed a new unsupervised land cover classification method based on probabilistic fusion theory. (2) Supervised method is the task of inferring a function from labeled training data. Hwang et al [6] proposed based on Gabor texture features and SVM to classify satellite imagery. (3) Semi-supervised method makes use of both labeled and unlabeled data for training that typically a small amount of labeled data with a large amount of unlabeled data. Tuia et al [7] successfully implemented this method with cluster kernels.

The traditional segmentation methods of remote electronic photograph are mainly applied based on edge and region. Although it is said that human can understand image only by edge, image graying may lost a lot of information. So in this paper, by contrast, we adopted K-means clustering algorithm to directly cluster original color image in L*a*b* color space. The term "K-means" was first introduced by James MacQueen [8] in 1967. In statistics and data mining, K-means clustering is a method of cluster analysis which aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean [9].

Our research on the processing of remote electronic photograph of CMOS detectors aims at getting the statistical analysis of segmentation results, such as the coverage of greenbelt and sandy desertification. It can provide detection means for some organizations such as land resource statistic bureau, and contribute to the proper monitor and planning of urban resource. In experiment, we discussed the difference between edge and region based traditional segmentation methods and K-means segmentation methods, and compared gray image segmentation with color based image segmentation; furthermore we did proper semantic understanding research on the classes of segmentation results and made relevant statistics. Thus we discovered the preferable condition, that applying K-means to the processing of remote electronic photograph.

2 Methods

2.1 Satellite Image Resource

Remote electronic photograph can be easily got by electronic robots and satellites, and many organizations can provide enormous remote electronic photograph resources, free ones such as Google Map [10] and Baidu Map [11], as well as a lot of commercial satellite image products [12]. We pay more attention to urban image with distinct feature, such as greenbelt, sandy desertification, architectural complex and river. Statistics about these image segmentations are more beneficial for town monitor and

resource management. In experiment, we chose some satellite image with complex topography (such as (Figure 1)) from Google Map as one of our experiment exemplars.

2.2 Traditional Satellite Image Segmentation

As already noted, traditional segmentation methods focus on image features such as edge and region in gray image. Procedures are as follows:

(1) Apply initial processing to image, including image graying, image enhancement and image smoothing.
(2) Use edge detection operators to extract image edge contour. Image edge is reflected by gray scale discontinuity.
(3) Adopt morphology methods to thin edge, and generate region, such as open operation and closed operation.

In experiment, we successfully adopted above-mentioned traditional edge detection and segmentation algorithms to process satellite imagery. It is used to be contrasted with latter detailed research on K-means clustering algorithm.

2.3 K-Means Satellite Image Segmentation

As color-based segmentation, K-means clustering algorithm clusters pixels which are close in one of the marked colors by calculating the Euclidean distance between each of those pixels and each marked color. The smallest distance tells that the pixel most closely matches that marked color. The detailed steps are as follows:

(1) Convert image from RGB color space to L*a*b* color space. Converting RGB values to L*a*b* values needs two steps. First it converts RGB to CIE XYZ values. Then it transforms the CIE XYZ values to L*a*b* values. The formulas in the process are available in [13].
(2) Classify the colors in 'a*b*' space using K-means clustering. Variations in the luminosity layer 'L*' ignored here, we only take the variations of chromaticity-layer 'a*' and chromaticity-layer 'b*' into consideration. As a result, the pixels in the image are just the points of a two-dimensional space with 'a*' and 'b*' components. First, we select k points randomly as the initial centroids. Then we use Euclidean distance metric to partition all the pixels in the image into k clustering buckets.
(3) Label every pixel in the image using the results from K-means, which aims at using index to separate the corresponding cluster from the image respectively, so as to pave the way for the statistical work and further refinement of the results.
(4) Create images that segment the satellite image by color. Apply the indexes marked in the last step to extract each of the clustered images from the original image respectively, which adds convenience to the observation and understanding of the extracted images.
(5) Refine the former results if needed. If the results are far from satisfactory, a further refinement, which the L*-component plays an important role in extracting the pixels with high L-value from the clustered images, will be a must.

K-means only clusters the pixels in the image, but all the classes resulted from the clustering must be interpreted by the user as to what the color patterns may mean in terms of classes that are actually present in the real world scene [14]. So we had all classes achieved from clustering understood semantically by a predefined mapping function which map clusters' centroids from their color space to the corresponding classes, such as greenbelt, sandy desertification, buildings, etc.

2.4 Satellite Image Semantic Understanding

We got several clusters from the original image through the K-means clustering algorithm. Each cluster contains a lot of cluster members and a sole centroid. We defined a mapping function, mapping centroids from their 'a*b*' color space to the corresponding classes. For example, in our experiment, the lower the centroid's a*-component is, the more possibly the centroid will be identified as the greenbelt class while the lower the centroid's b*- component is, the more possibly the centroid will be identified as blue-roof-building class. On the opposite side, the higher a*-component for a centroid may indicate that the centroid represents building class while the higher b*-component for a centroid may indicate that the centroid represents sandy desertification class. Because the semantic understanding of image is such a subjective work, so it should be done according to concrete conditions and requirements.

After discussion on the logical semantic understanding of each cluster in the image, we achieved the coverage of each cluster corresponding to its class in the region of the image, such as the coverage of greenbelt and the coverage of sandy desertification and so on, by calculating the respective ratio of the total number of pixels in the cluster to the total number of pixels in the whole image.

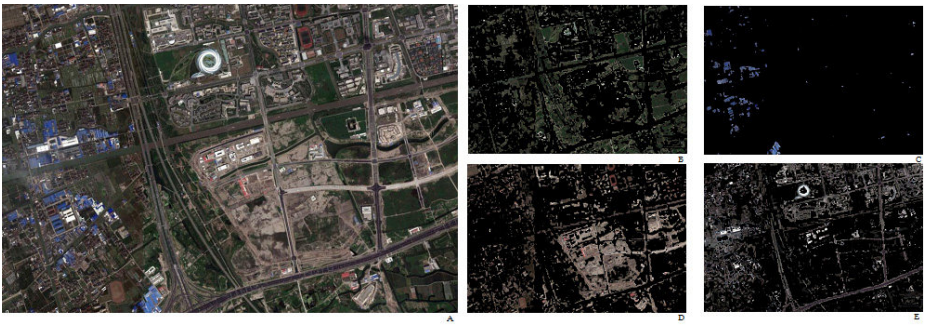


Fig. 1. A. Original image, from Google Map, locates at Zhangjiang High-tech Development Zone, Shanghai, China. Through the semantic understanding: B. Cluster 1, Greenbelt. C. Cluster 2, Roof. D. Cluster 3, Sandy Desertification. E. Cluster 4, Buildings.

