2011 China-Korea-Japan Electronics and Communications Conference (ECC 2011)

Oct. 26 - 28, 2011

Chengdu, Sichuan, P. R. China

Conference Program and Information Booklet

Organized by
University of Electronic Science and Technology of China (UESTC), China

Co-organized by
Yeungnam University, South Korea; and Tohoku University, Japan

Technical sponsored by
IEEE, IEEE Computer Society, and IEEE Technical Committee on Scalable Computing

Table of Contents

Welcome Message

Message from ECC 2011 General Chairs

Welcome to the 2011 China-Korea-Japan Electronics and Communications Conference (ECC 2011), held in Chengdu, Sichuan, China, on Oct. 26-28, 2011. On behalf of the organizing committee of ECC 2011, we would like to express our cordial welcome and high respect to all participants.

Electronics and communications play key roles in modern IT infrastructure, affecting every aspect of life, including services related to health, banking, commerce, defense, education and entertainment. It has been viewed that there are tremendously increasing demands for computing, communication and storage systems since last decades. To fulfill such requirements, modern IT infrastructures have been extended and scaled, which brings various related issues, such as electronic engineering, mobile communications, energy consumption, electromagnetic and wavelet, radar technology, etc.

ECC 2011 is an annual meeting which has been bringing together the leading scientists from academia and industry since 2002. Following the previous successful ECC conferences, ECC 2011 will focus on cutting edge research, development, and applications of electronics and communications areas.

ECC 2011 is supported by many people and organizations. We want to thank Prof. Houjun Wang, Prof. Yong Fan, and Prof. Jun Hu, the Steering Chairs of ECC 2011, for giving us the opportunity to hold this conference and for their guidance on organizing the conference. We also would like to express our appreciation to Program Chairs, Prof. Liping Li and Prof. Xingang Liu, for their hard and excellent work on organizing a very strong program committee. We appreciate all authors to submit your high-quality papers to ECC 2011. Thanks to all of you for participating ECC 2011, and hope you find the conference stimulating and instructive, at the same time enjoy the city of Chengdu, Sichuan, China.

Oct. 26th, 2011

General Chair of ECC 2011
Prof. Jin Pan
University of Electronic Science and Technology of China, China

Message from ECC 2011 Program Chairs

Welcome to the 2011 China-Korea-Japan Electronics and Communications Conference (ECC 2011), held in Chengdu, Sichuan, China, on Oct. 26-28, 2011. The goal of ECC 2011 is to provide a forum for scientists, engineers, and researchers to discuss and exchange their new ideas, novel results, work in progress and experiences on all aspects of electronics and communications, as well as to identify emerging research topics and define the future directions.

This time, ECC 2011 attracted numerous submissions, and all submitted papers went through a rigorous and objective selection process. Each paper received reviews from the program committee members and the external reviewers. As the result, papers are accepted by ECC 2011.

Similar with previous events, the oral and poster presentation types for accepted papers are adapted in ECC 2011. We are delighted to say that there are 24 oral papers and 20 poster papers. We also organized a keynote speaking by Prof. Qun Wan to address recent advances in electronics and communications. Besides the conference program, we have also organized several journal special issues, such as Journal of Universal Computer Science (SCIE Indexed), Concurrency and Computation: Practice & Experience (SCI Indexed), IET Communications (SCI Indexed), and the outstanding papers of ECC 2011 will be included in them with further review.

Several organizations provided financial and logistical assistances, and we are grateful for their support. We thank University of Electronic Science and Technology of China for hosting the conference and carrying the burden of local organization. We owe a great deal of thanks to the program committee members and the reviewers. The success of this year's ECC would not be possible without their hard work, and their professional contributions maintain the high standards of this year's technical program. We are also grateful to all the Steering Committee members for their helpful advices and constant support.

Finally, we extend a sincere "Thank You" to all the authors who entrusted their work to ECC 2011 and to everyone who attended the conference and enjoyed the program we helped to assemble. None of this would have been possible without the authors, program committee members, thoughtful reviewers, and our sponsors. Your continuous contribution, participation, and support to this conference make this entirely volunteer effort worthwhile.

Program Co-chairs of ECC 2011
Prof. Liping Li, UESTC, China
Prof. Xingang Liu, UESTC, China

Organization

ECC 2011 Organizing Committee

Steering Chairs
Houjun Wang, University of Electronic Science and Technology of China, China
Yong Fan, University of Electronic Science and Technology of China, China
Jun Hu, University of Electronic Science and Technology of China, China

General Chairs
Jin Pan, University of Electronic Science and Technology of China, China
Byoungchul Ahn, Yeungnam University, South Korea
Fumiyuki Adachi, Tohoku University, Japan

Program Chair
Liping Li, University of Electronic Science and Technology of China, China

Program Co-Chair
Xingang Liu, University of Electronic Science and Technology of China, China

International Advisory
Yiming Pi, University of Electronic Science and Technology of China, China
Ming Lv, University of Electronic Science and Technology of China, China
Hongsheng Zhong, University of Electronic Science and Technology of China, China
Yong-Wan Park, Yeungnam University, South Korea
Kook-Yeol Yoo, Yeungnam University, South Korea
Sung-Won Kim, Yeungnam University, South Korea
Takahiro Hanyu, Tohoku University, Japan
Masahiro Yamaguchi, Tohoku University, Japan
Hirohito Yamada, Tohoku University, Japan
Wei Peng, Tohoku University, Japan
Shunsuke Yamaki, Tohoku University, Japan
Yoshihiro Sugaya, Tohoku University, Japan

Publication Chair
Jingfu Bao, University of Electronic Science and Technology of China, China

Local Co-Chairs
Yunchuan Guo, University of Electronic Science and Technology of China, China
Shiwei Qu, University of Electronic Science and Technology of China, China

Conference Secretary
Keyu Chen, University of Electronic Science and Technology of China China (Chair)
Xiao Zhang, University of Electronic Science and Technology of China
Yun Liu, University of Electronic Science and Technology of China

Web and System Chair
Huiyong Li, University of Electronic Science and Technology of China, China

Conference Information

Registration and Conference Venue

Full Registration

Full registration includes one volume of book proceedings and CD where the registered paper is published, meals, reception, banquet, and souvenir.

Banquet Locations

Banquet locations are in the first floor of Hotel of UESTC.

Useful Telephone Numbers

ChengDu International Dialing Code: 86-28

Directory Enquiries: 114

Emergency Service (Police): 110

Emergency Service (Fire): 119

Emergency Service (Ambulance): 120

ChengDu International Airport, English (24 hours): 85205555

ECC 2011 Program

Conference Schedule

Oct. 26th 2011 (Wednesday)	
08:00 - 09:00	Registration
09:00 - 09:15	Welcome speaking: Prof. Jun Hu, Dean of School of EE, UESTC
09:15 - 10:00	Keynote speaking: Prof. Qun Wan, School of EE, UESTC
10:00 - 10:15	Coffee break
10:15 - 12:00	ECC2011 oral session I (Chair: Prof. Jingfu Bao, UESTC)
12:00 - 13:30	Lunch Time (Hongxing Restaurant)
13:30 - 14:30	ECC2011 poster session (Chair: Prof. Shiwei Qu, UESTC)
14:30 - 16:15	ECC2011 oral session II (Chair: Prof. Xingang Liu, UESTC)
16:15 - 16:30	Coffee break
16:30 - 18:15	ECC2011 oral session III (Chair: Prof. Peng Wei, Tohoku University)
18:15 - 20:30	Dinner (Hongxing Restaurant)
Oct. 27th 2011 (Thursday)	
09:00 - 10:45	ECC2011 oral session IV (Chair: Prof. Sung-Won Kim, Yeungnam University)
10:45 - 11:00	Coffee break
11:00 - 12:45	ECC2011 oral session V (Chair: Sugaya Yoshihiro, Tohoku University)
12:45 - 14:00	ECC2011 banquet (Hongxing Restaurant)
14:00 - 16:00	Visiting the new campus of UESTC
16:00 – 17:00	Professor meeting in new campus; free time visiting for students
Oct. 28th 2011 (Friday)	
07:00 - 19:00	Open-air discussion

Oral and Poster Presentation Guidelines

For Oral Presentation

There are five oral sessions during ECC 2011. Paper is allocated 20 minutes, with 15 minutes for presentation and 5 minutes for questions and answers.

Speakers should go to the session room at least 10 minutes before their session starts, introduce themselves to the session chairs and check their presentation material with the computer and audio-visual equipment.

The computer in session room can display MS PowerPoint and Adobe PDF files. Speakers can bring their presentation material on USB drives. If you use other digital storage devices not supported by the computer in the session room, please ask a conference helper to transfer the files.

For Poster Presentation

One poster session is included in ECC 2011. Poster paper is allocated 60 minutes, and the board (100cm*90cm) is prepared for material posting.

Authors should prepare printed ppt materials according to the size of the posting board. During the poster session, corresponding author of each paper should stay with the posting board and communicate with the visitors.

Conference Program

[No. 1 Education Building, UESTC, Oct. 26th, 2011]

Conference Registration

(08:00~09:00, No. 1 Education Building, UESTC)

Chair: Keyu Chen, UESTC

Welcome Speaking: Prof. Jun Hu, Dean of School of EE, UESTC

(09:00~09:15, No. 1 Conference Room, UESTC)

Chair: Prof. Jin Pan, UESTC

Keynote Speaking: Prof. Qun Wan, School of EE, UESTC

(09:15~10:00, No. 1 Conference Room, UESTC)

Chair: Prof. Jin Pan, UESTC

ECC 2011 Oral Session I

(10:00~11:45, No. 1 Conference Room, UESTC)

Session Chair: Prof. Jingfu Bao, UESTC, China

[ECC2011_O_01] Stability criteria for impulsive Cohen-Grossberg neural networks with time-varying delays
C. Zhou, H. Zhang (UESTC, China)

[ECC2011_O_02] Access Point Selection Mechanism to Avoid Rogue Access Point
N. Nazarov, S. Djuraev, S.Y. Nam (Yeungnam University)

[ECC2011_O_03] Sparse Channel Estimation for Multipath Fading Dual-hop Single Relay Cooperation Networks
G. Gui, W. Peng, Q. Wan, F. Adachi (Tohoku University, Japan)

[ECC2011_O_04] A Damping Approach in DMTL's Packaging for Reducing Release Time
C. DENG, J.F. BAO, X.H. ZHAO (UESTC, China)

[ECC2011_O_05] Transient Noise Excision for Skywave Over-The-Horizon Radar Based on Adaptive Gaussian Chirplet Decomposition
W. You, Q. He, Z. He, S. Wang (UESTC, China)

ECC 2011 Poster Session

(13:30~14:30, No. 2 Conference Room, UESTC)

Session Chair: Prof. Shiwei Qu, UESTC, China

[ECC2011_P_01] Stability Analysis of Discrete-Time Large-Scale Systems with Time-Delays
Q. Huang, L. Sheng, G. Yu (UESTC, China)

[ECC2011_P_02] Design and Implementation of an Emergency Response System using GPS
J.Y. Song, Y.W. Park (Yeungnam University, South Korea)

[ECC2011_P_03] Narrow spectral linewidth wavelength tunable laser diodes consisting of SOI ring resonators
K. NEMOTO, T. KITA, H. YAMADA (Tohoku University, Japan)

[ECC2011_P_04] Design of Wireless Monitoring System Based on ZigBee and GPRS Technology
X. Wang, H. Zhang (UESTC, China)

[ECC2011_P_05] Parking Assist System Using Fisheye lens
M.S. Kang, Y.W. Park (Yeungnam University, South Korea)

[ECC2011_P_06] Evaluation of V_{th}-Variation Effect on Multiple-Valued Current-Mode Circuits
K. Kashiuchi, M. Natsui, T. Hanyu (Tohoku University, Japan)

[ECC2011_P_07] A Novel MASH Structure for Fractional-N Synthesizer
B. Zhang, S. He (UESTC, China)

[ECC2011_P_08] Enhancing Radar Angular Resolution Algorithm with Maximum a Posteriori
J. Guan, Y. Huang, J. Yang (UESTC, China)

[ECC2011_P_09] Extraction of 3D Shape of Tooth with Three-Dimensional Contours of CT Images
T. NAGAOKA, R. YANAGISAWA, Y. SUGAYA, S. OMACHI (Tohoku University, Japan)

[ECC2011_P_10] Pattern-Search-based Optimized PM signal design
L. Li, L. Kong, X. Yang (UESTC, China)

[ECC2011_P_11] A Time-domain Sparse Signal Representation Perspective for Speech Denoising
J. Li, Q. Wan (UESTC, China)

[ECC2011_P_12] Delay-Aware Model-Based Accurate Simulator for Asynchronous NoC Design
Y. Watanabe, A. Matsumoto, T. Hanyu (Tohoku University, Japan)

[ECC2011_P_13] A Calibration for Gain and Phase Errors of Sensors Based on A Moving Sensor
T. Wang, P. Wei (UESTC, China)

[ECC2011_P_14] A NOVEL DESIGN FOR FOUR ELEMENTS TIME DOMAIN ULTRA-WIDE BAND ANTENNA ARRAY
P. Li, J. Pan, D. Yang (UESTC, China)

[ECC2011_P_15] Evaluation of Unsupervised Language Model Adaptation using Retrieval of Relevant Documents
S. Sato, R. Masumura, A. Ito (Tohoku University, Japan)

[**ECC2011_P_16**] **Two-Dimensional Simulation Approach on the Mechanism of Microwave Induced Thermo-Acoustic System**
J. Song, Z. Zhao (UESTC, China)

[ECC2011_P_17] A Tight Coupling Directional Coupler from Composite Right/Left-Handed

Transmission Lines
H. Wang, X. Tang, H. Zhang (UESTC, China)

[ECC2011_P_18] RF Magnetic Near Field Measurement using a Magnetic Force Microscope for the Potential Application to Evaluate LTE-Class RFIC
Y. Endo, M. Fukushima, M. Watanabe, K. Arai, Y. Shimada, M. Yamaguchi (Tohoku University, Japan)

[ECC2011_P_19] Two elements time domain antenna research based on trapezoidal cavity-backed design
D. Yang, J. Pan, P. Li, Z. Zhao, Z. Nie (UESTC, China)

[ECC2011_P_20] Analysis of Convergence Characteristic of Adaptive Notch Filters
A.UEKI, S. KOSHITA, M. ABE, M.KAWAMATA (Tohoku University, Japan)

ECC 2011 Oral Session II

(14:00~15:45 No. 1 Conference Room, UESTC)

Session Chair: Prof. Xingang Liu, UESTC, China

[ECC2011_O_06] Active-Gm-RC Bandpass Filter with 60MHz Center Frequency and a Combined Analog-Digital Tuning System
J. Shi, T. Konishi, S. Masui (Tohoku University, Japan)

[ECC2011_O_07] Dimension Reduction Based Direction of Arrive Estimation with an Uncalibrated Distributed Electromagnetic Component Sensor Array
H. Chen, Q. Wan, Y. Liu (UESTC, China)

[ECC2011_O_08] Enhancement of driver-dfficiency and the luminance uniformity in the light-emitting diode back-light units using RLC regulation and nubber circuits
M.S. Moon, J.S. Jang (Yeungnam University, South Korea)

[ECC2011_O_09] L-Coupling Noise Countermeasure for LTE-Class Receiver RFIC by Using Magnetic Thin-Film
W. Kodate, Y. Endo, M. Yamaguchi (Tohoku University, Japan)

[ECC2011_O_10] Multidimensional scaling-based passive emitter location from arrival of angular measurements
X. Shen, S. Qin, Q. Wan, W. Yang (UESTC, China)

ECC 2011 Oral Session III

(14:30~16:15 No. 1 Conference Room, UESTC)

Session Chair: Prof. Peng Wei, Tohoku University, Japan

[ECC2011_O_11] Implementation of Dimmable Visible Light Communication System
S.J. Lee, S.Y. Jung (Yeungnam University, South Korea)

[ECC2011_O_12] Cost Function for Uniformly Distributed Source Signals Using Kullback-Leibler Divergence for Independent Component Analysis
K. TANZAWA, S. KOSHITA, M. ABE, M. KAWAMATA (Tohoku University, Japan)

[ECC2011_O_13] A Novel Training based Channel Estimation Method in flat-block MIMO System
J. Yan, Q. Wan (UESTC, China)

[ECC2011_O_14] MIMO Based AOA Positioning with Phase Difference in Multipath Environments
I.H. Park, K.Y. Yoo (Yeungnam University, South Korea)

[ECC2011_O_15] Joint MMSE-FDE&Spectrum Combining for Antenna Diversity Reception on Single-carrier Transmission
T. OBARA, F. ADACHI (Tohoku University, Japan)

[No. 1 Education Building, UESTC, Oct. 27th, 2011]

ECC 2011 Oral Session IV

(09:00~10:45 No. 1 Conference Room, UESTC)

Session Chair: Prof. Sung-Won Kim, Yeungnam University, South Korea

[ECC2011_O_16] Design and Performance of a Lumped Resistively Loaded Spiral Antenna
M. Zou, J. Pan, M. Quan, L. Zuo (UESTC, China)

[ECC2011_O_17] Speech and Hand Gesture Fusion Interface for the Control of Home Automation System
J.G. Kim, S.Y. Suk, H.K. Kim, K.Y. Yoo, H.Y. Chung (Yeungnam University, South Korea)

[ECC2011_O_18] Resonance Characteristics of Cylindrical Cavities utilizing Curved Resonant Gratings as a Mirror
S. IIJIMA, Y. OHTERA, H. YAMADA (Tohoku University, Japan)

[ECC2011_O_19] Design of 32W Power Amplifier Based on packaged SiC MESFET
Y. Qiu, Y. Guo, Y. Xu, Y. Chen, R. Xu, W.Lin (UESTC, China)

[ECC2011_O_20] The Analysis of Maximum-Distance SW Update for WSNs
H.Y. Jeong, S.H. Cha, Y.M. Kwon, B.C. Ahn (Yeungnam University, South Korea)

ECC 2011 Oral Session V

(11:00~12:45 No. 1 Conference Room, UESTC)

Session Chair: Prof. Sugaya Yoshihiro, Tohoku University, Japan

[ECC2011_O_21] Complexity Reduced Near ML Block Signal Detection for Single-carrier MIMO Transmission
T. YAMAMOTO, F. ADACHI (Tohoku University, Japan)

[ECC2011_O_22] Study about the Parameters of the Holographic Impedance Modulated Surface Antenna
J. Zhang, J. Ouyang, F. Yang (UESTC, China)

[ECC2011_O_23] Separation of Decorative Characters into Skeleton Parts and Decoration Parts
Y. Soma, Y. Sugaya, S. Omachi (Tohoku University, Japan)

[ECC2011_O_24] A Novel MAC Protocol to Improve Performance in Mobile Ad Hoc Networks
M. R. Jonayed, S. W. Kim (Yeungnam University)

[ECC2011_O_25] An Assistant System for Writing in Easy Japanese
M. Zhang, A. Ito (Tohoku University, Japan)

Papers of ECC 2011

Stability criteria for impulsive Cohen-Grossberg neural networks with time-varying delays

Chenhui Zhou
School of Electronic Engineering,
University of Electronic Science and Technology of China,
Chengdu, 610054, P. R. China.
E-mail:zhou1232006@126.com

Hongyu Zhang
School of Electronic Engineering,
University of Electronic Science and Technology of China,
Chengdu, 610054, P. R. China.