The data achieved can be meaningful. The department concerned can apply these data to the management and statistics of the land resources. For example, one can tell whether the area of each class has grown or shrunk by observing the changing tendency in the data which show the coverage of each class in the same region at different time. Then more attention will be paid to those changing regions in the

images for a period of time so as to get more information. When there are enough samples, the analyzer can extract some attributes from them to describe the changes of the regions, which will help to analyze the reasons why the changes have taken place in these regions. As a result, if any unusual change occurs, such as the desertification of the natural green area, the illegal exploitation of the greenbelt, or the occurrence of illegal buildings, and so on, it will be detected conveniently. So we think that it may be one of the potential ways to supervise the land resources and the city construction.

3 Experiment and Result

We implemented the coding, the program debugging and the program running of both traditional segmentations in gray image and color-based K-means segmentations in Matlab. And we succeeded in segmenting the satellite image as well as getting statistics of the image using the method mentioned above. The preliminary results are shown in Figure 1.

Due to high discrimination in color scales Lab of marked feature image, the K-means clustering algorithm is of high performance.

After that some resource statistics have been done according to the classification results. And we attempted to understand all of the classes, the results of which turned out to be very close to the ones measured from manual work.

The program can be burned to the VLSI chips of electronic devices, as part of the intelligent detector in the satellite. Large amount of data do not need to transfer to the land station, while the classification result and statistic data can be reserved.

Table 1. The coverage of each entity classes

Entity Class	Coverage	Resource Understanding
Greenbelt	33.92%	Suburb
Roof	2.30%	Factories
Sandy Desertification	29.71%	Construction Fields
Buildings	34.07%	Development Zone

4 Discussion and Conclusion

K-means clustering have been widely used in a large area and ANIL et. al., [15] has discussed such segmentation method that is close related with what we are discussing in this paper. Here we focused more on the semantic understanding and relevant statistics to segmentation results, which are of more significance. Colors are important in the processing of remote electronic photograph for the reason that they usually correspond to some certain fixed classes. As a result, color-based image processing is better, compared to gray-level image processing. We can even enhance the true color composite of the image with a contrast stretch or a decorrelation stretch so as to

optimize the segmentation. Image understanding is a difficult task, because the subjective factors have a large impact on the result. In this paper, we proposed a mapping function which maps clusters' centroids from their 'a*b*' color space to the corresponding classes, attempting to solve the problem. Finally, we got a comparatively good result. The result can be burned to the VLSI chips of electronic devices, working as part of the intelligent machine with CMOS detector in the satellite.

In this paper, we researched on processing the remote electronic photograph of satellite detectors and achieved reasonable statistical results including the coverage of greenbelt, sandy desertification, etc. The methods are meaningful to the monitoring of land resource and the program can be integrated into the intelligent electronic devices. We plan to research on more intelligent classification methods such as SVM with the one presented above, an integrated optimal result will be expected. In view of the fact that the huge numbers of remote electronic photographs, parallel processing of massive data should also be considered in the future.

Acknowledgments. This research is partially supported by the Specialized Research Fund for Doctoral Program of Higher Education of China under Grant No.20100076120011 and the Shanghai Municipal Natural Science Foundation under Grant No.10ZR1410100. It is also partly supported by the Natural Science Foundation of China under Grant No.91118008, the State Key Laboratory of Rail Traffic Control and Safety (Beijing Jiaotong University) No.RCS2011K014, the Innovation Program of Shanghai Municipal Education Commission and ECNU Project "Heterogenous Network Convergence Technologies for CPS".

References

1. Tuia, D., Camps-Valls, G.: Recent advances in remote sensing image processing. In: 2009 16th IEEE International Conference on Image Processing, ICIP, November 7-10, pp. 3705–3708 (2009)
2. Paik, J.K., Katsaggelos, A.K.: Edge detection using a neural network. In: 1990 International Conference on Acoustics, Speech, and Signal Processing, ICASSP 1990, April 3-6, vol. 4, pp. 2145–2148 (1990)
3. Zhou, Y.T.: A circular smoothing edge detector. In: 1990 International Conference on Acoustics, Speech, and Signal Processing, ICASSP 1990, April 3-6, vol. 4, pp. 2153–2156 (1990)
4. Chehdi, K., Liao, Q.: Satellite image segmentation using edge-region cooperation. In: IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, May 19-21, vol. 1, pp. 47–50 (1993)
5. Nsaibi, M., Chaabane, F.: Image fusion of radar and optical remote sensing data for land cover classification. In: 3rd International Conference on Information and Communication Technologies: From Theory to Applications, ICTTA 2008, April 7-11, pp. 1–4 (2008)
6. Hwang, J.T., Chang, K.T., Chiang, H.C.: Satellite image classification based on Gabor texture features and SVM. In: 2011 19th International Conference on Geoinformatics, June 24-26, pp. 1–6 (2011)
7. Tuia, D., Camps-Valls, G.: Semisupervised Remote Sensing Image Classification with Cluster Kernels. IEEE Geoscience and Remote Sensing Letters 6(2), 224–228 (2009)

8. MacQueen, J.B.: Some Methods for classification and Analysis of Multivariate Observations. In: Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, pp. 281–297. University of California Press, Berkele (1967)
9. K-means clustering From Wikipedia, http://en.wikipedia.org/wiki/K-meansclustering
10. Google Map website, http://ditu.google.cn
11. Baidu Map website, http://map.baidu.com
12. DigitalGlobe website, http://www.digitalglobe.com
13. Matlab R2010b Product Help. Help/Video and Image Processing Blockset/Blocks/ Conversions/Color Space Conversion (2010)
14. Rekik, A., Zribi, M., Benjelloun, M., ben Hamida, A.: A k-Means Clustering Algorithm Initialization for Unsupervised Statistical Satellite Image Segmentation. In: 2006 1st IEEE International Conference on E-Learning in Industrial Electronics, December 18-20, pp. 11–16 (2006)
15. Chitade, A.Z., Katiyar, S.K.: Colour based image segmentation using k-means clustering. International Journal of Engineering Science and Technology (10), 5319–5325 (2010)

Research on the Extraction of Image Texture Based on Fractional Kalman Filter Algorithm in Electronic Image Processing

Zheng Wang and Xianmin Ma

School of Electric and Control Engineering
Xi'an University of Science and Technology
Xi'an, China
wendy7830@126.com, Maxm@xust.edu.cn

Abstract. This paper studies the Fractional Kalman Filter and its applications to the extraction of image texture in electronic image processing. Firstly, with the introduction of fractional calculus, Fractional recursive equation is presented. Next, a state space model of a given image and, based on this, the Fractional Kalman Filter algorithm are proposed. Finally, an example of electronic image processing is given to demonstrate the effectiveness of proposed algorithm and the simulation result shows that the details of the image are enhanced, while the background noise of the image is attenuated efficiently.

Keywords: Fractional Kalman Filter, texture, Image enhancement, state-space model.

1 Introduction

Fractional calculus is the theory of any order differential and integral. It is unified with integer order calculus, while is the generalization of the integer order calculus. It is proposed many years ago, but It is only 20 decades that it has been applied to the different realms broadly. At present, In the different realms, such as electronic image processing, electrochemistry, thermodynamic energy engineering, acoustic, electromagnetic field, machinery and control...etc, engineers are studying to create the mathematical models of system by fractional calculus equation so as to describe the dynamic characteristic of system better.

Gray values between Neighborhood pixels in a digital image have a high degree of self-similarity; a high degree of self-similar fractal image signal is often manifested by complex characteristics of texture detail. Different from integer order differential, fractional differential value of DC or low-frequency signal is not 0. fractional differential processing of signals both enhances nonlinearly the high frequency components of the signal, and to a certain extent strengthens nonlinearly the medium-frequency components of the signal, while also significantly retains nonlinearly the DC and low frequency components of the signal.

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 529–532.
springerlink.com

2 Fractional Calculus Modeling

The In a two-dimensional space, for example, two-dimensional linear discrete state-space fractional differential model:

$$\begin{cases} s(m,n) = \sum_{k,l=0}^{W} \sum_{j=0}^{m-1-k,n-1-k} A_{kl}^{j} s(m-1-k-j, n-1-l-j) + Bu(m,n) + \omega(m,n) \\ r(m,n) = Cs(m,n) + \upsilon(m,n) \end{cases} \tag{1}$$

Where $s(m,n)$ and $r(m,n)$ are system state and observed value, separately. $u(m,n)$ is controlling input, $\omega(m,n)$ is system noise, $v(m,n)$ is observed noise. C is transfer matrix, W is neighborhood width, A_{kj} is corresponding system parameter matrix. Based on the definition of fractional calculus, a-order fractional differential state equation:

$$A_{kj} s(m-k, n-l) = \sum_{j=0}^{m-1-k,n-1-k} A_{kl}^{j} s(m-1-k-j, n-1-l-j) \tag{2}$$

Where,

$$A_{kl}^{j} = \begin{cases} diag\left[-(-1)^{j+1} \begin{pmatrix} a_i^{mn} \\ j+1 \end{pmatrix} \right], k = l; \\ diag\left[-(-1)^{j+1} \begin{pmatrix} a_i^{m} \\ j+1 \end{pmatrix} \right], k > l; j > 0 \\ diag\left[-(-1)^{j+1} \begin{pmatrix} a_i^{n} \\ j+1 \end{pmatrix} \right], k < l; \end{cases} \tag{3}$$

$$A_{kl}^{0} = A_{kl}, \qquad j = 0$$

$$\begin{pmatrix} a_i^z \\ j \end{pmatrix} = \frac{a_i^z (a_i^z - 1) \cdots (a_i^z - j + 1)}{j!}, \qquad j = 1,2,3,\cdots \tag{4}$$

$a_i^z \in R^{*+} (z = m,n,mn)$ is the fractional order of $s_i(m,n)$, j is the length of data series.

3 Fractional Kalman Filter Algorithm Modeling

The application of Fractional calculus to image processing can enhance effectively the image texture feature. Fractional Kalman filter algorithm improves signal estimation accuracy, but only for one-dimensional space, and high-performance filtering of image signal system for two-dimensional space model remains to be further studied. This article is the comprehensive application of fractional calculus and Kalman filtering. Aiming at the image texture detail and the edge information fuzziness in the

denoising process by Kalman filter, we put forward a two-dimensional fractional Kalman filtering algorithm and apply the algorithm to denoise and enhance noise image, and finally get the validation of the algorithm.

Two-dimensional fractional order Kalman filter algorithm is:

$$\hat{s}(m,n) = \tilde{s}(m,n) + K(m,n)\left[r(m,n) - C\tilde{s}(m,n)\right] \tag{5}$$

$$K(m,n) = \tilde{P}(m,n)C^T\left[C\tilde{P}(m,n)C^T + R_{m,n}\right]^{-1} \tag{6}$$

$$P(m,n) = \left[I - K(m,n)C\right]\tilde{P}(m,n) \tag{7}$$

$$\tilde{P}(m,n) = \sum_{k,l=0}^{W}\sum_{j=0}^{m-k-1,n-l-1} A_{kl}^{j} P(m-k,n-l) A_{kl}^{j\,T} + \sum_{k,l=0}^{W}\sum_{j=0}^{m-k-1,n-l-1} A_{k,l}^{j} Cov\begin{pmatrix}\breve{s}(m-1-k-j,\\ n-1-l-j)\\ \breve{s}(m-1-k-j,\\ n-1-l-j)\end{pmatrix} A_{k',l'}^{j\,T} \tag{8}$$

$$+ Q_{m,n}$$

Where $\tilde{s}(m,n) = E\left[s(m,n)\middle|z_{-1}^{*}\right]$, $\hat{s}(m,n) = E\left[s(m,n)\middle|z^{*}\right]$, $K(m,n)$ is gain matrix, $R_{m,n}$ is Observation noise variance, $Q_{m,n}$ is noise variance. Z_{-1}^{*} and Z^{*} are observations sequences of the previous step and the current ,Respectively.

$$\breve{s}(m,n) = \hat{s}(m,n) - s(m,n) \tag{9}$$

There are three steps that fractional differential filtering algorithm is applied in image enhancement and denoising: the first step is to determine the back length l and the order parameters a, the second is to convert the original two-dimensional state space equation to two-dimensional linear discrete state-space fractional differential, and finally to process image with the algorithms.

4 Simulating Model and Resulting

We use asymmetric half-plane model to describe the image in electronic image processing. The corresponding two-dimensional discrete fractional differential state space model is:

$$\begin{aligned} s(m,n) &= \sum_{j=0}^{10} A_{0,0}^{j} s(m-j-1,n-j-1) + \sum_{j=0}^{10} A_{1,1}^{j} s(m-j-2,n-j-2) + \sum_{j=0}^{10} A_{2,2}^{j} s(m-j-3,n-j-3) \\ &+ \sum_{j=0}^{10} A_{1,0}^{j} s(m-j-2,n) + \sum_{j=0}^{10} A_{2,0}^{j} s(m-j-3,n) + \sum_{j=0}^{10} A_{2,1}^{j} s(m-j-3,n-1) + \sum_{j=0}^{10} A_{0,1}^{j} s(m,n-j-2) \\ &+ \sum_{j=0}^{10} A_{0,2}^{j} s(m,n-j-3) + \sum_{j=0}^{10} A_{1,2}^{j} s(m-1,n-j-3) + \omega(m,n) \end{aligned} \tag{10}$$

$$r(m,n) = Hs(m,n) + \upsilon(m,n) \tag{11}$$

Where A_{kl} is obtained by recursive least square estimation, H is the identity matrix, Given that the fractional coefficients is determined according to experiment: $a = 0.55$.