Abstract—**This paper is concerned with the problem of exponential stability for a class of impulsive Cohen-Grossberg neural networks with time-varying delays. Based on the Lyapunov method ,and the linear matrix inequality (LMI) approach, we get some sufficient conditions ensuring the global exponential stability of equilibrium point for impulsive Cohen-Grossberg neural networks with time-varying delays. Finally, a numerical examples are given to show the effectiveness of the theoretical results .**

Keywords—neural networks; time–varying delay; impulsive; global exponential stability; linear matrix inequality (LMI).

I. INTRODUCTION

Since Cohen-Grossberg neural networks (CGNNs) were first introduced by Cohen and Grossberg [1], many researchers have done extensive research work on this subject due to their important applications in many areas such as parallel computation, associative memory and optimization problemssee[1-17]. Therefore, it is important and necessary to investigate the stability of this class of CGNNs.

In implementation of neural networks, time delays are unavoidably encountered due to the finite switching speed of neurons and amplifiers. It has been found that, the existence of time delays may lead to instability and oscillation in a neural network.Therefore, the study of stability for delayed Cohen-Grossberg neural networks is of both theoretical and practical importance. In recent years, some results on stability of Cohen- Grossberg neural networks with delays have been obtained (see [2-6] and the references therein).

However, besides delay effect, impulsive effects are also likely to exist in neural networks[16]. In recent years, some results on stability of impulsive neural networks with delays have been obtained[7-17].For instance, in implementation of electronic networks, the state of the networks is subject to instantaneous perturbations and experiences abrupt change at certain instants, which may be caused by switching phenomenon, frequency change or other sudden noise, that is, it does exhibit impulsive effects. Therefore, it is necessary to consider both impulsive effect and delay effects on stability of neural networks.

Motivated by the above discussions, in this paper, we will investigate the global exponential stability of impulsive Cohen -Grossberg neural networks with time-varying delays. Based on the Lyapunov method and the linear matrix inequality (LMI) approach, we get some sufficient conditions ensuring the global exponential stability of equilibrium point for impulsive Cohen-Grossberg neural networks with time -varying delays.

The rest of this paper is organized as follows. Model description and preliminaries are given in Section 2. In Section 3, we give main results and their proof. Examples are given to illustrate our theory in Section 4. Finally, in Section 5 we give the conclusion.

II. PROBLEM FORMULATION AND ASSUMPTIONS

In this correspondence, we consider the following model:

$$\begin{cases} \dfrac{\partial u(t)}{\partial t} = -a_i(u_i(t))\Big[b_i(u_i(t)) - \sum\limits_{j=1}^{n} \alpha_{ij} g_j(u_j(t)) \\ \qquad\qquad - \sum\limits_{j=1}^{n} \beta_{ij} g_j(u(t-\tau_{ij}(t))) - \sum\limits_{j=1}^{n} c_{ij} v_j - J_i \Big] \\ \Delta u_i(t_k) = u_i(t_k^+) - u_i(t_k^-) = I_{ik}(u_i(t_k^-)) \\ u_i(t) = \phi_i(t), t \in [t_0 - \tau, t_0] \end{cases} \quad (1)$$

for i=1,2,...,n; k=1,2,.., $t \geq t_0$, $t \neq t_k$ where n corresponds to the number of units in the neural networks; $u(t) = (u_1(t), u_2(t), ..., u_n(t))^T$ corresponds to the state of the ith unit at time t and position x; $a_i(u_i(t))$ represents an ampliffication function; $b_i(u_i(t))$ is an appropriate behavior function; f_i is the activation functions; α_{ij} and β_{ij} are elements of feedback template, c_{ij} is feed-forward template, respectively; v_i and J_i denote input and bias of the ith neuron; $\tau_{ij}(t)$ corresponds to the transmission delay; t_k is called impulsive moment and satisffies $0 \leq t_0 \leq t_1 \leq t_2 \leq \cdots$, $\lim\limits_{k \to +\infty} t_k = +\infty$; $u_i(t_k^-)$ and $u_i(t_k^+)$

denote the left limit and right limit at t_k . $I_k(u(t_k))=(I_{1k}(u_1(t_k)),I_{2k}(u_2(t_k)),..,I_{nk}(u_n(t_k)))^T$ shows impulsive perturbation of the ith neuron at t_k .

For convenience, we introduce several notations. For $u=(u_1,u_2,...,u_n)^T\in R^n$, $|u|$ denotes the absolute-value vector given by $|u|=(|u_1|,|u_2|,...,|u_n|)^T$, For matrix $A=(a_{ij})_{n\times n}\in R^{n\times n}$, $|A|$ denotes the absolute-value matrix given by diag(b1, b2, . . . , bn) denotes the diagonal matrix with diagonal entries b1, b2, . . . , bn; $\|u\|$ denotes a vector norm defined by $\|u\|=\max_{1\le i\le n}|u_i|$, While $\|A\|$ denotes the absolute-value matrix given by $\|A\|=\max\limits_{1\le i\le n}\left\{\sum\limits_{j=1}^{n}|a_{ij}|\right\}$, For $A,B\in R^{m\times n}$, $A\ge B$ $(A>B)$ means that each pair of corresponding elements of A and B satisfies the inequality "$\ge(>)$". especially, A is called a nonnegative matrix if $A\ge 0$. $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ represent the maximum and minimum eigenvalues of a symmetric matrix A. The notation * always denotes the symmetric block in one symmetric matrix.Matrices,if not explicitly stated, are assumed to have compatible dimensions.

$PC[J,R^n]=\{v(t):J\to R^n \mid v(t)\ is\ continuous\ at\ t\ne t_k,$
$v(t_k^-)=v(t_k)\ and\ v(t_k^+)\ exists\ for\ t,t_k\in J,k\in N\}$

To obtain our results, we give the following assumptions and definitions.

(H1) Each function $a_i(\xi)$ is positive, continuous and bounded, i.e. there exist constants $\underline{a}_i$, $\overline{a}_i$ such that

$$0<\underline{a}_i<a_i(\xi)<\overline{a}_i<\infty,\qquad for\xi\in R,i=1,\dots,n.$$

(H2) $b_i(\xi)$ is continuous, and there exists a positive diagonal matrix $B=diag(b_1,b_2,\dots,b_n)$ such that

$$\frac{b_i(\xi_1)-b_i(\xi_2)}{\xi_1-\xi_2}\ge b_i\ge 0,$$

for any $\xi_1,\xi_2\in R$ and $\xi_1\ne\xi_2,i=1,\dots n.$

(H3) The activation functions $g_i(\xi)$ are continuous, and there exist three positive diagonal matrices $G=diag(G_1,..,G_n)$ such that

$$G_j=\sup_{\xi_1\ne\xi_2}\left|\frac{g_j(\xi_1)-g_j(\xi_2)}{\xi_1-\xi_2}\right|$$

(H4) The transmission delay $\tau(t)$ is time-varying and satises $0<\tau(t)<\tau$ and $0<\dot\tau(t)<\mu$, where τ and μ are some positive constants.

(H5) Let $r_k(u)=u+I_k(u)$ be Lipschitz continuous in R^n , that is, there exist nonnegative diagnose matrix $\Gamma_k=diag(\psi_{1k},\psi_{2k},\dots,\psi_{nk})$ satises

$$|r_k(u)-r_k(v)|\le\Gamma_k|u-v|$$

for all $u,v\in R^n,k\in N$.

Where $r_k(u(t_k,x))=(r_{1k}(u_1(t_k,x)),r_{2k}(u_2(t_k,x)),..,r_{nk}(u_n(t_k,x)))^T$, $I_k(u(t_k))=(I_{1k}(u_1(t_k)),I_{2k}(u_2(t_k)),\dots,I_{nk}(u_n(t_k)))^T$.

(H6) If $(u_1^*,u_2^*,...,u_n^*)^T$ is an equilibrium point of model (1), then the impulsive jumps of model (1) satisfy the following conditions:

$$I_{ik}(u^*)=0,\qquad i=1,...,n,\ k=1,2,...$$

Definition 1. An equilibrium point u^* of system (1) is said to be globally exponentially stable, if there exist constant $\lambda>0$ and $M\ge 1$ such that

$$\|u(t)-u^*\|_2\le M\|\phi(s)-u^*\|_\tau e^{-\lambda t},$$

For all $t\ge t_0$.

Definition 2. A function $u(t):[t_0-\tau,+\infty)\to R^n$ is called a solution of model (1) with the initial condition $u(s)=\phi(s)\in PC([t_0-\tau,t_0],R^n)$, if $u(t)$ is continuous at $t\ne t_k$ and $t\ge t_0$, $u(t_k)=u(t_k^-)$ and $u(t_k^+)$ exist, $u(t)$ satisfies model (1) for $t\ge t_0$, under the initial condition.

Especially, a point $u^*\in R^n$ is called an equilibrium point of model (1), if $u(t)=u^*$ is a solution of model (1).

Lemma 1 (Schur Complement). The following linear matrix inequality (LMI)

$$\begin{pmatrix}Q(x) & S(x)\\ S^T(x) & R(x)\end{pmatrix}>0$$

Where $Q(x)=Q^T(x)$, $R(x)=R^T(x)$ and $S(x)$ depends affinely on x, is equivalent to

$$(\mathrm{I})R(x)>0,Q(x)-S(x)R^{-1}(x)S^T(x)>0,$$
$$(\Pi)Q(x)>0,R(x)-S(x)Q^{-1}(x)S^T(x)>0.$$

Now let

$$y_i(t)=u_i(t)-u_i^*,\qquad i=1,2,...,n$$

It is easy to see that system (1) can be transformed into

$$\begin{cases} \dfrac{\partial y_i(t)}{\partial t} = -a_i(u_i(t))\Big[b_i(y_i(t)) - \sum_{j=1}^{n} \alpha_{ij}\tilde{g}_j(y_j(t)) \\ \qquad\qquad - \sum_{j=1}^{n} \beta_{ij}\tilde{g}_j(y(t-\tau_{ij}(t))) \Big] \quad t \ge t_0, t \ne t_k, k \in N \qquad (2) \\ y_i(t_k^+) = \tilde{r}_{ik}(y_i(t_k^-)), k \in N \\ y_i(t) = \phi_i(t) - u^*, t \in [t_0 - \tau, t_0], i = 1,2,...,n \end{cases}$$

where

$b_i(y_i(t)) = b_i(y_i(t)+u_i^*) - b_i(u_i^*), g_j(y_j(t)) = g_j(y_j(t)+u_j^*) - g_i(u_j^*),$

$r_{ik}(y_j(t_k)) = r_{ik}(y_i(t_k)+u_i^*) - r_{ik}(u_i^*)$

III. STABILITY ANALYSIS

In this section, we will establish a sufficient condition guaranteeing the globally exponential stability of impulsive Cohen-Grossberg neural networks with time-varying delays.

Theorem 1: If there exist diagnal and positive definite matrix P, and Q, such that the following LMIs hold

$$\Sigma = \begin{pmatrix} 2\varepsilon I - 2P\underline{A}B + 2P\bar{A}EG + Q^T G Q & P\bar{A}F \\ * & -(1-\mu)Q \end{pmatrix} < 0 \qquad (3)$$

$$\begin{pmatrix} -\lambda_k P & \Gamma_k \\ * & -P^{-1} \end{pmatrix} < 0 \qquad (4)$$

where $\underline{A} = diag\{\underline{a}_1, \underline{a}_2, ..., \underline{a}_n\}$, $\bar{A} = diag\{\bar{a}_1, \bar{a}_2, ..., \bar{a}_n\}$, $\Gamma_k = diag(\Gamma_{1k}, \Gamma_{2k}, ..., \Gamma_{nk})$, $E = (\alpha_{ij})_{n\times n}$, $F = (\beta_{ij})_{n\times n}$, $1 \le \lambda_k \le e^{\alpha(t_k - t_{k-1})}$, $\alpha > 0$ .Then the system (1) is globally exponentially stable.

Proof: Define a Lyapunov functional by

$$V(t) = e^{2\varepsilon t} y^T(t) P y(t) + \int_{t-\tau(t)}^{t} \tilde{g}^T(y(s)) Q \tilde{g}(y(s)) ds \qquad (5)$$

where $\tilde{g}(y(t)) = (\tilde{g}_1(y_1(t)), \tilde{g}_2(y_2(t)), ..., \tilde{g}_n(y_n(t)))^T$, $y(t) = (y_1(t), y_2(t), ..., y_n(t))^T$.

Taking the Dini derivative of $V(t)$ for $t \ge t_0$ and $t \ne t_k$.

$$\begin{aligned} & D^+V(t) \\ &= 2\varepsilon e^{2\varepsilon t} y^T(t) P y(t) + y^T(t) P \frac{\partial}{\partial t} y(t) + \tilde{g}^T(y(t,x)) Q_1 \tilde{g}(y(t,x)) \\ &\quad - (1-\dot{\tau}(t)) \tilde{g}(y(t-\tau(t))) Q_1 \tilde{g}(y(t-\tau(t))) \\ &\le 2\sum_{i=1}^{n} p_i y_i(t,x) \Big\{ -a_i(u_i(t,x)) \Big[b_i(y_i(t,x)) - \sum_{j=1}^{n} \alpha_{ij} \tilde{f}_j(y_j(t,x)) \\ &\quad - \sum_{j=1}^{n} \beta_{ij} \tilde{f}_j(y(t-\tau_{ij}, x)) \Big] \Big\} + \tilde{g}^T(y(t,x)) Q_1 \tilde{g}(y(t,x)) \\ &\quad - (1-\mu) \tilde{g}(y(t-\tau(t))) Q_1 \tilde{g}(y(t-\mu)) \end{aligned} \qquad (6)$$

Here assumption (H4) is used.

From assumption (H1), we have

$$\begin{aligned} & \sum_{i=1}^{n} p_i y_i(t) a_i(u_i(t)) b_i(y_i(t)) \\ &\ge \sum_{i=1}^{n} p_i y_i(t) \underline{a}_i b_i y_i(t) \\ &= y^T(t) P A B y(t) \end{aligned} \qquad (7)$$

By the same way, we can obtain

$$\begin{aligned} & \sum_{i=1}^{n} p_i y_i(t) a_i(u_i(t)) \sum_{j=1}^{n} \alpha_{ij} g_i(y_i(t)) \\ &\le \sum_{i=1}^{n} p_i y_i(t) \bar{a}_i \sum_{j=1}^{n} \alpha_{ij} g_i(y_i(t)) \\ &\le y^T(t) P \bar{A} E G y(t) \end{aligned} \qquad (8)$$

$$\begin{aligned} & \sum_{i=1}^{n} p_i y_i(t) a_i(u_i(t)) \sum_{j=1}^{n} \beta_{ij} g_i(y_i(t-\tau(t)) \\ &\le \sum_{i=1}^{n} p_i y_i(t) \bar{a}_i \sum_{j=1}^{n} \beta_{ij} g_i(y_i(t-\tau(t)) \\ &\le y^T(t) P \bar{A} F g(y(t-\tau(t)) \end{aligned} \qquad (9)$$

Therefore we obtain

$$\begin{aligned} & D^+V(t,x) \\ &\le e^{2\varepsilon t} y^T(t)(2\varepsilon I - 2P\underline{A}B) y(t) + y^T(t) P \bar{A} E G y(t) \\ &\quad + y^T(t) P \bar{A} F \tilde{g}(y(t-\tau(t))) + [y(t)]^T G^T Q G [y(t)] \\ &\quad + [\tilde{g}(y(t-\tau(t)))]^T (1-\mu) Q [\tilde{g}(y(t-\tau(t)))] \\ &= \begin{bmatrix} y(t) \\ \tilde{g}(y(t-\tau(t))) \end{bmatrix}^T \Sigma \begin{bmatrix} y(t) \\ \tilde{g}(y(t-\tau(t))) \end{bmatrix} < 0 \end{aligned}$$

Then, for $t \in (t_{k-1}, t_k]$, we have

$$V(t) \le V(t_{k-1}^+), \qquad k \in N \qquad (10)$$

On the other hand, when $t = t_k^+$, we have

$$\begin{aligned}
&e^{2\varepsilon t_k^+} y(t_k^+)^T P y(t_k^+) \\
&= e^{2\varepsilon t_k^-} \left[\tilde{r}_k(y(t_k^-))\right]^T P \left[\tilde{r}_k(y(t_k^-))\right] \\
&\le e^{2\varepsilon t_k^-} y(t_k^-)^T \Gamma_k^T P \Gamma_k y(t_k^-) \\
&= e^{2\eta t_k^-} \left[y(t_k^-, x)^T \Gamma_k^T P \Gamma_k y(t_k^-) - \lambda_k y(t_k^-)^T P y(t_k^-) \right. \\
&\quad \left. + \lambda_k y(t_k^-)^T P y(t_k^-) \right] \\
&= e^{2\eta t_k^-} \left\{ y(t_k^-, x)^T \left[\Gamma_k^T P \Gamma_k y(t_k^-) - \lambda_k P \right] y(t_k^-, x) \right. \\
&\quad \left. + \lambda_k y(t_k^-)^T \lambda_k P y(t_k^-) \right\}
\end{aligned} \tag{11}$$

Applying Lemma 1 to (4) yields

$$\Gamma_k^T P \Gamma_k y(t_k^-) - \lambda_k P \le 0 \tag{12}$$

Combining (10) and (11), we have

$$\begin{aligned}
&e^{2\varepsilon t_k^+} y(t_k^+)^T P y(t_k^+) \\
&< \lambda_k y(t_k^-)^T \lambda_k P y(t_k^-) \\
&= \lambda_k y(t_k)^T \lambda_k P y(t_k)
\end{aligned} \tag{13}$$