Fig. 1. The original image effect

Fig. 2. The image effect by fractional Kalman filtering

This article raises two-dimensional fractional Kalman filtering algorithm, and the algorithm can filter and enhance the image of Figure 1 which is mixed with Gauss white noise (mean value is zero, variance is 0.1). Image filtering effect is shown in Figure 2. Experimental results show that two-dimensional fractional Kalman filtering algorithm not only reduces the noise in the image, but also improves image detail and enhances image quality.

5 Conclusion

Based on fractional-order differential and Kalman filter theory, this article puts forward a two-dimensional fractional Kalman filtering algorithm of discrete state space. We Use two-dimension fractional Kalman filter algorithm for electronic image noise processing. Finally, the algorithm reduces effectively the image background noise and enhances the image feature details, so that the validity of the algorithm is verified in electronic image processing.

References

1. Podlubny, I.: Numerical solution of ordinary fractional differential equations by the fractional difference method. In: Proc. of the Int. Conf. in Difference Equations, London, pp. 507–516 (2007)
2. Zhang, S.Q.: The existence of positive solution for a nonlinear fractional differential equation. Math. Analysis Applied 252(4), 804–812 (2000)
3. Foster, I., Kesselman, C.: The Grid: Blueprint for a New Computing Infrastructure. Morgan Kaufmann, San Francisco (1999)
4. Babakhani, A., Daftardar Gejji, V.: Existence of positive solutions of nonlinear fractional differential equation. Math. Analysis Applied 278(2), 434–442 (2003)
5. Czogala, E.: Idea of a rough fuzzy controller and its application to the stabilization of a pendulum-car system. Fuzzy Sets and Systems 73, 61–73 (2005)
6. Takagi, T., Sugeno, M.: Fuzzy identification of systems and its applications to modeling and control. IEEE Trans. Syst., Man, Cybern. 15, 116–132 (2006)

A Novel Fuzzy Control Algorithm of Hybrid Digital Signal Processing

Gang Liu, Mei-jiao Liu, and Yong-jin Liu

School of Science, Shenyang Aerospace University,
110136 Shenyang, China
liuganglj1p@126.com

Abstract. In this paper, a new efficient control algorithm is proposed by coupling differential equations. Our unified model is natural for posing and solving hybrid analysis and controlling problems. We handle the famous asynchronous arbitrary problem in continuous and hybrid settings and analyze a hybrid digital signal processing that logically switches between two conventional controllers; meanwhile, global asymptotic stability for a realistic set of cases is proved. We systematize the notion of a hybrid digital signal processing by a hybrid controller using an optimal control framework. Theoretical results lead us to algorithms for synthesizing such hybrid controllers based on a generalized Bellman equation, impulse control algorithms, and linear programming.

Keywords: Fuzzy control, hybrid digital signal processing, stability of algorithms.

1 Introduction

The notion of dynamical system has a long history as an important conceptual tool in science and engineering [1]. Complex systems typically possess a hierarchical structure, characterized by continuous variable dynamics at the lowest level and logical decision-making at the highest. Virtually all control systems today perform computer-coded checks and issue logical as well as continuous-variable control commands. Such are hybrid systems. Traditionally, the hybrid nature of these systems is suppressed by converting them into either purely discrete or continuous entities [2]. Motivated by real-world problems, we identify the discrete phenomena that arise in hybrid systems and review previously proposed models.

2 Hybrid Dynamical Systems

As we all know, a dynamical system is a system

$$\sum=[X,\Gamma,\phi] \tag{1}$$

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 533–539.
springerlink.com

Where X is an arbitrary topological space, the state space of $\sum$. The transition semi-group Γ is a topological semi-group with identity [3]. The (extended) transition map $\phi : X \times \Gamma \to X$ is a continuous function satisfying the identity and semi-group properties. A transition system is a dynamical system as above, except that ϕ need not be continuous.

Briefly, a hybrid dynamical system is an indexed collection of dynamical systems along with some map for "jumping" among them [4,5]. Hence, the entire system can be thought of as a sequential patching together of dynamical systems with initial and final states, the jumps performing a reset to a initial state of a dynamical system whenever a final state is reached. A hybrid dynamical system is defined as follows:

Definition 1. A hybrid system is a general hybrid dynamical system with Q countable, and with $\Gamma_q \equiv R$ (or R_+) and $X_q \subset R^{d_q}$, $d_q \in R_+$, for all $q \in Q$. In the notation above, it is written as

$$[Q,[\{X_q\}_{q\in Q},R,\{f_q\}_{q\in Q}],A,G] \tag{2}$$

Where f_q is a vector field on $X_q \subset R^{d_q}$.

More formally, a general hybrid dynamical system (GHDS) is a system

$$H = [Q,\sum,A,G] \tag{3}$$

With its constituent parts defined as follows.

Q is the set of index states, also referred to as discrete states. $\sum=\{\sum_q\}_{q\in Q}$ is the collection of constituent dynamical systems[6],

where each $\sum_q = [X_q,\Gamma_q,\phi_q]$ is a dynamical system as above.

Here, the X_q are the continuous state spaces and ϕ_q are called the continuous dynamics.

$A = \{A_q\}_{q\in Q}, A_q \subset X_q$ for each $q \in Q$, is the collection of autonomous jump sets.

$G = \{G_q\}_{q\in Q}$, where $G_q : A_q \to \bigcup_{q\in Q} X_q \times \{q\}$, is the collection of (autonomous) jump transition maps. These are also said to represent the discrete dynamics of the hybrid dynamical system [7].

Thus, $S = \bigcup_{q\in Q} X_q \times \{q\}$ is the hybrid state space of H. For convenience, we use the following shorthand. $S_q = X_q \times \{q\}$ and $A = \bigcup_{q\in Q} A_q \times \{q\}$ is the autonomous jump set. $G : A \to S$ is the autonomous jump transition map, constructed component wise in the obvious way. The jump destination sets $D = \{D_q\}_{q\in Q}$ are given by

$D_q = \pi_1[G(A) \bigcap S_q]$, Where π_i is projection onto the ith coordinates. The switching or transition manifolds, $M_{q,p} \subset A_q$ are given by $M_{q,p} = G_q^{-1}(p, D_p)$, i.e., the set of states from which transitions from index q to index p can occur [8].