Therefore we obtain

$$\begin{aligned}
&V(t_k^+) \\
&= e^{2\varepsilon t_k^+} y(t_k^+)^T P y(t_k^+) + g^T(y(t_k^-)) Q g(y(t_k^-)) \\
&< \lambda_k y(t_k^-)^T \lambda_k P y(t_k^-) + \lambda_k g^T(y(t_k^-)) Q g(y(t_k^-)) \\
&\le \lambda_k V(t_k^-) = \lambda_k V(t_k)
\end{aligned} \tag{14}$$

Combining (10) and (14), we have

$$e^{2\varepsilon t_k^+} y(t_k)^T P y(t_k) \le V(t) \le V(t_{k-1}^+) \le \lambda_{k-1} V(t_{k-1}) \le \lambda_0 \lambda_1 \cdots \lambda_{k-1} V(t_0) \tag{15}$$

Since $1 \le \lambda_k \le e^{\alpha(t_k - t_{k-1})}$, therefore for $t \in (t_{k-1}, t_k]$,

$$\begin{aligned}
&e^{2\varepsilon t_k^+} y(t_k)^T P y(t_k) \\
&\le V(t) \le e^{\alpha(t_1 - t_0)} \cdots e^{\alpha(t_{k-1} - t_{k-2})} V(t_0) \le e^{\alpha(t - t_0)} V(t_0)
\end{aligned} \tag{16}$$

From (5), we have

$$\begin{aligned}
V(t) &\ge e^{2\eta t} \lambda_{\min}(P) \left\| y(t) \right\|^2 \\
&= e^{2\eta t} \lambda_{\min}(P) \left\| u(t) - u^* \right\|^2
\end{aligned} \tag{17}$$

$$\begin{aligned}
V(t_0) &\le e^{2\varepsilon t} \lambda_{\max}(P) \left\| \phi(t) - u^* \right\|^2 + \tau \lambda_{\max}(Q) \left\| \phi(t) - u^* \right\|_\tau^2 \\
&\le e^{2\varepsilon t} \left[\lambda_{\max}(P) + \tau \lambda_{\max}(Q) \right] \left\| \phi(t, x) - u^* \right\|_\tau^2
\end{aligned} \tag{18}$$

Combining (17) and (18), therefore for $t \in (t_{k-1}, t_k]$

$$\begin{aligned}
&e^{2\eta t} \lambda_{\min}(P) \left\| u(t) - u^* \right\|^2 \\
&\le e^{\alpha(t - t_0)} e^{2\varepsilon t} \left[\lambda_{\max}(P) + \tau \lambda_{\max}(Q) \right] \left\| u(t, x) - u^* \right\|_\tau^2
\end{aligned}$$

Let $M = \sqrt{\left[\lambda_{\max}(P) + \tau \lambda_{\max}(Q)\right] / \lambda_{\min}(P)}$, and $\varpi = (2\eta - \alpha)/2$,it is easy to see that

$$\left\| u(t, x) - u^* \right\| \le M \left\| \phi(t, x) - u^* \right\|_\tau e^{-\varpi(t - t_0)}, \quad t \ge t_0$$

This complete the proof.

Remark 1: if $u_i(t_k^+) = u_i(t_k^-)$ in Cohen-Grossberg neural networks (1), we have the following result.

Corlllary 1: If there exist diagnal and positive definite matrix P, and Q , such that the following LMIs hold

$$\begin{pmatrix} 2\varepsilon I - 2P\underline{A}B + 2P\bar{A}EG + Q^T G Q & P\bar{A}F \\ * & -(1-\mu)Q \end{pmatrix} < 0$$

where $\underline{A} = diag\{\underline{a}_1, \underline{a}_2, ..., \underline{a}_n\}$, $\bar{A} = diag\{\bar{a}_1, \bar{a}_2, ..., \bar{a}_n\}$, $E = (\alpha_{ij})_{n \times n}$, $F = (\beta_{ij})_{n \times n}$, $1 \le \lambda_k \le e^{\alpha(t_k - t_{k-1})}$, $\alpha > 0$ .Then the system (1) is globally exponentially stable.

IV. ILLUSTRATIVE EXAMPLE

In this section, we will use an example to show the effectiveness of the obtained results.

Example 1: Consider the neural network (1) with two neurons,where $i, j = 1, 2$, $g_j(t) = (|t+1| - |t-1|)/2$, $\varepsilon = 0.25$, $\tau(t) = 0.05\sin^2(t)$, $\bar{a}_i = \underline{a}_i = 1$, $\lambda^k = 1.5$, respectively

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

$$E = \begin{bmatrix} -1.5 & 0.5 \\ 1 & -1.5 \end{bmatrix}$$

$$F = \begin{bmatrix} -1 & 1 \\ 2 & 0.9 \end{bmatrix}$$

$$\Gamma_k = \begin{bmatrix} 0.1317 & 0 \\ 0 & 0.1038 \end{bmatrix}$$

Obviously, $\dot{\tau}(t) = \mu = 0.1$, Applying Theorem 1 of this paper, we have

$$P = \begin{bmatrix} 33.2381 & 0 \\ 0 & 22.6946 \end{bmatrix}$$

$$Q = \begin{bmatrix} 71.1634 & 0 \\ 0 & 67.9872 \end{bmatrix}$$

Therefore, the concerned neural network (1) is globally exponential stable.

V. CONCLUSION

Under some suitable assumptions and using Lyapunov

functional, some LMI-based sufficient conditions have been derived to ensure the global exponential stability of impulsive Cohen–Grossberg neural networks with time-varying delay. Some remarks and a numerical example have been employed to show the concerned neural network is globally exponential stable.

REFERENCES

[1] M. Cohen, S. Grossberg, Absolute stability and global pattern formation and parallel memory storage by competitive neural networks, IEEE Trans. Syst. Man Cybern. 13 (1983) 815–826.

[2] C. C. Hwang, C. J. Cheng, and T. L. Liao, "Globally exponential stability of generalized Cohen–Grossberg neural networks with delays," *Phys. Lett. A*, vol. 319, no. 1/2, pp. 157–166, Dec. 2003.

[3] T. Chen and L. Rong, "Robust global exponential stability of Cohen–Grossberg neural networks with time delays," *IEEE Trans. Neural Netw.*, vol. 15, no. 1, pp. 203–206, Jan. 2004.

[4] J.Cao,L.Liang,Boundedness and stability for Cohen-Grossberg neural network with time-varying delays, J. Math. Anal. Appl. 296 (2004) 665-685.

[5] M. Gao and B. Cui, "Robust exponential stability of interval Cohen–Grossberg neural networks with time-varying delays," *Chaos Solitons Fractals*, vol. 40, no. 4, pp. 1914–1928, May 2009.

[6] C. Huang and J. Cao, "On *p*th moment exponential stability of stochastic Cohen–Grossberg neural networks with time-varying delays," *Neurocomputing*,vol. 73, nos. 4–6, pp. 986–990, Jan. 2010.

[7] J. Li and J. Yan, "Dynamical analysis of Cohen–Grossberg neural networks with time-delays and impulses," *Neurocomputing*, vol. 72, nos.10–12, pp. 2303–2309, Jul. 2009.

[8] M. Tan and Y. Zhang, "New sufficient conditions for global asymptotic stability of Cohen–Grossberg neural networks with time-varying delays," *Nonlinear Anal. RWA*, vol. 10, no. 4, pp. 2139–2145, Aug. 2009.

[9] Z. Wang, H. Zhang, and W. Yu, "Robust stability criteria for interval Cohen–Grossberg neural networks with time varying delay," *Neurocomputing*, vol. 72, nos. 4–6, pp. 1105–1110, Jan. 2009.

[10] J. Cao, J. Wang, Global asymptotic stability of a general class of recurrent neural networks with time-varying delays, IEEE Trans.Circuit Syst.—I 50 (1) (2003) 34–44.

[11] Q. Song and Z. Wang, "Stability analysis of impulsive stochastic Cohen–Grossberg neural networks with mixed time delays," *Phys. A*, vol. 387,no. 3, pp. 3314–3326, May 2008.

[12] H.Zhang and Y. Wang, "Stability analysis of Markovian jumping stochastic Cohen–Grossberg neural networks with mixed time de lays," *IEEE Trans. Neural Netw.*, vol. 19, no. 2, pp. 366–370, Feb.2008.

[13] Z. Guan and G. Chen, "On delayed impulsive Hopfield neural networks," *Neural Netw.*, vol. 12, no. 2, pp. 273–280, Mar. 1999.

[14] Z. Chen, J. Ruan, Global stability analysis of impulsive Cohen–Grossberg neural networks with delay, Phys. Lett. A 345 (2005) 101–111.

[15] Z. Chen, J. Ruan, Global dynamic analysis of general Cohen– Grossberg neural networks with impulse, Chaos, Solitons and Fractals 32 (2007) 1830–1837.

[16] D. Xu, Z. Yang, Impulsive delay differential inequality and stability of neural networks, J. Math. Anal. Appl. 305 (2005) 107–120.

[17] Z. Yang, D. Xu, Impulsive effects on stability of Cohen–Grossberg neural networks with variable delays, Appl. Math. Comput. 177 (2006) 63–78.

[ECC2011_O_02]

Access Point Selection Mechanism to Avoid Rogue Access Point

Nodir Nazarov
Department of Information and Communication Engineering
Yeungnam University,
Republic of Korea
E-mail: nodir@yu.ac.kr

Sirojiddin Djuraev
Department of Information and Communication Engineering
Yeungnam University,
Republic of Korea
E-mail: sirojiddin87@yahoo.com

Seung Yeob Nam
Department of Information and Communication Engineering
Yeungnam University,
Republic of Korea
E-mail: synam@ynu.ac.kr

Abstract-**Rogue Access Point (AP) detection is still a challenging issue since most approaches require high deployment cost and suffer from high false positive and false negative rates. In this paper, we propose AP selection mechanism to prevent wireless users from using rogue APs. Since intelligent attackers can spoof identities of AP, such as BSSID, MAC and IP addresses, a strong authentication mechanism is required. We use signal strength-based authentication technique to support the evasion attempts of intelligent rogue APs. Experiment results show the accuracy, effectiveness and robustness of our rogue AP detection system. The proposed scheme can detect rogue APs accurately with a negligibly small false positives and false negatives.**

Keywords-IEEE 802.11 wireless LAN, Rogue access point detection, signal strength estimation

I. Introduction

Wireless LAN provides high productivity, convenience and mobility so that wireless users can access a wide range of data from different locations such as Internet cafes, airports, shopping malls, universities, offices, homes and campuses with a variety of mobile devices. As 802.11 technologies continue to become more popular, less expensive and easier to install, the threat to corporate or enterprise network security increases. One of the major security concerns in Wireless LAN faced by many enterprise networks today is the Rogue Access Point (RAP).

A rogue access point is an unauthorized Wi-Fi access point attached to corporate network without permission of network administrators. RAP can be deployed casually, for example by innocent users who have weak background on network security, looking for convenience of wireless access, or they can be set for malicious purpose as an insider attack. Whether a rogue AP is placed intentionally or unintentionally, it provides a backdoor for unauthorized users into a private enterprise network. Unauthorized users might get direct access to sensitive data by penetrating conventional security measures. Once an innocent user is connected to a rogue AP, the adversary can manipulate and monitor the incoming and outgoing traffic of the user, and also launch different kinds of attacks. For example, the attacker can easily launch phishing attacks by redirecting the user's web page request to a fake one to steal the user's sensitive information such as bank account number and password [1]. Furthermore, they may interfere with nearby well-planned APs and cause performance problems inside the network [2]. Wired network security solutions, such as firewalls, central authentication, wired IDS are completely blinded to the threat caused by rogue APs. Even the use of encryption and authentication on official wireless LAN may not be enough to prevent Rogue AP [3].

Rogue APs can be classified into two categories:

1) Wired RAP: typical consumer-grade wireless router connected directly into an Ethernet jack. Wireless LAN card installed on desktop or laptop PC that is connected to wired network can also be wired RAP, where WLAN card pretends to be an access point.

2) Wireless RAP: personal computer with two wireless cards, one connected to a real AP and the other configured as an AP to provide Internet access to WLAN stations.

We consider both types of rogue APs in this paper.

We focus on selecting legitimate APs to prevent wireless clients from using rogue APs. The basic idea of our scheme can be described as follows. The list of legitimate APs is managed by a special node, AP registration center (APRC), to distinguish legitimate APs from rogue ones. Thus, when a new wireless client joins the 802.11 network managed by the APRC, it sends a query message to APRC to check if each AP is rogue one, not registered in the list of legitimate APs managed by APRC. Intelligent attackers may try to evade the defense system. They can spoof RAPs' identities, such as BSSID, MAC address, IP address to pretend to be a legitimate AP. Hence, strong authentication mechanism is required. Although attackers may forge the identities mentioned above, they cannot predict signal strength of other APs at a particular wireless node. Thus, we attempt to authenticate APs based on the signal strength of the frames received from each AP.

The remainder of this paper is organized as follows. Section 2 surveys the existing rogue AP detection approaches. Section 3 describes the proposed AP selection mechanism, in details. We show experimental results in Section 4. Finally, Section 5 concludes the paper.

II. Related work

Due to the popularity and flexibility of 802.11 a/b/g/n technologies rogue AP detection systems have become an essential component of wireless network security measures. Most of the current RAP detection approaches can be classified as follows.

A. Wireless Approach

The simplest and most classical approach among rogue AP detection systems is wireless approach. One way to search rogue AP is to scan all radio frequency (RF) channels by carrying portable laptop or handheld device with sniffer software (e.g., AirMagnet [5] and NetStumbler [6]) within the range of campus or enterprise. This approach suffers from different drawbacks, such as scalability, deployment cost, effectiveness and accuracy. Since these wireless analyzer tools require to monitor Wi-Fi signals in entire environment, it can be very time consuming to walk through all of facilities in search of rogues. Also, with 802.11 hardware operating at different frequencies (802.11a - 5 GHz and 802.11b/g – 2.4 GHz), IT personnel must upgrade their detection devices to accommodate multiple frequencies. Furthermore, scans are easy to elude, since a rogue AP can be unplugged when the scan occurs. This method may be workable in a very small area networks.

Another way to search rogue APs is automatic scanning using full time probes or special APs that also works as probes (e.g., ORiNCO AP [7]). However, this technique requires high deployment cost, considering that one must place a large number of sensors for entire enterprise to monitor wireless network. This method is also ineffective, if the organizations do not have wireless access point in the network. Furthermore, it may not be effective if an attacker uses a directional antenna, or reduces the signal strength to cover the small range within his/her office.

B. Wired Approach

The usual approach monitors mixture of wired and wireless traffic at a traffic aggregation point (e.g., a gateway router) and determines whether a host uses wired or wireless connection.

Several research efforts use inter-packet spacing to determine whether the traffic originated on a wireless link or a wired link. In [2], the authors developed their algorithm using sequential hypothesis tests and made prompt decision as TCP ACK pairs are observed. The focus of [9] is on using of temporal traffic characteristics to detect rogue APs at a central location. The study of [11] differentiates wireless traffic by utilizing variations in packet inter-arrival time. They developed Hidden Markov Model by analyzing Denial of Service (DoS) attack and the traffic characteristics of 802.11 based Wireless Local Area Network. In other research [12], a train of packets arriving at the gateway router is used to measure inter-arrival time.

RTT metric is also used to differentiate wireless and wired traffic. In [2], the authors consider RTT as a method of determining if one is using wireless link in the communication path. The work of [13] uses temporal TCP/IP characteristics of SYN, FIN, and ACK local round trip times (LRTT).

All of those methods rely on the difference in the temporal characteristics of wired and wireless traffic. Due to the variations in channel conditions, wireless link capacity varies over time and random delays are introduced. The difference in link speed between wired and wireless links also affects the characteristics significantly. Most of these schemes do not discriminate between wireless traffic from authorized and unauthorized APs, but only detects existence of wireless stations.

C. Agent based Approach

This type of detection technique requires deploying additional network components, such as agents, central administrator and client-server application. The main idea of [15] is to set distributed agent-based intrusion detection and response system for wireless LANs that can detect unauthorized wireless elements. Multi-Agent based methodology is proposed in [16] to detect and eliminate rogue access points. However, it may require a lot of agents for good coverage, which leads to a high deployment cost.

D. Client side Approach

The authors of [14] discovered new rogue AP detection scheme that can be implemented purely by end users. Their RAP detection protocol uses timing information based on the RTT. The intuition is to let the user probe a server in the local network and measure the RTT from the response. If the mean of RTT is statistically larger than a certain threshold, associated AP is defined as a rogue AP. However, their mechanism suffers from various drawbacks. First, they considered only wireless rogue APs in a public places, like airports. It cannot do anything about other types of RAPs, especially wired RAPs. Then, since they use DNS lookup as a probe message, an attacker can easily break the system by constructing its own DNS cache. In this case, clever RAP responses to clients' DNS lookup requests by itself in order to reduce RTT value. Third, the accuracy of this algorithm may be decreased down to 60% in heavy traffic conditions. Moreover, it takes a rather long time (17.2 sec when number of probe packets is 100 and 50.7 sec when it is 300) to detect rogue APs in saturated network traffic conditions. They improve the efficiency of the algorithm by increasing the number of samples from 100 to 300. However, too many probe packets may create huge traffic that might affect the throughput of existing data traffic.

Our proposed scheme helps wireless clients to select

authorized AP while avoiding unauthorized ones. It can detect all types of rogue APs with a high accuracy.

III. Proposed AP selection mechanism

Our AP selection method uses voting request and reply messages exchanged between a wireless user and APRC through several APs. We employ APRC to check legitimacy of an AP based on the pre-defined list of authorized APs. The wireless node relies on the voting to check the authorization status of APs detected within the range. The wireless node repeats this process multiple times by connecting to different APs and makes a decision based on the collected replies. If the number of answers telling a certain AP is legitimate is larger than the number of opposing answers, we conclude that this AP is legitimate. Finally, the wireless node selects AP with the strongest signal among the authorized candidate APs.