3 Algorithms of Hybrid Dynamical Systems

3.1 Travertine Model

Travertine introduces and discusses so-called differential automata. A differential automaton, A, is a system (S, f, v) where S is the state-space of $A, S = R^n \times Q, Q \tilde{=} \{1, 2, \cdots, N\}$ is the discrete state space of A, and R^n is the continuous state space of A [9];

f is a finite family $f(\cdot, q): R^n \to R^n, q \in Q,$ of vector fields, the continuous dynamics of A; $v: S \to Q$ is the discrete transition function of A.

Let $v_q \equiv v(\cdot, q), q \in Q.$ Define $I(q) = v_q(R^n) \setminus \{q\}$, that is, the set of discrete states "reachable in one step" from q. We require that for each $q \in Q$ and each $p \in I(q)$ there exist closed sets

$$M_{q,p} \equiv v_q^{-1}(P) \tag{4}$$

The sets $\partial M_{q,p}$ are called the switching boundaries of the automaton A. Define $M_q = \bigcup_{p \in I(q)} M_{q,p}$ and define the domain of capture of state q by

$$C(q) \equiv R^n \setminus M_q = \{x \in R^n \,|\, v(x, q) = q\} \tag{5}$$

The equations of motion are

$$\dot{x}(t) = f(x(t), q(t)) \tag{6}$$

$$q^+(t) = v(x(t), q(t)) \tag{7}$$

with initial condition [10] $[x(0), q(0)]^T \in \bigcup_{q \in Q} C(q) \times \{q\}$.

The notation t^- indicates that the discrete state is piecewise continuous from the right. Thus, starting at $[x_0, i]$, the continuous state trajectory $x(\cdot)$ evolves according to $\dot{x} = f(x, i)$. If $x(\cdot)$ hits some $\partial M_{i,j}$ at time t_1, then the state becomes $[x(t_1), j]$, from which the process continues.

Travertine places restrictions on the model above: First each $f(\cdot, q), q \in Q$, is assumed to be globally Lipschitz so that the continuous dynamics are well-behaved. Also, for each $q \in Q$ and $p \in I(q)$, the set $M_{q,p}$ is required to be connected and there must exist a function

$g_{q,p} \in C^1(R^n, R)$ with 0 in its image a regular value such that

$$M_{q,p} = \{x \in R^n \,|\, g_{q,p}(x) \geq 0\}. \tag{8}$$

Thus, $v_q^{-1}(p)$ is a $n-$ sub manifold of R^n with boundary

$$\partial M_{q,p} = \{x \in R^n \,|\, g_{q,p}(x) = 0\} \tag{9}$$

which is an $(n-1)-$ sub manifold of R^n.

3.2 Back-Guckenheimer-Myers Model

The framework proposed by Back, Guckenheimer, and Myers in [1] is similar in spirit to the Travertine model. This is done through transition functions defined on the switching boundaries. Also, the model allows a more general state space.

More specifically, the model consists of a state space

$$X = \bigcup_{q \in Q} X_q, \quad Q \tilde{=} \{1, 2, \cdots, N\} \tag{10}$$

Where each X_q is a connected, open set of R^n. Notice that the sets X_q are not required to be disjoint. The continuous dynamics are given by vector fields $f_q : X_q \to R^n$. Also, one has open sets U_q such that $\overline{U_q} \subset X_q$ and ∂U_q is piecewise smooth. For $q \in Q$, the transition functions $G_q : X_q \to X \times Q$ given the jumps that take place when the state in X_q hits ∂U_q.

They must satisfy $\pi_1(G_q(x)) \in \overline{U}_{\pi_2}(G_q(x))$,

Where π_k k is the kth coordinate projection function. Thus, $\pi_1(G_q(x))$ is the "continuous part" and $\pi_2(G_q(x))$ is the "discrete part" of the transition function.

The dynamics are as follows. The state starts at point x_0 in U_i. It evolves according to $\dot{x} = f_i(x)$. If $x(\cdot)$ hits some ∂U_i at time t_1, then the state instantaneously jumps to state ξ in $\overline{U}_j$, where $G_i(x(t_1)) = (\xi, j)$. From there, the process continues. We refer to this as the BGM model[12].

It is assumed that the switching boundaries ∂U_q have a concrete representation in terms of the zeros of $h_q \equiv \min\{h_{q,1}, \cdots, h_{q,N_q}\}$

where the $h_{q,i} : X_q \to R$ are smooth. The convention then is such that $h_q > 0$ on U_q. Thus, the switching boundaries are $(n-1)$-dimensional Lipschitz continuous manifolds.

However, its expressiveness does allow the possibility of some seemingly "anomalous" behavior. For example, since one allows jumping to the boundary of the sets U_i, trajectories may infinitely "cycle" if $G_j(x) = (y, i)$ and $G_i(y) = (x, j)$.

For simplicity, though, the destination sets are specified a priori instead of by the collection of set-valued maps F. Also, with no real loss of generality, we consider $U_i \equiv U, V_i \equiv V, i \in Z_+$. However, we do generalize to allow the vector field to depend on the continuous state at the last jump time.

Specifically, our discrete state space is $Q = Z_+$. The continuous state space for $x(\cdot)$ is $X = \{X_i\}_{i=0}^{\infty}$ where each X_i is a subset of some Euclidean space R^{d_i}, $d_i \in Z_+$. We also specify a priori regions $A_i, C_i, D_i \subset X_i, i \in Z_+$. These are the autonomous jump sets, controlled jump sets, and jump destination sets, respectively. Let A, C and D denote the unions

$$\bigcup_i A_i \times \{i\}, \bigcup_i C_i \times \{i\}, and \bigcup_i D_i \times \{i\}, i \in Z_+,$$

respectively. Let U, V be the sets of continuous and discrete controls, respectively. The following maps are assumed to be known:

1. Vector fields $f_i : X_i \times X_i \to R^{d_i}, i \in Z_+$.
2. Jump transition maps $G_i : A_i \times V \to D$.
3. Autonomous transition delay $\Delta_{a,i} : C_i \times D \to R_+$.
4. Controlled transition delay $\Delta_{c,i} : C_i \times D \to R_+$.

As shorthand, we may define $G : A \times V \to D$ in the obvious manner. Similarly, for Δ_a and Δ_c.