A. APRC (AP Registration Center)

The proposed scheme uses administrator-centric approach, where the management node APRC can tell a wireless node if a certain AP is authorized or not. When wireless clients scan the wireless network and detect multiple APs within their communication range, they ask APRC whether an associated AP is legitimate or not. Since APRC manages the list of authorized APs, APRC can easily discriminate rogue APs from legitimate ones based on the MAC addresses of APs. Fig. 1 illustrates wireless network architecture with various types of APs.

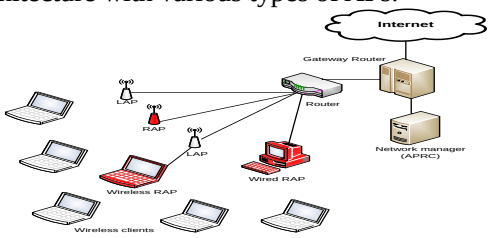


Fig. 1 Wireless neztwork architecture including a legitimate AP (LAP) authorized by APRC and two types of RAPs, i.e. wireless and wired RAP.

The IP address of APRC can be advertized through beacon frames to wireless nodes. It is possible to add this information to the end of the beacon frame by modifying the driver of AP. Since we assume every wireless user know the IP address of APRC through the beacon messages, each wireless node can ask APRC if a specific AP is legitimate or not. When APRC receives query message from a wireless node, it replies with simple YES/NO answer based on the list of legitimate APs.

B. Extension of beacon frames

The IEEE 802.11 standard allows AP vendors to add 253 bytes of vendor-specific Beacon Information Elements (BIEs) in its beacon. This information is vendor specific, which is commonly used by each vendor to put their own proprietary information within the Beacon frames as shown in Fig. 2. We use this feature to define a special BIE for broadcasting IP address of APRC. IP address information requires only 32 bit space. The main drawback of this approach is that it requires driver modification for client devices. We will investigate the issue of implementation, more specifically driver modification, in future study.

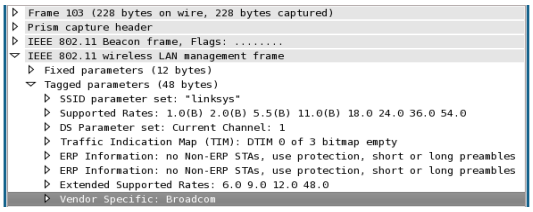


Fig. 2 Vendor-specific field in IEEE 802.11 beacon frame.

C. Attacker model

We consider possible attack scenario in Fig. 3.

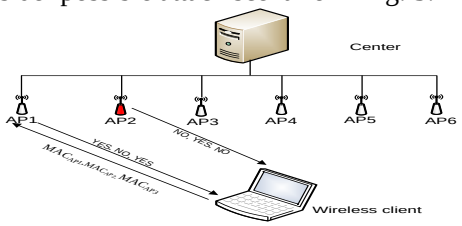


Fig. 3 Spoofing attack: AP1, AP3, AP4, AP5 and AP6 are authorized APs; AP2 is attacker's rogue AP that spoofs BSSID, MAC and IP addresses of voting reply messages.

In this case, AP1, AP3, AP4, AP5 and AP6 are authorized APs that are registered in the list of legitimate APs in APRC. Only AP2 is a malicious rogue AP. Attacker's rogue AP can be wired or wireless AP depending on the attack environment. If there is available Ethernet jack to connect to the wired network, the attacker can establish wired RAP, otherwise he can set a wireless RAP. Since a wireless node trusts the answer on voting reply message coming from APRC through an associated AP, the attacker may attempt to modify voting reply message to advertise himself as legitimate AP and other APs as rogue ones when a wireless node is connected to the attacker's AP. Moreover, the attacker may try to interfere communications between the wireless node and other APs. The adversary node can listen to the communication between other nodes and APs and generate fake response messages that advertise legitimacy of itself and the illegitimacy of other APs. Hence, strong authentication mechanism is required. Even if an attacker can spoof identities of AP, such as BSSID, MAC and IP addresses, he cannot guess the signal strength of frames from other APs at a certain location. Therefore, the wireless node can identify from which AP the voting reply messages arrive based on the signal strength. We assume that there exists more than one AP within the range of a wireless node. We further assume that the number of APs that are not under control of attacker is larger than the number of attacker's APs.

D. AP selection algorithm

In this subsection, we explain the proposed access point selection mechanism to avoid rogue access points in more detail. If there is no rogue access point, then we had better select the access point whose signal strength is highest at the selected wireless node, since higher throughput can be provided by that access point. However, the AP corresponding to the highest signal strength might be a rogue one. Thus, we attempt to select an access point whose signal strength is high among the non-rogue APs within the range. The proposed mechanism consists of two phases. In the first phase, three candidate access points are selected among the APs within the range of the wireless node based on the signal strength. In the second phase, we detect rogue APs among the three candidate APs, and select the final one among the non-rogue APs based on the signal strength.

In the first phase, the wireless node measures the signal strength of beacon frames from every AP within the communication range. T denotes the duration of the measurement period (in seconds). Since beacon frames are usually sent at the interval of 100 msec, $10T$ samples can be collected for each AP during an interval of duration T. the default value of T is 2 seconds in this paper. After collecting enough number of samples for each AP, we can calculate the mean and the standard deviation of signal strengths for each AP. μ_i and σ_i denote the mean and the standard deviation of the signal strengths for AP_i $(1 \le i \le n)$, where n is the number of APs within the range of the selected wireless node. After obtaining the list of (μ_i, σ_i) for every AP within the range, we select three candidate access points in the following way. We select the access point whose average signal strength is highest at the wireless node as the first candidate AP, AP_a.

Let us consider the following two set $S_b^{'}$ and S_b:

$$S_b^{'} = \{i \mid \mu_i + \sigma_i < \mu_a - \sigma_a, 1 \le i \le n\},$$

$$S_b = \{AP_i \mid |\mu_i| \ge |\mu_j|, \forall j \in S_b^{'}\}.$$

Then, S_b contains the access point whose signal strength is highest among the access points whose signal strength ranges are not overlapping with that of AP_a. S_b usually contains only one AP as an element, and that AP is selected as the second candidate AP, AP_b. If S_b has more than one element, then one of them is randomly selected as AP_b. The third candidate AP, AP_c is also selected in a similar way, i.e. randomly, from the following set S_c:

$$S_c = \{AP_i \mid |\mu_i| \ge |\mu_j|, \ \forall j \in S_c^{'}\},$$

where $S_c^{'} = \{i \mid \mu_i + \sigma_i < \mu_b - \sigma_b, 1 \le i \le n\}$.

We can easily know that the signal strength ranges of AP_a, AP_b, and AP_c do not overlap, and this property is required to authenticate access points in the second phase.

In the second phase, the wireless node sends query messages to APRC to know the legitimacy of three candidate APs. However, if that query is delivered through a rogue access point, then the rogue AP might attempt to modify the response message to change it favorable to the rogue AP. Thus, we send three query messages to APRC through three different APs, i.e. candidate APs. Three query messages are required to check the authorization status of each AP. Since we need to investigate the legitimacy of three candidate APs, nine messages are required in total. However, it is possible to merge three query messages delivered through the same access point. Thus, only three messages are enough to query the authorization status of three candidate APs through three different APs.

If the query message containing the three MAC addresses of candidate APs, (MAC_a, MAC_b, MAC_c) is sent to APRC, then APRC investigates the existence of each MAC address in the list of MAC addresses of authorized APs. If the MAC address is found in the table, then the AP corresponding to that MAC address is verified to be a legitimate one.

Even though the rogue AP attempts to modify the response message sent from APRC, if the number of normal APs is larger than that of rogue APs among the candidate APs, then a correct decision can be made on the legitimacy of each candidate AP based on the majority of votes, i.e. the answers in the reply messages. However, in the worst case the rogue AP might generate fake response messages pretending to come from other candidate APs to increase the number of votes supporting the rogue AP itself. For example, if the rogue AP can generate two more fake response messages and they are counted as valid votes by the wireless node, then the rogue AP can win the vote since there will be only two replies or votes from the normal APs and three fake votes from the single rogue AP.

We resolve this problem by classifying the voting replies into three categories corresponding to three APs based on the signal strength of each frame, and merging the answers belonging to the same category into one. Table I show the mean and the standard deviation of the signal strength for the beacon frames arriving from three candidate APs.

TABLE I RANGE OF SIGNAL STRENGTH OF FRAMES ARRIVING FROM THREE CANDIDATE APS

AP number	Mean of signal strength	Standard deviation
a	μ_a	σ_a
b	μ_b	σ_b
c	μ_c	σ_c

If r_i is the signal strength of the reply message from APRC, then that message is considered to come from the AP whose mean signal strength is closest to r_i. After all the reply messages are classified according to this rule, we merge the answers belonging to the same category in the following way. If all the reply messages contain the same answer, then we merge all those answers into one for the

selected category since normal APs are not likely to deliver multiple reply messages. If the answers of those replies are not consistent, then we discard all the answers for that category.

Even though a rogue AP generates multiple fake reply messages with spoofed source MAC addresses, all those messages will be classified into the same category corresponding to the rogue AP, and they will be merged into one according to the above rule. Thus, signal strength-based message classification or AP authentication can effectively mitigate fake reply message problem of a rogue AP.

The final access point selection rule can be summarized as follows. Let A denote the set of three candidate access points, and let R denote the set of rogue APs among the access points in A. Then, we select the access point as an element of the following set S:

$$S = \{i \mid i \in A \cap R^c, |\mu_i| \geq |\mu_j| \; \forall j \in A \cap R^c\}$$

In other words, we select the AP whose signal strength is strongest among the non-rogue candidate access points.

IV. Experimental results

To evaluate the performance of proposed algorithm, we built experimental test bed as shown in Fig. 4.

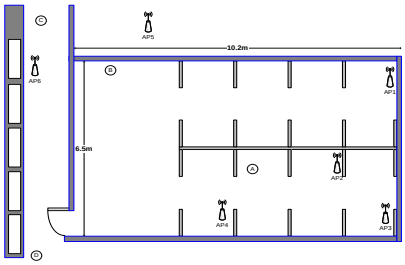


Fig. 4 Location of APs: Wireless users are connected to network from locations A, B, C and D.

All APs are separated from each other by at least a few meters. We perform experiment at different locations in order to show that it is usually possible to find three APs whose signal strengths are clearly different from each other. Table II illustrates the distance between wireless nodes and APs in our test bed.

LOCATION OF WIRELESS STATIONS

Wireless node	Distance from AP (m)					
	AP1	*AP2*	*AP3*	*AP4*	*AP5*	*AP6*
A	5.7	2.7	4.7	2.4	6.15	9
B	8.9	8.1	10.3	6.34	2.4	4
C	12.9	12.1	14.4	10	4.7	1.7
D	15	11.65	13.2	7.7	10.15	7.1

We measure the signal strength of beacon frames received by a wireless machine in several locations. Fig. 5 shows the change of signal strength of the frames received by a wireless node at position A. Fig. 6 shows the range of signal strength of beacon frames arriving from different APs, i.e. $(\mu - \sigma; \mu + \sigma)$, where μ and σ are the average and the standard deviation of the measured signal strength.

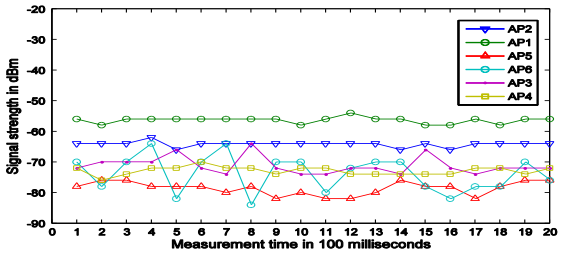


Fig. 5 Signal strength of beacon frames measured at position A.

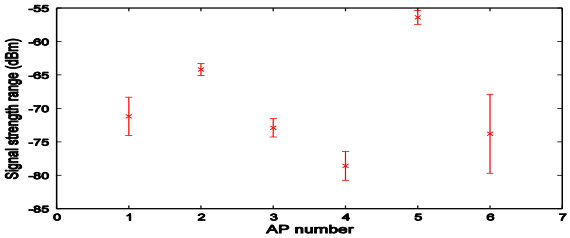


Fig. 6 The range of beacon frames signal strength.

We use UDP packet to deliver voting request and response messages. MAC addresses of all beacon frames are appended to the payload of UDP packet for the voting request message. The APRC replies with UDP packet that has the sequence of YES/NO answers representing the legitimacy of each AP.

Access points AP1, AP2 and AP5 are selected for the voting process since the ranges of their signal strengths are not overlapped. After performing the voting process by associating with those three candidate APs one by one, we obtain records shown in Table III.

TABLE II VOTING REPLY MESSAGES BEFORE MERGING

Received frame number	Voting response		Legitimacy checked by APRC		
	MAC address	Signal strength (dB)	*AP1*	*AP2*	*AP5*
1	MAC_{AP1}	-72	YES	NO	YES
2	MAC_{AP1}	-62	NO	YES	NO
3	MAC_{AP5}	-56	YES	NO	YES
4	MAC_{AP2}	-64	NO	YES	NO

As shown in that table, the second entry belongs to the spoofed voting response, i.e., MAC address is same with AP1 but the signal strength is close to the range of AP2. The result of reply message classification and merging is shown in Table IV.

TABLE III VOTING ANSWERS AFTER MERGING OF VOTING REPLY MESSAGES

Merged answer number	Voting response		Legitimacy checked by APRC		
	MAC address	Signal strength (dB)	*AP1*	*AP2*	*AP5*
1	MAC_{AP1}	-72	YES	NO	YES
2	MAC_{AP2}	-62	NO	YES	NO
3	MAC_{AP5}	-56	YES	NO	YES

Finally, we run access point selection algorithm. Since the number of YES answers is larger than NO answers for AP1, AP1 is considered as an authorized AP. With the same reasoning AP5 is detected as an authorized AP and AP2 is categorized as a rogue AP. Finally, AP5 is selected because the signal strength of frames from AP5 is higher than for AP1.

We performed our detection algorithm 20 times in various locations. We obtained 100% accuracy in all experiment, i.e., there is no false negatives and false positives.

We find that it takes 10 to 130 msec to perform single voting procedure, i.e. from voting request packet transmission time to voting reply packet reception time. We also find that it usually takes less than 16 sec to perform the whole process in the pre-described experiment environment, and thus, the proposed scheme requires a similar or less amount of time compared with the RTT-based approach of [14]. Furthermore, since only six messages, three query and three response messages, are exchanged during the whole procedure, the control message overhead is much less than for the RTT-based approach of [14].

V. Conclusion

In this paper, we propose a new access point selection mechanism to avoid rogue APs. The proposed scheme consists of two stages. In the first stage, we select three candidate access points based on the signal strength measured at the wireless node. Since we select three access points whose signal strengths are clearly different from each other, the signal strength can be used to authenticate APs. In the second stage, we detect rogue access points by consulting a management center, APRC, about the authorization status of three candidate APs. In order to cope with the fake message attack by rogue APs, we send multiple query messages through the three selected candidate APs, and authenticate the sender of the reply message based on the signal strength. The experiment results show that the proposed scheme can select an authorized AP with strong signal strength while avoiding rogue APs with a high probability.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technologies (No. 2011-0013083).

References

Hao Han; Chiu C. Tan; Qui LI; Sanglu Lu. "A Timing-Based Scheme for Rogue AP Detection" *IEEE Transactions on Parallel and Distributed Systems*

Wei Wei; Kyoungwon Suh; Bing Wang; Yu Gu; Kurose, J.; Towsley, D.; Jaiswal, S. "Passive Online Detection of 802.11 Traffic Using Sequential Hypothesis Testing with TCP ACK-Pairs" *Mobile Computing, IEEE Transactions*

http://www.rogueap.com/

[4] Mohan K Chirumamilla, Byrav Ramamurthy "Agent Based Intrusion Detection and Response System for Wireless LANs" 0-7803-7802-4/03/$17.00 © 2003 IEEE.

[5] www.airmagnet.com

[6] www.netstumbler.com

[7] White paper. "Rogue Access Point Detection: Automatically detect and Manage Wireless Threats to Your Network" *Proximcorportion*

[8] Watkins, L. ; Beyah, R. ; Corbett, C. "A Passive Approach to Rogue Access Point Detection" *Global Telecommunications Conference*, 2007

[9] Beyah, R., Kangude, S., Yu, G., Strickland, B., and Copeland, J. "Rogue Access Point Detection Using Temporal Traffic Characteristics." *In Proceedings of IEEE GLOBECOM 2004*

[10] Wei, W.; Jaiswal, S.; Kurose, J.; Towsley, D. "Identifying 802.11 Traffic from Passive Measurements Using Iterative Bayesian Inference." *In the Proceedings of INFOCOM, 2006*

[11] Shivaraj, G.; Min Song; Shetty, S. "A hidden Markov Model based approach to detect Rogue Access Points" *Military Communications Conference, 2008*

[12] Shetty, Sachin; Song, Min; Ma, Liran. "Rogue Access Point Detection by Analyzing Network Traffic Characteristics" *Military Communications Conference 2007*

[13] Guangzhi Qu; Nefcy, M.M. "RAPiD: An Indirect Rogue Access Points Detection System" *Performance Computing and Communications Conference (IPCCC), 2010 IEEE 29th International*

[14] Hao Han; Bo Sheng ; Tan, C.C. ; Qun Li ; Sanglu Lu. "A Timing-Based Scheme for Rogue AP Detection" *IEEE Transactions on Parallel and Distributed Systems*

[15] Chirumamilla, M.K.; Ramamurthy, B. "Agent based Intrusion Detection and Response System for Wireless LANs" *Communictaions, 2003. ICC '03. IEEE International conference*

[16] Sriram, V.S.S.; Sahoo, G.; Agrawal, K.K. "Detecting and Eliminating Rogue Access Points in IEEE-802.11 WLAN - a multi-agent sourcing Methodology" Advance Computing Conference (IACC), 2010 IEEE 2nd International

Sparse Channel Estimation for Multipath Fading Dual-hop Single Relay Cooperation Networks

Guan Gui[1,2], Wei Peng[2], Qun Wan[1], and Fumiyuki Adachi[2]
[1] Department of Electronic Engineering, University of Electronic Science and Technology of China
Chengdu, 611731, China
[2] Department of Electrical and Communication Engineering, Tohoku University, Sendai, 980-8579, Japan
{gui, peng}@mobile.ecei.tohoku.ac.jp; wanqun@uestc.edu.cn; adachi@ecei.tohoku.ac.jp

Abstract—**Coherent detection and demodulation at the receiver requires channel state information. In this paper, we investigate channel estimation problem in sparse multipath dual-hop single relay (DHSR) cooperation networks. Conventional linear channel estimation methods, e.g., least square (LS), based on the rich multipath assumption in cooperation convoluted channels, are robust and simple while at the cost of large amount training resource. However, recent physical channel measurements, have shown that the multipath channel, in point-to-point (P2P) wireless system, exhibits great sparsity in some high-dimensional space, e.g., time-delay and Doppler-spread domain. Based on the mathematical channel model of P2P and sparseness measure function, we show that the cooperation convoluted channels also exhibit sparsity using Monte Carlo runs and propose several sparse channel estimation methods to exploit the inherent sparse structure in channel delay spread domain multipath DHSR cooperation networks using amplify-and-forward protocol. Simulation results confirm the superiority of the proposed methods to LS-based linear estimation method.**

Keywords-dual-hop single relay (DHSR), compressive sensing, cooperation convoluted channel, cooperation networks, sparse multipath fading.

I. INTRODUCTION

Relay-based cooperation communication networks [1-4] have been intensively studied in the last decades due to its capability of enhancing the transmission capacity and providing the spatial diversity for single-antenna receivers by employing the relay nodes as virtual antennas [4]. It is well known that utilizing multiple input multiple output (MIMO) transmission can boost the channel capacity [5-6]. In addition, MIMO diversity techniques can mitigate fading and hence enhance the quality of service (QoS) [7-8]. However, it is very hard to mount multiple antennas onto a small handheld terminal. To resolve the contradiction between then, one could choose relay-based cooperation networks can be introduced which have been investigated in last decade years [1, 4]. The main reason is that relay nodes in the network can be exploited as diversity antennas, relay nodes can either be provided by a network operator or be obtained from cooperating terminals of other users.

In the relay network, data transmission is usually divided into two phases. During phase I, the source broadcasts its own information to relay. During Phase II, the relay forwards its received signal to the destination. Usually, there has two kinds of protocols in cooperation networks, one is purely amplify the received signal at relay and forward it to desination, which is termed as amplify-and-forward (AF); and the second is to demodulate the received signal and modulated again and retransmit to destination, which is often termed as decode-and-forward (DF). Due to coherent demodulation, accurate channel state information (CSI) is required at the destination (for AF) or at both relay and destination (for DF). For DF cooperation networks, the channel estimation methods developed for P2P communication systems can be applied. However, extra computation of channel estimation will increase the computational burden at relay and broadcasting the estimated channel information will result in further interference at destination. Based on the different protocols, cooperative networks can be classified into three categories:

(1) $\mathbb{S}$ communicates with $\mathbb{R}$ and $\mathbb{D}$ during the first time slot. In the second time slot, both S and $\mathbb{R}$ communicate with $\mathbb{D}$.

(2) $\mathbb{S}$ communicates with $\mathbb{R}$ and $\mathbb{D}$ during the first time slot. In the second time slot, only the $\mathbb{R}$ communicates with $\mathbb{D}$.

(3) $\mathbb{S}$ sends signal to $\mathbb{R}$, assuming that $\mathbb{D}$ cannot receive the direct signal from $\mathbb{S}$ during the first time slot. At the second time slot, the amplified signal is broadcast to destination $\mathbb{D}$.

In this paper, we focus our attention on category (3), which is shown in Fig. 1. Linear channel estimation for the relay-based AF cooperation networks has been proposed 4 assuming a rich multipath channel. Hence, low spectral efficiency is unavoidable due to that large space is allocated to transmit training sequence and relatively small space is left to carry user data. If an insufficient length training sequence is used, sufficiently accurate channel estimate cannot be obtained. Recent channel measurements have reported that the wireless channels often exhibit inherent sparse or cluster-sparse structure in delay-spread domain. In order to take advantages of channel sparsity, we propose a novel compressive channel estimation scheme. Simulation results will confirm the effectiveness of the proposed methods.

Section II introduces the AF cooperation channel model and the sparseness measure of convoluted channel. Section III discusses compressive channel estimation of convoluted channel and presents various CCE methods. In section IV, we give various simulation results and discuss the performance of the estimators. Concluding remarks are presented in Section V.

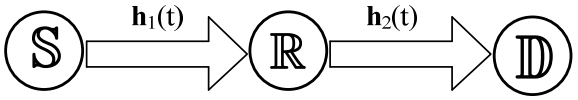


Fig. 1. A typical sparse multipath cooperation network.

II. SYSTEM MODEL

Due to the limited transmit power and multipath fading, we assume that there is no direct propagation path between $\mathbb{S}$ and $\mathbb{D}$ as shown in Fig. 1. And frequency-selective multipath channels will generate multiple delayed and attenuated copies of the transmitted waveform. Source and relay are assumed to have average power constraints in P_S and P_R respectively. The L_1 length discrete multipath channel vector $\mathbf{h}_1$ between $\mathbb{S}$ and $\mathbb{R}$ is represented by

$$\mathbf{h}_1 = \sum_{l=0}^{L_1-1} h_{1,l}(\tau)\delta(\tau-\tau_{1,l}), \tag{1}$$

Where $h_{1,l}$ and $\tau_{1,l}$ denote the complex-valued path gain with $E[\sum_{l=0}^{L_1-1}|h_{1,l}|^2]=1$ ($E[.]$ denotes the expectation operation) and the symbol-spaced time delay of the l^{th} path, respectively. According to the channel representation Eq. (1), a N−dimensional (complex) signal $\mathbf{x}$ transmitted in equivalent complex cooperation channel leads to a received signal at the relay given by [12,14]

$$\mathbf{y}_1 = \mathbf{H}_1\mathbf{x}+\mathbf{n}_1, \tag{2}$$

where $\mathbf{H}_1$ is an $N\times N$ circulant channel matrix with $[\mathbf{h}_1^T\ \mathbf{0}_{1\times(N-L)}]^T$ as its first column, $\mathbf{n}_1$ represents the complex additive Gaussian white noise (AWGN) with zero-mean and covariance matrix $E[\mathbf{n}_1\mathbf{n}_1^H]=\sigma_n^2\mathbf{I}_N$, ($\mathbf{I}_N$ denotes a $N\times N$ identity matrix). Due to the same channel property as Eq.(1), channel vector $\mathbf{h}_2$ between $\mathbb{R}$ and $\mathbb{D}$ is given by

$$\mathbf{h}_2 = \sum_{l=0}^{L_2-1} h_{2,l}(\tau)\delta(\tau-\tau_{2,l}), \tag{3}$$

where $h_{2,l}$ and $\tau_{2,l}$ denote the complex-valued path gain with $E[\sum_{l=0}^{L_2-1}|h_{2,l}|^2]=1$ and the symbol-spaced time delay of the l^{th} path, respectively. Hence, the received signal vector at the destination $\mathbb{D}$, which is from relay $\mathbb{R}$, can be written as

$$\mathbf{y}_2 = \beta\mathbf{H}_2\mathbf{y}_1+\mathbf{n}_2 = \beta\mathbf{H}_2\mathbf{H}_1\mathbf{x}_1+\mathbf{n}, \tag{4}$$

where $\mathbf{n}=\beta\mathbf{H}\mathbf{n}_1+\mathbf{n}_2$ is the composite noise with zero mean and covariance matrix $E\{\mathbf{n}\mathbf{n}^H\}=(\beta^2|\mathbf{H}_2|^2+\mathbf{I}_N)\sigma_n^2$, where $\mathbf{I}_N$ is an $N\times N$ identity matrix, and the amplified positive coefficient β is given by

$$\beta = \sqrt{\frac{P_R}{\sigma_{\mathbf{h}_1}^2 P_S+\sigma_n^2}}, \tag{5}$$

According to the matrix theory [15], circulant channel matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ can be decomposited as $\mathbf{H}_1=\mathbf{W}^H\mathbf{\Lambda}_1\mathbf{W}$ and $\mathbf{H}_2=\mathbf{W}^H\mathbf{\Lambda}_2\mathbf{W}$, respectively, where $\mathbf{W}$ is the (unitary) discrete Fourier transform matrix (DFT) with $\mathbf{W}^{mn}=1/\sqrt{N}\,e^{-j2\pi mn/N}$, $m,n=0,1,..,N-1$. Hence, system model (4) can be rewritten as

$$\mathbf{y}_2 = \mathbf{W}^H\beta\mathbf{\Lambda}_2\mathbf{\Lambda}_1\mathbf{W}\mathbf{x}+\mathbf{n}. \tag{6}$$

If the (6) is left-multiplied by $\mathbf{W}$, is can yield

$$\mathbf{y} = \mathbf{X}\mathbf{h}+\hat{\mathbf{n}}, \tag{7}$$

where $\mathbf{h}\triangleq\beta(\mathbf{h}_2*\mathbf{h}_1)$ is the cooperation convoluted channel vector, $\mathbf{X}=diag(\mathbf{W}\mathbf{x})\mathbf{F}$ denotes the training signal matrix, $\mathbf{F}$ is a matrix taking the first $(2L-1)$ columns of $\sqrt{N}\mathbf{W}$ and $\hat{\mathbf{n}}=\mathbf{\Lambda}_2\mathbf{W}\mathbf{n}_1+\mathbf{W}\mathbf{n}_2$ is a realization of a complex Gaussian random vector with zero mean and covariance matrix of $E\{\mathbf{n}\mathbf{n}^H\}=(\beta^2|\mathbf{\Lambda}|^2+\mathbf{I}_N)\sigma_n^2$. The maximum likelihood (ML) algorithm based coherent detection is then obtained from

$$\begin{aligned}\tilde{\mathbf{x}} &= \arg\max_{\mathbf{x}} P(\mathbf{y}|\mathbf{x})\\ &= \arg\max_{\mathbf{x}} \frac{1}{\pi\sigma_n^2(\beta^2|\mathbf{h}_1|^2+1)}\times\exp\left\{-\frac{|\mathbf{y}-\mathbf{X}\mathbf{h}|^2}{\sigma_n^2(\beta^2|\mathbf{h}_1|^2+1)}\right\}\\ &= \arg\min_{\mathbf{x}} |\mathbf{y}-\mathbf{X}\mathbf{h}|^2 .\end{aligned} \tag{8}$$

For coherent detection in AF cooperation network, destination $\mathbb{D}$ performs the ML detection using cooperation channel estimator We will calculate the sparseness of cooperation channel h at first. A measure function [16] based on the relationship between the ℓ_1-norm and the ℓ_2-norm is utilized to calculate the sparseness of channel vectors. Consider a sparse channel vector $\mathbf{h}$, its sparseness is calculated by

$$\text{Sparseness}(\mathbf{h}) = \frac{\sqrt{L} - \|\mathbf{h}\|_1 / \|\mathbf{h}\|_2}{\sqrt{L} - 1}, \tag{9}$$

where L is the length of $\mathbf{h}$. This function equals to unity if and only if $\mathbf{h}$ contains only a single non-zero component. Fig. 3 shows the sparseness of channel vectors $\mathbf{h}_i, i = 1, 2$ and its cooperation channel vector $\mathbf{h}$. To compare the sparseness measure of different channel vectors in Fig. 2, we assume that the length of $\mathbf{h}_i, i = 1, 2$, are $L_1 = 32$ and $L_2 = 36$, respectively. Hence, the length of cooperation channel $\mathbf{h}$ is $L = L_1 + L_2 - 1 = 67$. From Fig.3, it can be found that the sparseness of cooperation channel $\mathbf{h}$ is even sparser than $\mathbf{h}_i$. Hence, compressed channel sensing can be utilized to get the estimation of $\mathbf{h}$.

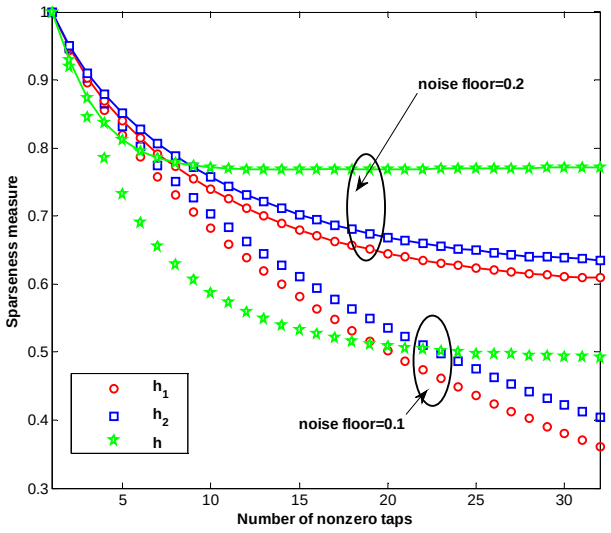


Fig. 2. Sparseness measure of different channel vectors over the noise floor.

III. COMPRESSIVE CHANNEL E ESTIMATION

From (7), the optimal sparse channel estimator $\tilde{\mathbf{h}}_{opt}$ t is given by [18]

$$\tilde{\mathbf{h}}_{opt} = \arg\min_{\mathbf{h}} \left\{ \frac{1}{2} \|\mathbf{y} - \mathbf{X}\mathbf{h}\|_2^2 + \lambda \|\mathbf{h}\|_0 \right\}, \tag{10}$$

where λ is a regularized parameter which tradeoffs the estimation error and sparsity of $\mathbf{h}$, $\|\mathbf{h}\|_2^2$ denotes the ℓ_2-norm which is given by $\|\mathbf{h}\|_2^2 = \sum |h_i|^2$ and $\|\mathbf{h}\|_0$ denotes ℓ_0-norm which counts the number of nonzero coefficients. Unfortunately, (10) is nonconvex optimization problem and is NP-hard [18]. In other words, optimal sparse estimators are unlikely to be calculated efficiently. However, numerous practical suboptimal algorithms exist for the cooperation convoluted channel h if the training measurement matrix satisfies the restricted Isometry property (RIP) [19]. Usually, these sub-optimal algorithms can be classified in two kinds. The first kind is greedy algorithms such as orthogonal matching pursuit (OMP) [10] and compressive sampling matching pursuit (CoSaMP) [11], which select each dominant coefficient in channel by iteration. The second kind is convex relation methods such as Lasso [9]. Here we also consider the LS channel estimator (known channel position of $\mathbf{h}$) for comparison. The LS estimator (known position) is given by

$$\tilde{\mathbf{h}} = \begin{cases} \mathbf{X}_T^{\dagger}\mathbf{y}, & T \subseteq \text{supp}(\mathbf{h}) \\ 0, & \textit{others} \end{cases} \tag{11}$$

where supp($\mathbf{h}$) denotes the nonzero taps supporting the channel vector $\mathbf{h}$, $\mathbf{X}_T$ is the submatrix constructed from the columns of $\mathbf{X}$, and T denotes the selected subcolumns corresponding to the nonzero index set of the convoluted channel vector $\mathbf{h}$. The MSE of LS estimator $\tilde{\mathbf{h}}$ is given by [13]

$$MSE(\tilde{\mathbf{h}}) = \sigma_n^2 \,\text{Tr}\left\{ \left(\mathbf{X}_T^{?} \mathbf{X}_T \right)^{-1} \right\}. \tag{12}$$

By utilizing CS recovery algorithms for compressive channel estimation, we propose CCS-Lasso, CCS-OMP and CCS-CoSaMP. The three methods for cooperation convoluted channel estimate are described as follows.

A. CCS-Lassso estimator

Given the received signal y, the unitary DFT matrix $\mathbf{W}$ and $\mathbf{F}$, training signal $\mathbf{x}$ (training signal matrix $\mathbf{X} = diag(\mathbf{W}\mathbf{x})\mathbf{F}$), the regularized parameter is set as $\lambda = \sigma\sqrt{2\log N}$. The CCS-Lasso channel estimator $\tilde{\mathbf{h}}_{Lasso}$ is given by

$$\tilde{\mathbf{h}}_{Lasso} = \arg\min_{\mathbf{h}} \left\{ \frac{1}{2} \|\mathbf{y} - \mathbf{X}\mathbf{h}\|_2^2 + \lambda \|\mathbf{h}\|_1 \right\}. \tag{13}$$

B. CCS-OMP estimator

Given the received signal $\mathbf{y}$, $\mathbf{W}$ and $\mathbf{F}$, and $\mathbf{x}$, CCSOMP estimator performs as follows: **Initialize**. Set the nonzero coefficient index set $T_0 = \varnothing$, the residual estimation error $\mathbf{r}_0 = \mathbf{y}$ and put the initialize iteration counter as $k = 1$. **Identification**. Select a column index n_k of $\mathbf{X}$ that is most correlated with the residual:

$$n_k = \left| \langle \mathbf{r}_{k-1}, \mathbf{X}_n \rangle \right|, \textit{and} \quad T_k = T_{k-1} \cup n_k. \tag{14}$$

Estimation. Compute the best coefficient for approximating the channel vector with chosen columns

$$\mathbf{h}_k = \arg\min_{\mathbf{h}} \left\| \mathbf{y} - \mathbf{X}_{T_k} \mathbf{h} \right\|_2. \tag{15}$$

Iteration. Update the estimation error:

$$\mathbf{r}_k = \mathbf{y} - \mathbf{X}_{T_k}\mathbf{h}_k. \tag{16}$$

Increment the iteration counter k. Repeat (14)-(16) until stopping criterion holds and then set $\tilde{\mathbf{h}}_{OMP} = \mathbf{h}_k$.

C. CCS-CoSaMP estimator

Given the received signal y, the unitary DFT matrix $\mathbf{W}$ and $\mathbf{F}$, training signal $\mathbf{x}$ (training signal matrix $\mathbf{X} = diag(\mathbf{Wx})\mathbf{F}$), the maximum number of dominant channel coefficients is assumed as S. The CCS-CoSaMP performs as follows:

Initializtion. Set the nonzero coefficient index set $T_0 \neq \varnothing$ the residual estimation error $\mathbf{r}_0 = \mathbf{y}$ and put the initialize iteration counter as $k = 1$.

Identification. Select a column index n_k of $\mathbf{X}$ that is most correlated with the residual:

$$n_k = \left|\left\langle \mathbf{r}_{k-1}, \mathbf{X}_n \right\rangle\right|, and\ T_k = T_{k-1} \cup n_k. \tag{17}$$

Using LS method to calculate a channel estimator as $T_{LS} = \arg\min\left\|\mathbf{y} - \mathbf{Xh}\right\|$, and select T maximum dominant taps $\mathbf{h}_{LS}$. The positions of the selected dominant taps in this substep are denoted by T_{LS}.