The dynamics of the control system can now be described as follows. There is a sequence of pre-jump times $\{\tau_i\}$ and another sequence of post-jump times $\{\Gamma_i\}$ satisfying

$$0 = \Gamma_0 \le \tau_1 < \Gamma_1 < \tau_2 < \Gamma_2 < \cdots \le \infty,$$

such that on each interval $[\Gamma_{j-1}, \tau_j)$ with non-empty interior, $x(\cdot)$ evolves according to the differential equation

$$x(t) = \xi(t), \quad t \ge 0 \quad \text{in some } X_i, i \in Z_+.$$

3.3 Algorithms of Optimal Hybrid Control

In this section, we give explicit algorithms for computing the solutions to such optimal control problems.

We now discuss explicit solutions of the GQVIs. First, note that instead of thinking of V defined on the generalized state-space in the previous section, $V : \bigcup_i X_i \to R_+$, we may revert to considering it component-wise: $V_i : X_i \to R_+$.

The Boundary-Value algorithms outlined in this section begin with a guess for each V_i .

The algorithm is summarized as follows :

1. Given: a guess for each V_i .

2. For each i , compute the constrained boundary values arising from the GQVIs:

(a) For $x \in A_i$ impose the constraint

$$V_i(x) = \min_v \{c_a(x,v) + e^{-a\Delta_a} V(G(x,v))\}$$

(b) For $x \in C_i$ impose the constraint

$$V_i(x) = \min\{V_i(x), \min_{z \in D}\{c_c(x,z) + e^{-a\Delta_c(x,z)} V(z)\}\},$$

3. Solve the GQVI Equation to update each V_i separately, imposing the constraints of Step 2.

4. If convergence, exit; else go to Step 2.

4 Conclusions

In this paper, we discuss topological issues that arise in hybrid systems analysis. Then We proposed a very novel framework for hybrid control algorithm that was shown to encompass all hybrid phenomena considered in this thesis and all hybrid models reviewed from the literature. A specific control problem in hybrid digital signal processing was studied in this framework, leading to an existence result for optimal controls.

Acknowledgement. This research is supported in part by the Young Teacher's Foundation of Shenyang Aerospace University (201109Y), China. We also thank the anonymous referees for useful suggestions on the earlier draft of the manuscript.

References

1. Hirsch, M.W., Smale, S.: Differential Equations, Dynamical Systems, and Linear Algebra. Academic Press, San Diego (1974)
2. Luenberger, D.G.: Introduction to Dynamic Systems: Theory, Models, and Applications. Wiley, New York (1979)
3. Tanaka, K., Sugeno, M.: Stability analysis and design of fuzzy control systems. Fuzzy Sets Syst. 45, 136–156 (1992)
4. Castillo, O., Melin, P.: Intelligent adaptive model-based control of robotic dynamic systems with a new hybrid neuro-fuzzy-fractal approach. In: Proceedings of FLINS 2000, pp. 351–358. World Scientific, Bruges (2000)
5. Li, W., Chang, X., Farrell, J., Wahl, F.: Design of an enhanced hybrid fuzzy p+id controller for a mechanical manipulator. IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics 31 (December 2001)

6. Sun, L., Er, M.: Hybrid fuzzy control of robotics systems. IEEE Transactions on Fuzzy Systems 12, 755–765 (2004)
7. Badamshin, R.A., Ilyasov, B.G., Cherniakhovskaya, L.R.: The problems of the complex dynamic objects control in critical situations on the base of knowledge. Mashinostroenie, Moscow (2003)
8. Zadeh, L.A.: Outline of a new approach to the analysis of complex systems and decision processes. IEEE Trans. on Systems, Man and Cybernetics 3(1), 28–44 (1973)
9. Chen, B.S., Tseng, C.S., Uang, H.J.: Mixed fuzzy output feedback control design for nonlinear dynamic systems: An LMI approach. IEEE Trans. Fuzzy Syst. 8, 249–265 (2000)
10. Zadeh, L.A.: Fuzzy sets. Information and Control 8, 338–353 (1965)
11. Shing, J., Jang, R., Chuen-Tsai, S.: Neuro-fuzzy Modeling and Control. Proceeding of IEEE (March 1995)
12. Wang, L., Langari, R.: Complex systems modeling via Fuzzy logic. IEEE Trans. Syst. Man, Cybern. 26, 100–106 (1996)

Maya-Based Simulation of Tree Movement with Electronic Image Processing

ZhanLi Li and JiaXuanZi Yan

College of Computer Science and Technology, Xi'an University of Science and Technology, 710054 Xi'an, China
yanjiaxuanzi@sina.com

Abstract. According to the motion characteristics of the trees in the wind, the paper builds a realistic virtual simulation of three-dimensional plant movement on Maya with electronic image processing. The main purpose is to provide realistic effect for many applications fields, such as games, movies and so on. Firstly, the tree model is established on basis of polygon modeling tools with electronic image processing. Secondly, the force is analyzed and bone is created so as to set reasonable bone controller to drive the movement of the tree model. Finally, we realized the swing and fracture based on the principle of dynamic and cloth solver is adopted to make the swing and falling of the tree leafs more natural. The whole process mainly includes the gravity field; the wind field and the key frame technology. Experiments show that the method not only reduces the development time and work, but also makes the movement of trees more realistic.

Keywords: Plant modeling, Skinned mesh, Dynamics, Key frame, Maya, Electronic Image Processing.

1 Introduction

Maya is considered as the most powerful 3D development software, especially, it is particularly appropriate for complicated scenarios. This paper uses Maya technique to complete the simulation of tree movement with electronic image processing with electronic image processing. Maya is characterized by high speed and high performance. The character animation system has the weight brush stroke adjustment, and the nonlinear character editing system, powerful particles, rigid body dynamics and flexible body dynamics can be used to analyze object movements.

Computer-aided 3D movement simulation of plants has become an important research field in computer graphics. Along with the development of computer animation and increasing demands in game field, the movement simulation of trees in the wind has become a hot topic. Domestic and foreign studies on movement simulation of trees focus mostly on overall movement simulation of trees, and the studies on exfoliation of leaves and fracture of branches in the strong wind are still a few. This paper, based on Aerodynamics and Mechanics of Materials [1-3], provides the solutions to movement simulation of dynamic trees based on Maya dynamics

D. Jin and S. Lin (Eds.): Advances in Mechanical and Electronic Engineering, LNEE 178, pp. 541–547.
springerlink.com

principles and key frame techniques. Maya can reduce the computing volume, better achieve 3D dynamic tree model, and greatly improve the sense of reality of virtual natural environment.

2 Maya-Based Tree Modeling with Electronic Image Processing

2.1 Tree Modeling with Electronic Image Processing

Use the Polygons modeling tool to make the model [4]; mainly use the points, lines and surfaces to control the polygon; use the extruding surface to make tree model. Firstly, create a basic body. The branch cross-section is circular, so we can choose the cylinder to make the model [5]. Modify the polygon to adjust the basic shape, click Extrude the polygon and then get the upper shape of branches; make forked branches, and then continue to extrude the branches.