Merge. The positions of dominant taps are merged by $T_k = T_{LS} \cup T_k$.

Estimation. Compute the best coefficient for approximating the channel vector with chosen columns,

$$\mathbf{h}_k = \arg\min_{\mathbf{h}} \left\|\mathbf{y} - \mathbf{X}_{T_k}\mathbf{h}\right\|_2. \tag{18}$$

Pruning. Select the T_k largest channel coefficients

$$\mathbf{h}_k = \left[\mathbf{h}\right]_S, \tag{19}$$

and replace the left taps by zeros.

Iteration. Update the estimation error

$$\mathbf{r}_k = \mathbf{y} - \mathbf{X}_{T_k}\mathbf{h}_k. \tag{20}$$

Increment the iteration counter k. Repeat (17)-(20) until stopping criterion holds and then set $\tilde{\mathbf{h}}_{CoSaMP} = \mathbf{h}_k$.

IV. SIMULATION RESULTS

In this section, we will compare the performance of the proposed estimators with LS-based linear estimator and adopt 10000 independent Monte-Carlo runs for average. The length of training sequence is $N = 64$. All of the nonzero taps of sparse channel vectors $\mathbf{h}_1$ and $\mathbf{h}_2$ are generated following Gaussian distribution and subject to $\left\|\mathbf{h}_1\right\|_2^2 = \left\|\mathbf{h}_2\right\|_2^2 = 1$. The length of the two channel is $L_1 = L_2 = 32$, and the positions of nonzero channel taps are randomly generated. Transmit power is set as $P_S = N$ and AF relay power is set as $P_R = N$. The signal to noise ratio (SNR) is defined as $\mathbb{S}$ as $10\log(P/\sigma_n^2)$. When the number of nonzero taps in cooperation channels $\mathbf{h}_i$, $i = 1,2$ is changed, the simulation results are shown in Fig. 4~Fig. 6. Channel estimators (LS, CCE-Lasso, CCE-OMP, and CCE-CoSaMP) are evaluated by average mean square error (MSE) which is defined by

$$\text{average MSE}(\Delta\mathbf{h}) = \frac{\left\|\mathbf{h} - \hat{\mathbf{h}}\right\|_2^2}{M(2L-1)}, \tag{21}$$

where $\mathbf{h}$ and $\hat{\mathbf{h}}$ denote channel vector and its estimator, respectively, M is the number of Monte Carlo runs and $(2L-1)$ is the overall length of channel vector $\mathbf{h}$. In Fig. 4, the number of nonzero taps of $\mathbf{h}_i, i = 1,2$, is set to 2 (very sparse), according to the sparseness measurement in Eq.(9), the cooperation convoluted channel also has sparsity. Fig. 3 shows that performance of the proposed CCS estimators is much better than LS estimator and is close to the ideal LS estimator by using known position of the channel. In addition, in low SNR situation (less than 10dB), the proposed CCE-Lasso estimator achieves even better performance than the LS estimator (known position). In Fig. 4, the number of nonzero taps of $\mathbf{h}_i, i = 1,2$, is set to be 4, and in Fig. 5, the number of nonzero taps of hi is set to be 4 and 8, respectively. When comparing the simulation results in the three figures (Fig. 3~Fig. 5), it can be seen that the proposed estimators can exploit the channel sparseness. Better estimator by exploiting more sparser information in a channel and vice versa. It is worth noting that if channels are dense rather than sparse, all of the proposed estimators will have the same performance as LS estimator.

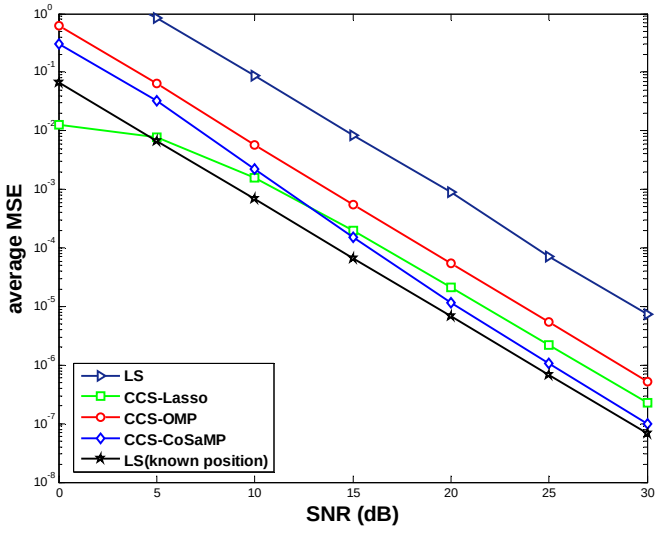


Fig. 3. MSE performance as a function of SNR (Number of nonzero taps of $\mathbf{h}_1$ and $\mathbf{h}_2$ are 2).

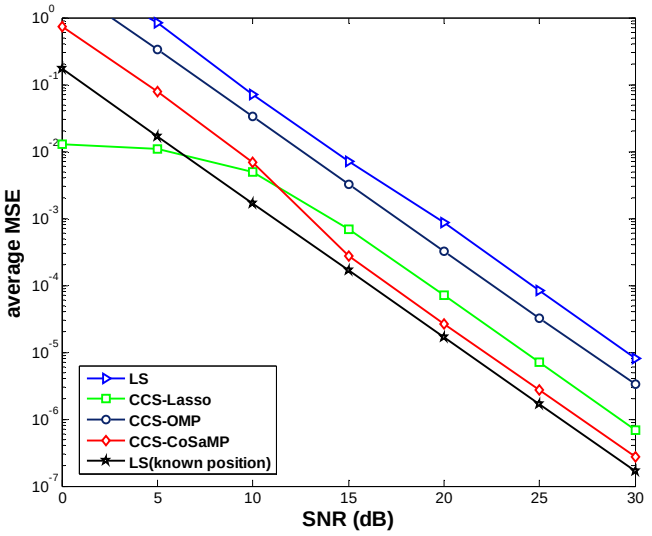


Fig. 4. MSE performance as a function of SNR (Number of nonzero taps of $\mathbf{h}_1$ and $\mathbf{h}_2$ are 4).

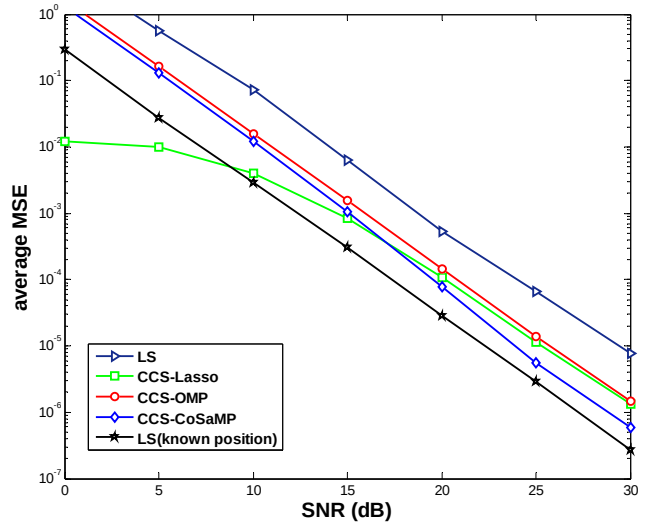


Fig. 5. MSE performance as a function of SNR (Number of nonzero taps of $\mathbf{h}_1$ and $\mathbf{h}_2$ are 4 and 8, respectively).

V. CONCLUSION

In this paper, we have investigated channel estimation problem for relay-based sparse multipath dual-hop single relay cooperation networks. Differing from the conventional linear channel estimation methods, we proposed sparse channel estimation methods for the AF networks under AF protocol. Sparseness of cooperation convoluted channel was confirmed by a measure function. The proposed methods have exploited the sparsity in cooperation channel. Simulation results have confirmed the performance superiority of the proposed method to the conventional linear LS method. The proposed method can also be extended to sparse multipath multi-antenna cooperation networks.

REFERENCES

[1] Cover T. M. and El Gamal A. A., "Capacity theorems for the relay channel," IEEE Transactions on Information Theory, IT-25, 1979.

[2] J. N. Laneman, D. N. C. Tse and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," IEEE Transactions on Information Theory, Vol. 50, No. 12, 2004.

[3] Yiu S., Schober R. and Lampe L., "Distributed space-time block coding," IEEE Transactions on Communications, Vol. 54, No. 7, 2006.

[4] Gao F., Cui T. and Nallanathan A., "On Channel Estimation and Optimal Training Design for Amplify and Forward Relay Networks," IEEE Transactions on Wireless Communications, Vol. 7, No. 5, 2008.

[5] Telatar E., "Capacity of multi-antenna Gaussian channels," European Transactions Telecommunication, Vol. 10, No. 6, 1999.

[6] Goldsmith A., Jafar S. A., Jindal N. and Vishwanath S., "Capacity limits of MIMO channels," IEEE Journal on Selected Areas in Communications, Vol. 21, No. 5, 2003.

[7] Tarokh V., Seshadri N. and Calderbank A. R., "Space time codes for high data rate wireless communication: performance criterion and code construction," IEEE Transactions on Information Theory, Vol. 44, No. 2, 1998.

[8] Alamouti S., "A simple transmit diversity technique for wireless communications," IEEE Journal of Selected Areas in Communication, Vol. 16, No. 8, 1998.

[9] Ribshirani T., "Regression shrinkage and selection via the lasso," Journal of the Royal Statistical Society (B), Vol. 58, No. 1, 1996.

[10] 10. Tropp J. A. and Gilbert A. C., Signal recovery from random measurements via orthogonal matching pursuit, IEEE Transaction on Information Theory, 53, 12 (2007).

[11] 11. Needell D. and Tropp J. A., CoSaMP: Iterative signal recovery from incomplete and inaccurate samples, Applied and Computational Harmonic Analysis, 26, 3 (2009).

[12] Proakis J. G., Digital Communications, 4th edition, New York, NY: McGraw-Hill, 2001.

[13] Bajwa U. W., Haupt J., Raz G. and Nowak R., "Compressed channel sensing," CISS'08, Mar. 19-21, 2008.

[14] Adachi F., Tomeba H. and Takeda K., "Introduction of frequency domain signal processing to broadband single-carrier transmissions in a wireless channel," IEICE Transactions on Communication, Vol. E92-B, No. 09, 2009.

[15] Gray R. M., "Toeplitz and Circulant Matrices: A Review," Foundations and Trends in Communications and Information Theory, Vol. 3, No. 2, 2006.

[16] Hoyer P. O., Non-negative Matrix Factorization with Sparseness Constraints, "The Journal of Machine Learning Research," Vol. 5, No. 2004, 2004.

[17] Candes E., Romberg J. and Tao T., "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," IEEE Transaction on Information Theory, Vol. 52, No. 2, 2006.

[18] Donoho D. L., "Compressed sensing," IEEE Transactions on Information Theory, Vol. 52, No. 4, 2006.

[19] Candes E. J., "The restricted isometry property and its implications for compressed sensing," Compte Rendus de l'Academie des Sciences, Paris, Serie I, 346, 2008.

A Damping Approach in DMTL's Packaging for Reducing Release Time

DENG CHENG[1], BAO JING-FU[1], DU YI-JIA[1], ZHAO XING-HAI[2]

[1](School of Electronic Engineering in University of Electronic Science and Technology of China, No.2006 Xiyuan Avenue of Hi-tech west zone in Chengdu 611731 email- baojingfu@uestc.edu.cn)

[2](China Academy of Engineering Physics, Mianyang 621900,China）

Abstract This paper reports on the design, simulation, and model of a damping approach that is adding a extra bulge inside the top package layer in DMTL(distributed microelectromechanical transmission-line). With this new structure, performances of this phase shifter are improved, especially its release time that reduces to 358 μs while the original is 670 μs. As this package has a damping effect on the MEMS bridge, the speed of DMTL nearly increases two times and the capacitance fluctuation is greatly inhibited.

Key words DMTL; release time; speed of DMTL; damping approach; MEMS package, mechanical equivalent

1 INTRODUCTION

Microelectromechanical system (MEMS) electronically variable phase shifters have been developed for their usefulness in phased-array radars [1,2], communication systems [3] and measurement instrumentation [4] during the past ten years. The advantages of using MEMS over FET's or p-i-n diodes are their low-loss performance [5,6] and lack of measurable intermodulation distortion [7,8]. However, the application of MEMS phase shifters in wider range is restricted by its low power handling and long switching time which don't match high-performance passive phased arrays. Although various MEMS phase shifters were demonstrated at different bandwidths, especially the digital distributed micro electromechanical transmission-line (DMTL) at X-band [9], Ku-band [10,11], K-band [12], Ka-band [13,14], V-band [15] and even W-band [16] frequencies by several groups, most of results don't indicate the switching time which determines the speed of MEMS phase shifters. Switching time of FET's or p-i-n diodes is much higher (approximate 1-100ns) than MEMS's (1-300 μs). In order to reduce the MEMS's switching time, the optimization of materials and some parameters, such as driven voltage, height of gap, and width of CPW signal line is tried. However, switching time of MEMS phase shifters decreases a little because these parameters affect the pull-in time of decades μs order much more than release time of hundreds μs order[17][18]（Switching time is the sum of pull-in time and release time）. What is remarkable about switching time of the DMTL approach is how greatly one can reduce the release time. The release time is required for the MEMS bridge to move from the down state to its normal height (or within 5% of this value).

In this paper, a novel approach—etching a extra bulge inside the top package layer in DMTL—is introduced. This bulge is above the MEMS bridge and nearly contacts with the bridge at its original height. In the release of the bridge, a damping force appears when the bridge reaches above its original height. This force stops the bridge's motion effectively and the release time reduces by half. The total switching time also drops. Additionally, this approach prolongs the MEMS bridge's lifetime by preventing it from vibration.

This work was supported by Joint Fund of the National Natural Science Foundation of China and the China Academy of Engineering Physics(11176006), National Ministries Foundation of China (Grant No. 9140A23070311DZ0210) and Chinese Academy of Engineering Physics Science and Technology Development Foundation (Grant No. 2008A0403016).

2 BASICS OF RELEASE TIME

2.1 DMTL structure

One of the most attractive MEMS phase shifters is DMTL which contains many periodically distributed (Fig.2(a)), an electrostatic force occurs between the bridge and the bottom electrode when the bias voltage is applied between them.

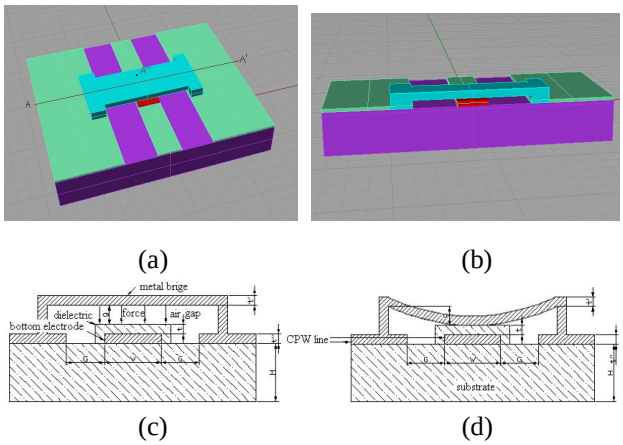


Fig.1. Schematic of a typical unit in MEMS phase shifters in (a) vertical view; (b) front view (cut along A-A' in (a)) and sectional view in (c) up and (d) down state.

Then the bridge is bent until the cantilever abruptly falls to the bottom contact. The time that it takes from original height to the bottom contact is named as pull-in time. Fig.2 presents a schematic diagram of MEMS bridge state before actuation and after falling down, respectively.

2.2 Release time

Pull-in time depends on the applied voltage, the effective stiffness of the bridge and the gap between the bridge and bottom electrode. It can be calculated by finite element method (FEM) simulations [19] or nonlinear differential equation[20]. Release time is relative to the effective stiffness of the bridge, the gap between the bridge and bottom electrode and some other factors. The first part of Release time—opening time can be estimated by the principle of energy conservation [21]. The result should be discussed with and without considering the damping respectively. Ignoring the damping, the energy conservation principle is

$$\frac{1}{2}m\dot{x}^2 = \frac{1}{2}kd^2 - \frac{1}{2}kx^2 \qquad (1)$$

The velocity can be solved

$$\dot{x} = -\sqrt{\frac{kd^2 - kx^2}{m}} \qquad (2)$$

The opening time can be gained from integrating $dt = dx / \dot{x}$ from x=d to x=0

$$t_o = \int_d^0 \frac{dx}{\dot{x}} = \frac{\pi}{2\omega_0} \qquad (3)$$

Ignoring the inertial effects (m=0) and considering the damping ($\gamma \neq 0$), we have employed classical Newtonian mechanics to predict the opening time

$$m\ddot{x} + \gamma\dot{x} + kx = 0 \qquad (4)$$

The solution is $x = de^{-(\gamma/k)t}$, then t= γ /k ln5.

After the bridge reaches its normal height, it requires much more time for the bridge to stop vibrating, especially settle within 5% of its original height. In this paper, a novel approach in package is introduced to greatly reduce time consumption and increase MEMS phase shifters' speed.

3 DESIGN

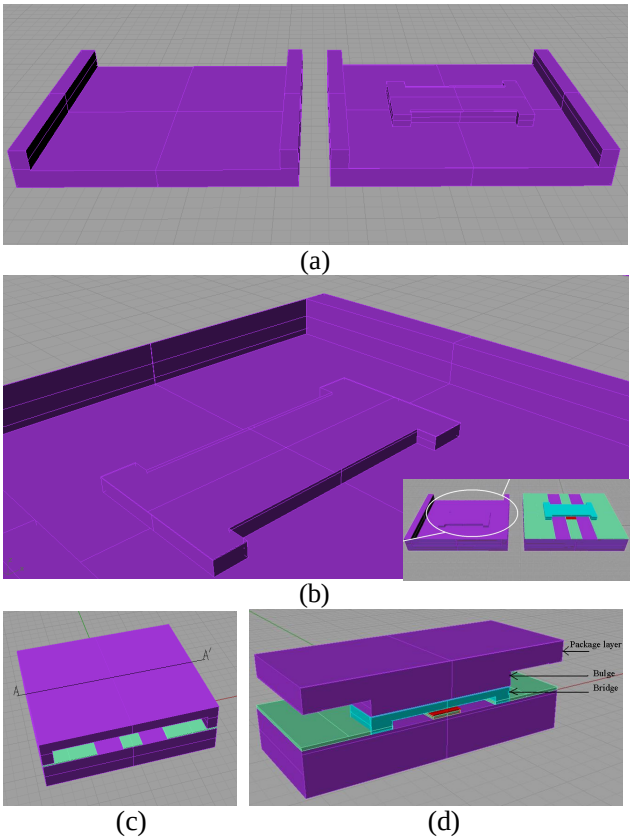


Fig.2. Schematic of (a) traditional (left) and novel(right) package layer; (b) detail of the bulge on the cap layers; (c) a cap layer covering on MEMS bridge unit; (d) front view (cut along A-A' in (c)).