Subsequently, it is the production of complicated forked branches. In consideration of leave model and placement, combine with growth law of real trees in the nature and tree pictures; try to make the model more realistic. When you make the leave model, you should firstly consider the form of rear map. This paper uses transparent map to make the leaves, so you need not make the shape of the model, and later it will use the maps to make the shape. This paper only needs to use the polygon modeling tool to make the slice. After a slice, copy the slice to place, click Duplicate Special to copy. Leave models are placed according to the growth relationship between true branches and leaves [6]. Leaves and branches are connected according to the bones of branches. The final effect is shown in Fig 1.

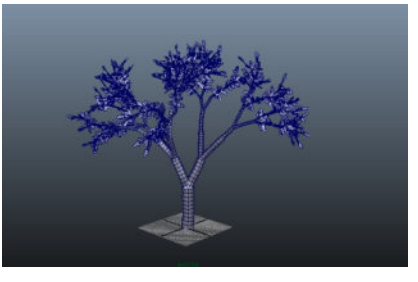


Fig. 1. Overall effect of tree model with electronic image processing

2.2 Texture Mapping of Branches and Leaves

Make the map of the branches [7], before mapping, use the Cylindrical Mapping to project the lower cylinder, select Polygons mode and then click Cylindrical Mapping to allocate UV for models, and achieve accurate alignment of maps. Open the UV editor, and split into the plane. The rest of branches are split like this. After UV is split, paste the pictures on branches. Open the Material Editor, create material balls and then give them to branches. Subsequently, add a texture for material balls, and find a branch picture for models. Next, create the leave map. This paper uses transparent map. The transparent map can provide transparent edges.

3 Tree Movement with Electronic Image Processing Method

Skeletal animation, a 3D graphics technique[8], is often used to simulate the movement of human body or animal body. It uses joint structure as the skeleton, connects the role bind onto the skeleton, and uses joint movement to control bind deformation. Skeletal animation only needs to store the data of animation control of the skeleton, whereas the number of joint nodes in the skeleton is far less than that of skin mesh vertices. Due to such advantage, it can quickly achieve real-time bind deformation on low-end hardware. This paper will use the skeletal animation to simulate the plant simulation movement, as the above mentioned, it is still applicable to the joint structure as the skeleton, connecting tree skin onto the skeleton and using joint movement to control tree deformation.

3.1 Force Analysis of Tree Movement with Electronic Image Processing

Trees model can be as a whole because bone has been completed bind with model so that the movement of the bones led the entire movement of the tree model. We just only need to analyze the stress of trunk controller point, instead of stress analysis for each node in the branch, because bone controllers made mechanical analysis of the branches movement simplified. Controller can set the key frame according to the point of stress analysis and the calculated results, so as to drive the movement of the whole tree. In order to facilitate the calculation of the deformation of the branches, the local coordinate is established for tree, which is shown in Fig 2.

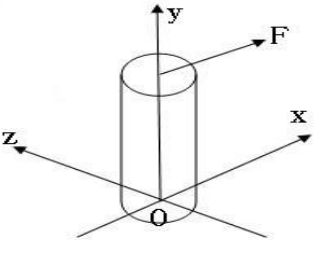


Fig. 2. The local coordinate of tree

The starting point of the local coordinate system O is set to the starting point of the branches。 the z-axis and y-axis perpendicular to the direction of motion of branches, while the z-axis perpendicular to the force F and the y-axis is set to the direction of the branches.

The force F can be decomposed into a force Fx and force Fy , the deformation caused by FZ is not taken into account. We use the following formula to calculate the deformation caused by the force Fx [3].

$$x = 4F_x L^3 / 3\pi ER^3 \tag{1}$$

L is the length of the branches of a tree, R is the radius of the branches, E is the coefficient of elasticity of the branches. the rotational angle θ of the branches has been simplified because the deformation of the branches is smaller. Branches deformation x can be seen as the arc, and then the deflection angle of branches can be calculated by the following formula:

$$\theta = arcSin(x/L) \tag{2}$$

We can get rotation angle θ and swing displacement in the x-axis of the independent trunk through this method so that the angle of rotation and swing displacement at different time can be calculated according to forces of trunk. We can calculate the movement at different points according to the above method, and then the key frames can be set by bones controller of the trunk based on these data, so that the produced animation will be more natural and realistic.

3.2 Control of Tree Movement with Electronic Image Processing

Firstly, enter the Animation module and find the Joint tool in Skeleton menu; create the Skeleton. Given that the skeleton of trees is created according to the path of branches, the consistent direction should be considered at creating the skeleton, thus there will be no errors when generating the animation. Create the skeleton from the trunk, and then make the skeleton according to branches. There are six skeletons on main trunks. The skeleton hierarchy is arranged according to the sequence of parent-child relationship. The root joint is the highest level among the skeleton. There is only one root joint in the skeleton. The child joint is under the parent joint. In consideration of skeleton fracture later, the skeleton should be made according to the branch location. For instance, the skeleton is created from the branch location, as shown in Fig 3.

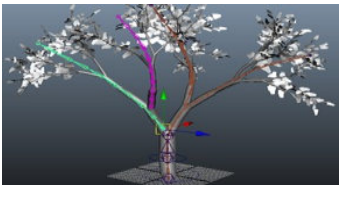


Fig. 3. Skeleton created on branches

After the model and skeleton are created, the created skeleton still cannot boost the movement of branches. We need to connect the skin onto the skeleton, and then use the skeleton to control the skin and generate the right deformation. Maya includes Rigid Bind and Smooth Bind [9]. If you use the Rigid Bind, each vertex of the model will be affected by a joint. When the joint is crooked, it needs to use flexor deformation or change the vertex weight to get a soft curved surface. If you use the Smooth Bind, Each vertex of the model will be affected by multiple joints, the default weighted value of each vertex is 1, and this value is allocated by the affecting joints

according to a certain percentage. The higher the percentage is, the stronger the control will be for the vertex. In consideration of actual conditions of tree movement with electronic image processing in the wind, and in combination with the natural law of tree bending changes, this paper uses the Smooth Bind.

After the skinned mesh is completed, set the controller to freely adjust the animation. Firstly, the movement of branches should be taken into account, and this paper uses IK curve editor to control the movement of branches. Therefore, you'd better add IK curve editor to each branch. Subsequently, create the skeleton, and use the skeleton to bind this curve, thus the skeleton can control the IK. Create a circle curve control for the skeleton, and then exercise parent-child constraint over the circle curve control and the skeleton. Then the controllers are made with the same method.

Subsequently, we should consider the movement rules of branches, so that the animation will be more natural. This paper uses the soft curve IK in the Maya dynamics to control the skeleton. When you create a soft body, the system will automatically create a particle object [10]. To some extent, the soft body is only an aggregation of particles .When the particles move, the points in the geometry will also move. This paper mainly uses the curve to control the branch blowing effect.

4 Simulation of Tree Movement with Electronic Image Processing

4.1 Simulation of Branch Movement

It is necessary to set the key frame animation [11] for the movement of branches. The movement of branches is mainly to set the key frames for the skeleton controller. According to the changes of weight of created soft curve particles, the movement of branches can control the swing of branches. How to make the branches and trunks move together? It only needs to establish the parent-child relationship between branch controller and trunk controller, so if you set the key frames for trunks, the branches will move accordingly.

In the production of branch movement, we should control the animation effect according to rotation angle θ and swing displacement in the x-axis. You can use the animation curve editor to set the changes of animation speed. This paper, according to the computing data of force analysis of branches and leaves set out in Reference [1], adjusts the handle to control the rotation angle θ and swing displacement in the x-axis.

We can use active rigid body and passive rigid body to achieve the effect of branch fracture [12]. Create the active rigid body for branches to be fractured; select the ground to create the passive rigid body. The fractured parts of the branches are fractured according to the segmentation position at modeling. We can control the fracture time according to the key frames. The movement trace of fractured branches is determined according to the movement state of active rigid body, select the branches to be fractured, and then create the gravitational field, so that the branches will move downward driven by the gravitational field; and add a wind field, thus it will form a natural movement trace. The movement trace is controlled mainly according to strength and direction of gravitational field and wind field.

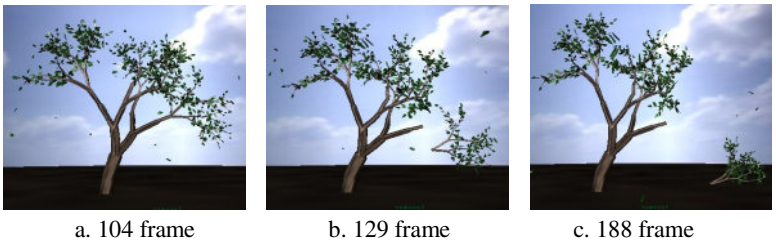
4.2 Simulation of Leave Movement

We can use Maya's Cloth [13] to generate the movement of leaves, select leaves and add the cloth for leaves to simulate the falling effect of leaves; use the Cloth to create the real cloth effect. Through the computation of cloth; carry out dynamics computation according to character animation, and simulate the movement effect of real cloth. The Cloth can easily achieve the special effect, even leave falling, balloon explosion, and object squeezing, etc.

Cloth object is geometry for cloth simulation. Given that the leave is a slice made previously with polygon modeling tool, it is unnecessary to convert it into a polygon object, and it can be a cloth object directly. We can adjust the parameter hardness of clothes and friction to control the movement of leaves. Like the movement of branches, we also need to add the gravitational field and disruption field to the leaves, so that the leaves can fall. The exfoliation of leaves is based on a property of clothes, and such property is used to control the command of cloth settlement, and set the key frames for this property. When it reaches the first threshold value which is set according to the wind force, the leaves will begin to fall. If you activate such property at the falling time, the leaves will begin to compute the clothes.

5 Conclusion

According to the theory above-mentioned, based on the Maya, this paper simulated the tree movement with electronic image processing, including scene simulation, tree swing simulation under the breeze, the tree exfoliation and branch fracture simulation under strong wind, under the action of different wind. The following simulation results show the different state of the tree movement with electronic image processing. as shown in Fig 4.



a. 104 frame b. 129 frame c. 188 frame

Fig. 4. The simulation results of the tree movement with electronic image processing

Based on the study on plant movement in the wind, this paper mainly discusses the key technical issues in the simulation of virtual 3D plant movement which is created with Maya, a professional animation software. Firstly, use the polygon modeling, especially the Extrude function, to create the tree models. And then complete and bind the tree skeleton structure, so that the skeleton can drive tree models to move together. In consideration of real effect of branch swing in the wind, this paper adopts the Smooth Bind to reach more realistic movement effect, uses the soft body dynamics to achieve the swing effect of the branches and cloth computation

to achieve the exfoliation of leaves; the overall simulation process of tree movement with electronic image processing involves the setting of the key frames; the fracture process of branches and exfoliation process of leaves all depend upon the gravitational field and wind field. According to the principle of dynamics and the key frame technique, we can develop Maya-based platform to simulate the movement of trees in different winds, all of these must rely on Maya's powerful technical platform.

Acknowledgement. Project supported by the Special Research Foundation of Education Bureau of Shaanxi Province, China (Grant NO.09JK577).

References

1. Li, Z.L., Liu, Y.T.: Visual Simulation of tree movement with electronic image processing in the Strong Winds. Journal of Computer-Aided Design and Computer Graphics 23, 1365–1372 (2011)
2. Li, Z.L., Norishige, C.: Research on Real-time Animation of Trees Swaying in Wind. Journal of System Simulation 20, 2085–2091 (2008)
3. Li, Z.L., Zhang, M.: Simulation of tree motion based on Vega Prime. Key Engineering Materials 474-476, 1296–1299 (2011)
4. Gao, W.S.: Maya Foundation and Practical Course. Tsinghua University, Beijing (2009)
5. Xiong, Z., Wang, R.J., Chen, L., Yao, L.Y.: Real-time Simulation Design and Realization of Dynamic 3D Trees. Computer Engineering and Design 30, 79–80 (2009)
6. Shi, D.: Study on Scenario-based Plant 3D Reconstruction Techniques. Master Degree Thesis of Jiangsu University, Beijing (2006)
7. Li, G.H., Lu, F.Z., Yu, Q.F.: 3D Tree Shaping Technique Based on Object Methods. Journal of Liaoning University of Technology 22 (2002)
8. Wuhan, R.M.: Maya6 Super Manual (3) Animation Chapter. Weapons Industry Press, Beijing (2005)
9. Wang, Q.: Standard Training Course (I) for Autodesk Maya8. Posts and Telecom Press, Beijing (2007)
10. Cai, Y.: Realization of Galaxy Animation Simulation Based on Maya Platform Particle System. Engineering Master Degree Thesis of Shanghai Jiaotong University (2008)
11. Ye, S.J., Wan, J.: Simulation of Virtual Human Walking Movement Based on Maya Key Frame Technique. Computer Knowledge and Technology (2008)
12. Zhang, B.R.: Maya: Character Animation Chapter. Weapons Industry Press (2007)
13. Wuhan, R.M.: Maya6 Super Manual (4) Dynamics Chapter. Weapons Industry Press (2005)

Author Index