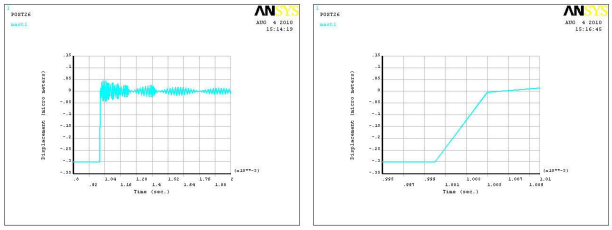
A novel bulge has been designed on the cap layer (top-most layer) of package. Package layer is common in MEMS devices (Fig.2(a)). Based on the existent cap

layer, a new layer with a bulge is designed in ANSYS10.0 (Fig.2(a)). Its detail is showed in Fig.2(b). The shape of the bulge is same as the bridge's in order to enlarge their contact area at the bridge's height (Fig.2(d)). In the release process, the bridge collides with the bulge and its vibration is impeded greatly; therefore release time drops.

In this design, a SiO_2 cap layer is chosen and thought to be etched with the bridge's shape (Fig. 2(b)) on the inner side of the layer. When it covers on the bridge unit (Fig. 2(b)), the cap layer (top-most layer) over a DMTL unit has a bulge nearly contacts with the up-state bridge (Fig.2(c)). After the bridge reaches to its normal height from the bottom and contact with the bulge, a collision happens with a damping effect (Fig. 2(c)).

4 SIMULATION

4.1 Performance of traditional structure



(a) (b)

Fig.3. Displacement of point A'' (see Fig. 1(a)) in (a) the release process and (b) detail of the time range from 995 μ s to 1010 μ s.

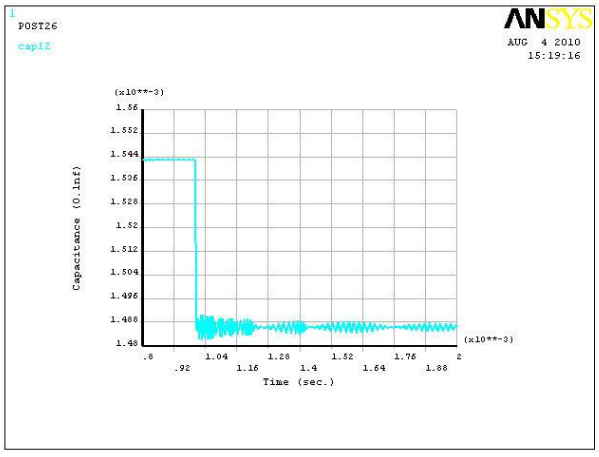


Fig.4. Capacitance of the MEMS bridge unit in the release process

ANSYS10.0—a commercial software, is used for simulating the release time of MEMS bridges. We have simulated the release time of a DMTL's bridge, which needs very short release time as well as pull-in time. In the simulation, the displacement of point A''(see Fig. 1(a))—the center of the bridge—is illustrated in Fig.3.

From Fig.3, the vibration of the bridge near its original height is intense after the applied voltage is removed. As shown in Fig.4, the capacitance of the MEMS bridge unit is fluctuant because of this vibration and this fluctuation is harmful in the operation of DMTL phase shifters.

Fig. 3 show the release time is approximate 670 μ s, while it takes the bridge only 4 μ s to reach its original height. That is to say, what is remarkable about release time of the DMTL is how greatly one can reduce the former one (670 μ s). Considered it's caused by the vibration near the bridge's normal height, a novel approach in package is introduced, as described in part 3. With a bulge on the inner side of the cap layer, this kind of vibration in bridge's release is restrained as Fig.2(d) shows.

4.2 Performance of novel structure

Measured the time when the bridge starts to stay within 5% of its normal height, the release time of DMTL unit with bulge cap layer is obtained (Fig.5). Compared with normal DMTL phase shifter (Fig.3(a)), the release time of the bridge reduces to only 358 μ s and save 312 μ s to settle.

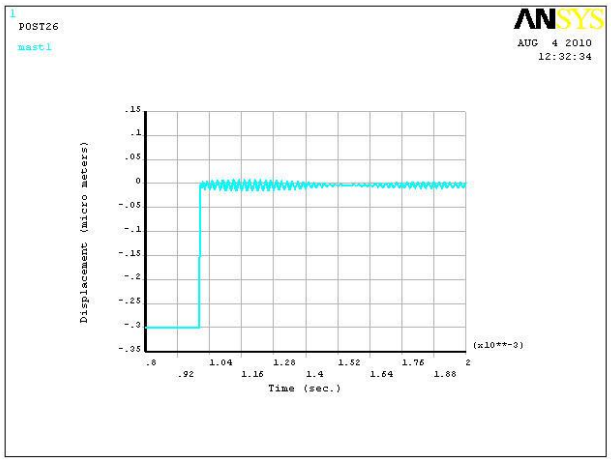


Fig.5. The displacement of point A'' with block shield in the release process

5 MECHANICAL EQUIVALENT

To explain the simulation results, we use a simple equivalent mechanical model composing of a discrete spring, mass and damper (Fig.6). This model treats the bridge as a single lumped mass with a spring, normal damper and a new damper that represents the influence of the bulge in package layer. The new damper is above the mass and suspended on the top clamp. The original height of the damper's one end is equal to the normal height of the bridge. When the bridge moves above its original height, its movement is impeded. But if the bridge is below this height, the damper doesn't work. The classical Newtonian mechanics can be used to predict this behavior under the bulge damping force. The model would accurately predict the dynamic behavior of the MEMS bridge with this new package structure.

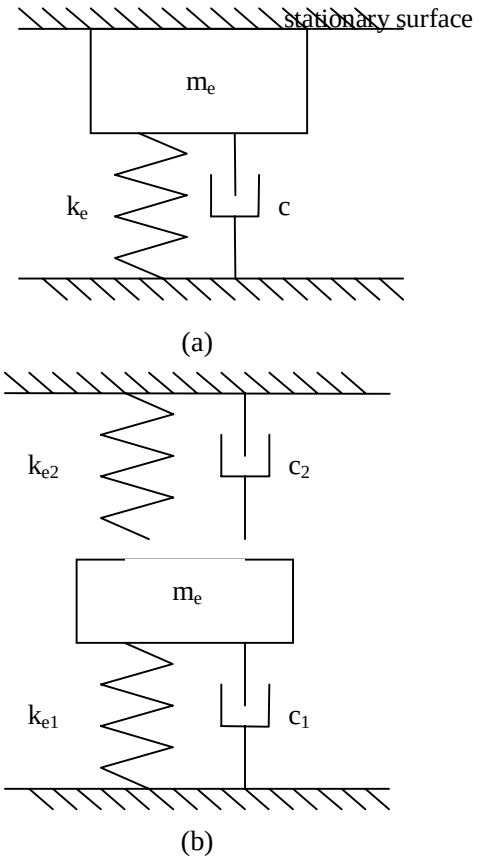


Fig.6. Equivalent mechanical model of the bridge with (a) rigid and (b) flexible block shield.

Fig.6 shows that a rigid block shield is equivalent to a stationary surface (Fig.6(a)) and a flexible block shield is equivalent to a combination of a spring k_{e2} and a damper c2 (Fig.6(b)). k_{e2} and c_2 are the equivalent elastic and damping coefficient of the novel structure, respectively. k_{e1} and c_1 are the equivalent elastic and damping coefficient of MEMS bridge. m_e is the equivalent mass of MEMS bridge .

6 CONCLUSION

In this paper, a novel approach in package is introduced in the design of DMTL units and improves the performance of DMTL phase shifters, especially its release time. With this bulge structure on the package layer, the release time of MEMS bridge approximately reduces by half. Future work will be centered on the establishment of its accurate mechanical model and researches of the lifetime's improvement.

参 考 文 献

[1] Constant B, Nicole P, Menager L et al.. MEMS technology for Radar front end modules. 2009 International Radar Conference Radar. IEEE. 2009, pp. 6.

[2]Malmqvist R, Samuelsson C, Gustafsson A et al.. On the use of MEMS phase shifters in a low-cost ka-band multifunctional ESA on a small UAV. 2007 Asia-Pacific Microwave Conference. IEEE. 2008, pp. 969-72.

[3] Fernaacutendez M, Vasylchenko A, Dainesi P et al.. Dipole antenna and distributed MEMS phase shifter fully integrated in a single wafer process for beam steering applications. Microelectronic Engineering, vol.87, no.5-8, May-Aug. 2010, pp. 1290-3.

[4] J. S. Hayden and G. M. Rebeiz, "Very low-loss distributed X-band and Ka-band MEMS phase shifters using metal–air–metal capacitors," IEEE Trans. Microwave Theory Tech., vol. 51, pp. 309–314, Jan. 2003.

[5] Thakur K, SumithraDevi A, Ranjitha I. Performance of low loss RF MEMS Fixed – Fixed capacitive switch characterization. AEMC, 2009 P 1 – 4.

[6]Sinjari A, Chowdhury S. High performance 77 GHz Single Pole Triple Throw (SP3T) MEMS switch. IEEE NEWCAS-TAISA '09, 2009.

[7] Kaynak, M.; Ehwald, K.E.; Scholz, R.; Korndorfer, F.; Wipf, C.; Sun, Y.M.; Tillack, B.; Zihir, S.; Gurbuz, Y.; Characterization of an embedded RF-MEMS switch. Silicon Monolithic Integrated Circuits in RF Systems (SiRF), 2010. Page(s): 144 – 147.

[8] Mafinejad, Y.; Kouzani, A.Z.; Mafinezhad, K.; Nabovatti, H.; Design and simulation of a low voltage wide band RF MEMS switch. Systems, Man and Cybernetics, 2009. Page(s): 4623 – 4627.

[9] Morton MA, Papapolymerou J. A packaged MEMS-based 5-bit X -band high-pass/low-pass phase shifter. IEEE Transactions on Microwave Theory and Techniques, vol.56, no.9, Sept. 2008, pp. 7.

[10] Perruisseau J, Topalli K, Akin T et al.. Low-loss Ku-band artificial transmission line with MEMS tuning capability.

IEEE Microwave and Wireless Components Letters, vol.19, no.6, June 2009, pp. 377-9.

[11] Yamane D, Yamashita K, Seita H et al.. A dual-SPDT RF-MEMS switch on a small-sized LTCC phase shifter for Ku-band operation. 2009 Asia Pacific Microwave Conference (APMC 2009). IEEE. 2009, pp. 555-8.

[12] Y. Liu, A. Borgioli, A. S. Nagra, and R. A. York, “K-band 3-bit low-loss distributed MEMS phase shifter,” IEEE Microwave Guided Wave Lett., vol. 10, pp. 415–417, Oct. 2000.

[13] Rock C, Hudson T, Wolfson Bet al.. Design, Fabrication and Evaluation of a MEMS-Based, Ka-Band, 16-Element Sub-Array. 2010 IEEE Aerospace Conference. IEEE. 2010, pp. 17.

[14] Gautier W, Ziegler V, Stehle A et al.. RF-MEMS phased array antenna on low-loss LTCC substrate for Ka-band data link. 2009 39th European Microwave Conference (EuMC). IEEE. 2009, pp. 914-17.

[15] Hong-Teuk K, Jae-Hyoung P, Sanghyo L et al.. V-band 2-b and 4-b low-loss and low-voltage distributed MEMS digital phase shifter using metal-air-metal capacitors. IEEE Transactions on Microwave Theory and Techniques, vol.50, no.12, Dec. 2002, pp. 2918-23.

[16] Stehle A, Georgiev G, Ziegler V et al.. RF-MEMS switch and phase shifter optimized for W-Band. 2008 38th European Microwave Conference. IEEE. 2008, pp. 104-7.

[17] Schauwecker B ; Strohm M; Mack T; Simon W; Luy. Single-pole-dpuble-throw switch based on toggle switch, J.-F. Electronics Letters, v 39, n 8, p 668-670, April 17, 2003.

[18] Peroulis D, Pacheco P,Sarabandi K, Katehi B. Electromechanical Considerations in Developing Low-Voltage RF MEMS Switches. IEEE Trans. Microwave Theory Tech., vol. 51, NO. 1, JANUARY 2003.

[19] Chan K, Kan C, Dutton W. Nonlinear dynamic modeling of micromachined microwave switches. Proceedings of IEEE MTT-S Symposium, 1997.

[20] Huang M, Liew M, Wong H et al.. Mechanical design and optimization of capacitive micromachined switch. Sensors and Actuators A 93: 273–285.

[21] Kaajakari V. Closed form expressions for RF MEMS switch actuation and release time. Electronics Letters, v 45, n 3, p 149-150, 2009.

Transient Noise Excision for Skywave Over-The-Horizon Radar Based on Adaptive Gaussian Chirplet Decomposition

Wei You
Department of electronic engineering,
University of Electronic Science and Technology of China
Chengdu, China
yoyouv@163.com

Qian He
Department of electronic engineering，
University of Electronic Science and Technology of China
Chengdu, China

Zishu He
Department of electronic engineering,
University of Electronic Science and Technology of China
Chengdu, China

Shuangling Wang
Department of electronic engineering ，
University of Electronic Science and Technology of China
Chengdu, China

Abstract—**Removing transient noise from returned signal is an important task for OTH radar signal processing. Compared with the traditional techniques, the adaptive Gaussian basis representation (AGR) based method does not require data interpolation and thereby enhances the transient interference excision performance. In this paper, we propose to employ adaptive Gaussian chirplet decomposition (AGC) to replace the AGR. The new approach inherits the advantage of the AGC based method where data interpolation is not required. Further, it is superior in handling nonlinear radar returns. This is a big advantage for OTH radar systems, where clutter component in the returned signal is usually nonlinear due to ionosphere contamination. The effectiveness of the propose method is verified using simulated and experimental results.**

Keywords-skywave over-the-horizon radar; transient noise excision; adaptive Gaussian chirplet decomposition;

I. INTRODUCTION

Skywave over the horizon radar (OTHR) transmits signal via the ionosphere. OTHR works in HF band (3-30MHz). It can see over the horizon and detects targets over several thousand kilometers away [1]. As it transmits and receives return signal to and from the sky, it often encounters strong transient noise, whose time duration is very short compared with the radar coherent processing interval (CPI) [2]. Transient noise is broadband which may mask the moving target and thus degrade the detection performance of the radar. The ocean clutter is another main interference of the OTHR returned signal. Different from the transient noise, the clutter return is often narrow band with longer time duration. In order to improve the detection performance of the OTHR, it is necessary to get rid of the transient noise and clutter components from the received signals. Consider the case where transient noise excision is performed before the clutter elimination procedure. Since the subsequent clutter elimination depends heavily on the clutter characteristic, to guarantee the elimination performance, it is required not to distort the clutter components during the transient noise excision process.

In this paper, we focus on the transient noise excision. Some algorithms have been proposed in the literature to excise the transient noise. The authors in [3] proposed to locate the transient noise after clutter suppression. Once the location determined, the transient noise can be removed and recovered using Burg's linear prediction algorithm. In [4], the authors proposed to detect the transient noise using SVD technique. In [5-6], the transient noise was modeled using an impulse signal, while the clutter was modeled using a complex cosine component. In their algorithm, the transient noise was detected or excised in frequency domain.

Most of the algorithms proposed in literature need to use linear prediction technique to recover data segment corrupted by the transient noise. The performance of these algorithms will be degraded when the noise occupies long time interval. Recently, Guo et al [7] proposed an algorithm based on the linear adaptive Gaussian representation (AGR). The AGR has the advantage of no data interpolation is needed over the traditional methods. Here, we propose to employ the adaptive Gaussian chirplet (AGC) decomposition to replace the AGR based method. For the skywave OTHR, non-linear phase can be caused by the ionosphere. If the transient noise goes through the ionosphere, then it becomes the nonlinear chirp [8]. Comparing with the AGR based method, AGC with non-linear phase not only inherits the advantage of AGR, it is more suitable to match the contaminated skywave OTHR return

This work is supported by the National Science Foundation of China under Grant 11076006, 61032010, and 61102142, and by the Fundamental Research Funds for the Central Universities under Grant ZYGX2009J019.

signal and thus has better performance in excising the transient noise.

This paper is organized as follows. Section Ⅱ introduces the signal model used in the paper, the excision algorithm based on the AGC decomposition is presented in section Ⅲ. In section Ⅳ, the simulation results will be presented to verify the effectiveness of the algorithm. And section Ⅴ concludes the paper.

II. SIGNAL MODEL

The return signal of the OTHR can be expressed as:

$$r(t) = c(t) + i(t) + s(t) + v(t) \tag{1}$$

where $c(t)$ is the clutter component, $i(t)$ is the transient noise, $s(t)$ is the target return, and $v(t)$ is the complex Gaussian noise.

Experimental result has shown that when the ocean surface is illuminated by HF electromagnetic waves, two Bragg components center at the transmitted frequency can be received and the clutter can be modeled as [9]

$$c(t) = A_I e^{j2\pi f_b t} + A_Q e^{-j2\pi f_b t} \tag{2}$$

where f_b is the Bragg frequency, A_I and A_Q contain the magnitude and initial phase of the two components, respectively. For skywave radars, as the signal is transmitted through the ionosphere, the phase in (2) is often non-linear due to the ionosphere contamination [10].

The transient noise in (1) may include the lighting and/or the echo of meteor trail. These two types of transient noise are typically 20-40 dB stronger than the internal noise of the receivers which significantly degrades the target detection performance of the OTHR.

The target return $s(t)$ in (1) is usually modeled as a Doppler component with a frequency determined by the target velocity and the radar working frequency.

III. PROPOSED ALGORITHM

A. AGC Decomposition

In AGC decomposition, a signal is decomposed using a series of AGC functions. The $k-th$ function is characterized by four parameters: variance σ_k^2, time center t_k, frequency shift w_k, and the frequency modulation factor β_k. The AGC base function can be expressed as [11]:

$$g_k(t) = \left(\frac{1}{\pi\sigma_k^2}\right)^{\frac{1}{4}} \exp\{-\frac{1}{2\sigma_k^2}(t-t_k)^2 + jw_k(t-t_k) + j\frac{\beta_k}{2}(t-t_k)^2\} \tag{3}$$

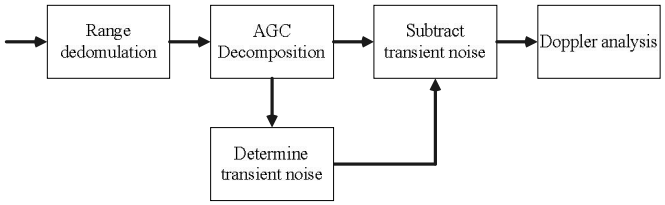


Figure 1. The projection factor of AGC decomposition(simulated data)

where $\sigma_k > 0, t_k, w_k, \beta_k \in R$.

For signal $x(t)$, the four parameters of which are determined by maximizing the following projection energy

$$\max \left| \frac{< x(t), g_k(t) >}{< g_k(t), g_k(t) >} \right| \tag{4}$$

The base function $g_k(t)$ which maximizes the equation will be used to decompose the signal. Note that the calculation in (4) is a four dimensional search, the genetic algorithm is introduced for the computation and find the optimal solution of the parameters.

B. Transient noise excision

Using the AGC decomposition, the signal $x(t)$ can be approximated by the sum of a series of AGC base functions. Each base function is characterized by the previously mentioned four parameters. From (3) it can be seen that a smaller σ_k^2 implies a shorter pulse, and vice versa. In the signal model of the OTHR return in (1), the transient noise can be modeled using a short pulse, while the ocean clutter can be modeled using a longer pulse. After the AGC decomposition, the transient noise can be excised simply by subtracting the terms with relatively small σ_k^2.

The procedure of the excision algorithm based on AGC decomposition is shown in Figure 1. The return signal is firstly range demodulated (pulse compression), and the range-Doppler data is derived. Then the return signal of each range cell is decomposed using the AGC algorithm, which is described as follows:

Initialize: For the return signal $r(t)$ as per (1), let $k = 1$, $x_k(t) = r(t)$.

(a) In Eq. (4), compute the four parameters σ_k, t_k, w_k, β_k for $x_k(t)$. Genetic algorithm is used to find the optimal solution. The bounds of the four parameters are determined as $\sigma_k \in [1/F_s, T_c/4]$, where F_s is the sampling rate, T_c is the coherent integration time, $t_k \in [1/F_s, T_c - 1/F_s]$, $w_k \in [-\pi F_s, \pi F_s]$, and $\beta_k \in [-f_{d\max}F_s/100, f_{d\max}F_s/100]$, $f_{d\max}$ is the largest frequency shift of each AGC base function, the value of which depends on how the ocean clutter is distorted by the ionosphere.

(b) Let $x_{k+1}(t) = x_k(t) - A_k g_k(t)$, where projection factor $A_k = < x_k(t), g_k(t) > / < g_k(t), g_k(t) >$

(c) If the residual energy of $x_{k+1}(t)$, $\|x_{k+1}(t)\|^2 = < x_{k+1}(t), x_{k+1}(t) >$ is small enough, go to (d), otherwise, let $k = k+1$ and go to (a). The threshold used is $\|r(t)\|^2 / 500$.

(d) Let $K = k$, and check the σ_k of each AGC function. Those functions with small σ_k are regarded as transient noise and subtracted from the radar return. The threshold for σ_k is $\sigma_{\max} / 10$, where $\sigma_{\max}$ is $\max\{\sigma_1, \sigma_2, ..., \sigma_K\}$.

C. Transient noise detection

As it can be seen, the excision algorithm based on AGC decomposition can only announce the existence of the transient noise after the return signal is fully decomposed. And thus the excision procedure can be triggered. This is not effective as the transient noise may only exist in several range cells, only return signal in these cells need to be processed.

Differential operator is a good arithmetic for detecting sharp transients [12]. In this paper, it is used to detect the transient noise in the return signal. If noise is detected by the differential operator, then return signal in this range cell will be processed further via AGC decomposition, otherwise, it will go on to process the next cell. For a discrete signal $r[n]$, the output of the differential operator is

$$d[n] = r[n] - r[n-1] \tag{5}$$

For the clutter signal, as it is modeled as low frequency Doppler components, it is smooth in time domain. The output of the differential operator will be small. For the transient noise, in time domain, it is a sharp pulse; the output will also be sharp. With this property, the transient noise can be detected easily by comparing the output with a given threshold. If the output is larger than the threshold, a transient noise is announced, otherwise not. This step can be done before AGC decomposition. As the computation cost for the differential operator is very small comparing with the AGC decomposition, this procedure can dramatically decrease the computation time for transient noise excision.

IV. SIMULATION RESULTS

In this section, the simulation results will be presented to show the performance of the proposed transient noise excision algorithm. The result is compared with the already proposed AGR algorithm.

A. Result of simulated data

The simulation parameters are: working frequency is 15.8 MHz, the clutter is contaminated by a multiplicative noise $\cos(2\pi * 0.4t)$, PRF is 40Hz, CNR is 45dB, and 256 samples are processed in one CPI. The magnitudes of the Bragg components are 7 and 8, respectively. Two transient noise, $400w(t-1.25)\exp(j2\pi(t+20t^{3.4}))$, $400w(t-2.5)\exp(j2\pi(-3t-20t^{3.7}))$ are embedded.

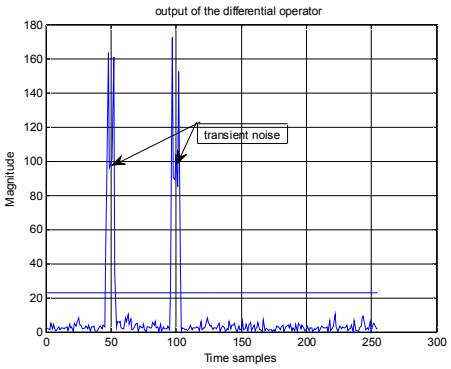


Figure 2. The output of the differntial operator(simulated data)

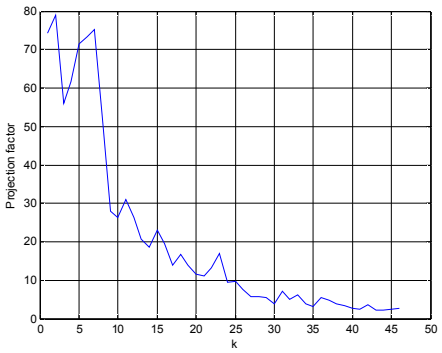


Figure 3. The projection factor of AGC decomposition(simulated data)

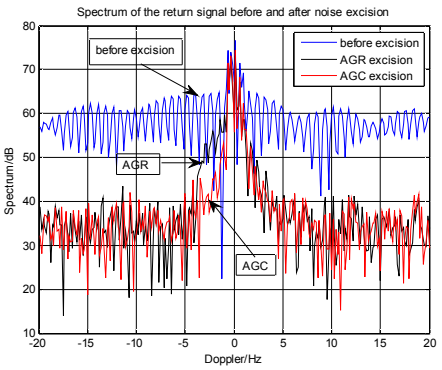


Figure 4. The spectrum before and after noise excision(simulated data)

Figures 2-4 show the processing result using the simulated data. Following the algorithm described in section III, the returned signal is firstly processed by the differential operator. If transient noise detected, then the signal will be processed by the AGC decomposition.

Figure 2 shows the result output of the differential operator, the output of the simulated return signal is plotted using solid line while the threshold used to detect the transient noise is plotted using dashed line. From the result, the transient noise can be detected, and the result matches with the embedded transient noise positions.

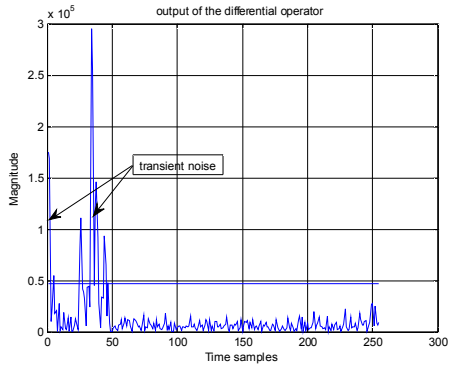


Figure 5. The output of the differntial operator(raw data)

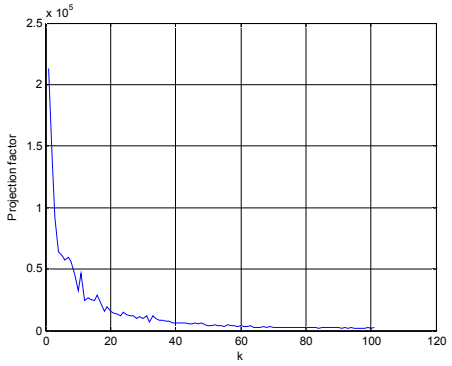


Figure 6. The projection factor of AGC decomposition(raw data)

As the transient noise is detected by the differential operator, the returned signal is further processed by the AGC algorithm. Figure 3 shows the projection factor of the AGC decomposition. For the AGC decomposition, 47 functions are used, while 52 functions are used by the AGR algorithm.

After the AGC decomposition, the functions with small variance are regarded as transient noise and subtracted from the original return signal. Figure 4 shows the spectrum before and after noise excision. The spectrum gain of AGC is 21.6 dB, and it is 20.88 dB for the AGR algorithm.

B. Result of real data

Figures 5-7 show the processing result of the real raw data. The data was captured in China using an experimental skywave OTH radar operating at 15.813 MHz, whose PRF is 47.6Hz and 256 samples are processed in one CPI.

Figure 5 shows the output of the differential operator. From the result, after the processing of the operator, two sharp pulses can be detected by comparing with the threshold which is plotted in dashed line. The result in the figure matches well with the real transient noise positions in the radar return signal.

Figure 6 shows the projection factor calculated by the algorithm. From the figure, the projection factor converges along with the decomposition when more AGC functions are used for the decomposition. It also shows that 102 functions are used, while for the AGR based algorithm, 104 functions are used.

Figure 7 shows the return spectrum before and after transient noise excision. By calculation, the spectrum gain of AGC is 8.9 dB, while for AGR based algorithm, the gain is 7.4 dB.

Both the results show that AGC based method is superior to the AGR when the returned signal is contaminated by the ionosphere. The reason is that when the radar return signal is contaminated by the ionosphere, the ocean clutter will have nonlinear phase and the transient noise becomes nonlinear chirp modulated. AGC base function with nonlinear phase can better match the return signal thus has better performance in excising the noise.

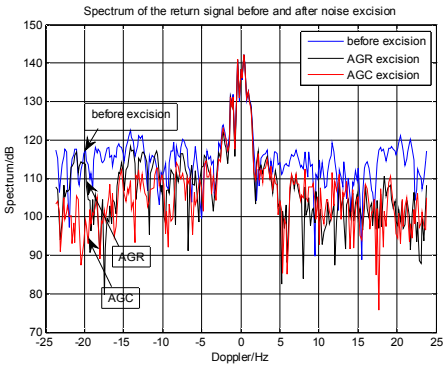


Figure 7. The spectrum before and after noise excision(raw data)

V. CONCLUSION

A transient noise excision algorithm based on the AGC decomposition was proposed in this paper. When the returned signal contaminated by the ionosphere contains nonlinear phase components, the AGC based method can better match the radar return and has better performance in excising the transient noise, compared with the old AGR based method. Genetic algorithm was employed to find the optimal parameters of the AGC functions. Simulations were carried out using real experimental skywave OTHR raw data and simulated data, which verified the effectiveness of the proposed method.

REFERENCE

[1] J. Headrick, HF over-the-horizon radar, in: Radar Handbook, 2nd ed., Merrill I. Skolnik, Ed. McGraw-Hill [M], 1990.

[2] T. Liu, Y. Gong, et al, OTHR impulsive interference suppression in strong clutter background, IEICE Trans. vol.E92-A, no.11, 2009, pp. 2866-2873.

[3] M.D Xing, Z. Bao, Y. Qiang , Transient interference excision in OTHR, Acta Electronica Sinica, vol. 30, no. 6, 2002, pp.823–826.

[4] X.X. Chen, Y.H. Huang, A SVD-Based approach of suppressing transient interference in high-frequency Radar, Journal of Electronics and Information Technology[J], vol.27, no.12, 2005, pp.1879-1882.

[5] T. Liu, J. Wang, OTHR impulsive interference detection based on AR model in phase domain, WSEAS Trans. on Signal Processing [J], vol.5, no.4, 2009, pp.147-156.

[6] T. Liu, X.X. Chen, J. Wang, Y.H. Gong, Subspace impulsive interference suppression in OTHR, Progress In Electromagnetics Research C [J], 2009, vol.7, pp.167-181.

[7] X. Guo, H. Sun, T. Yeo, Transient interference excision in over-the-horizon radar using adaptive time-frequency analysis, IEEE Trans. on Geoscience and Remote Sensing, vol. 43, no. 4, 2005, pp: 722–735.

[8] National Instruments, Signal Processing Toolset User Manual, chapter 3, 2001, pp. 1-5.

[9] R. Khan, Ocean-clutter model for high-frequency Radar, IEEE Journal of Oceanic Engineering, vol.16, no.2, 1991, pp.181-188.

[10] K. Lu, X.Z. Liu, Y.T. Liu, Ionospheric decontamination and sea clutter suppression for HF skywave radars, IEEE Journal of Oceanic Engineering [J], vol. 30, no. 2, 2005, pp.455-462,

[11] Q. Yin, S. Qian, et al, A fast refinement for adaptive Gaussian chirplet decomposition, IEEE Trans. On Signal Processing, vol.50, no. 6, June 2002, pp.1298-1306.

[12] J. Qian, J.S. Barlow, M.P. Michael, A simplified arithmetic detector for EEG sharp transients-prilimrary results, IEEE Trans. on Biomedical Engineering, vol.35, no. 1, 1988, pp.11-18.

Active-G_m-RC Bandpass Filter with 60MHz Center Frequency and a Combined Analog-Digital Tuning System

Jingbo Shi, Takayuki Konishi, Shoichi Masui

Tohoku University, Sendai, Miyagi, 980-8577, Japan
Email: jbshi@riec.tohoku.ac.jp

***Abstract* — An active-G_m-RC bandpass filter (BPF) with a combined analog-digital tuning system is presented. The active-G_m-RC BPF biquad features a positive feedback capacitor to enhance the quality factor, and its required number of OTA per filter-order becomes 0.5. A design example for the application to an on-chip equivalent SAW filter with 60MHz center frequency and 8MHz bandwidth is presented. The proposed tuning system can be enabled with a configurable active-G_m-RC BPF/LPF. Behavioral simulations present the proper tuning for the process variation of ±20% in resistors and capacitances, the power can be reduced by a factor of 18.8 over the active-RC BPF.**

***Index Terms* — Active-G_m-RC, bandpass filter, lower power design, digital tuning, wireless receiver.**

I. INTRODUCTION

With a rapid development of modern communication systems, demands for wireless transceiver have been imposing researches on low-power multiband/multimode highly-integrated CMOS-based ICs with reduced bill of material (BOM). The commonly targeted architectures for the receiver part of multiband/multimode transceivers are a SAW(Surface Acoustic Wave)-less zero IF and low-IF; then, various approaches to minimize the blocker sensitivity and noise figure degradation have been proposed for the applications of cellular and digital TVs. Further challenges to BOM and investigate new receiver architectures have been made by integrating an on-chip equivalent SAW filter by active-RC bandapss filter (BPF) [1]. The approach has been published as a 5th-order BPF with 60MHz center frequency and 8MHz bandwidth to cover the standard IF band (44MHz) operation of a conventional SAW filter; however, the obtained power consumption increases up to 23.5mW since the number of used operational transconductance amplifiers (OTAs) is 2 per filter-order (10 for 5th-order BPF), and the associated OTA unity gain frequency becomes 3GHz. This paper proposes an active-G_m-RC BPF to the equivalent SAW filter application with 60MHz center frequency and 8MHz bandwidth. The resulting power consumption can be reduced to 3mW in actual implementation.

In the integrated filter design, the active-RC and G_m-C architectures are commonly used as IF and channel selection filters with the cut-off frequency of several MHz or below in wireless receivers. The application of the active-RC filter to the equivalent SAW-filter is not considered as appropriate from the standpoint of the power consumption. Instead, G_m-C architecture enables lower power and noise as a result of a simple transconductor topology [2]; while its linearity characteristics are inferior to the active-RC filters.

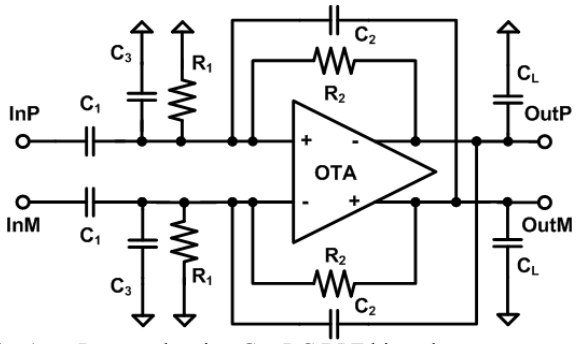


Fig. 1. Proposed active-Gm-RC BPF biquad.

The adopted active-G_m-RC architecture can merge the advantages from the conventional active-RC and G_m-C filters [3][4]. By taking those advantages of the active-G_m-RC architecture, this paper explores an optimized active-G_m-RC BPF and proposes a combined analog-digital tuning system.

II. II. ACTIVE- G_m-RC BPF BIQUAD DESIGN

We propose an optimized active-G_m-RC BPF biquad as shown in Fig. 1. A typical transfer function of a general BPF biquad, $H_{BPFb}(s)$, is given by

$$H_{BPFb}(s) = -\frac{A_{biq}\dfrac{\omega_{c,biq}}{q_{biq}}s}{s^2 + \dfrac{\omega_{c,biq}}{q_{biq}}s + \omega_{c,biq}^2} \tag{1}$$

where A_{biq}, q_{biq}, and $\omega_{c,biq}$ are gain at center frequency, quality factor, and center frequency, respectively.

$$A_{biqp} = \frac{\omega_u R_1 R_2 C_1}{R_1 + R_2 - \omega_u R_1 R_2 C_2} \tag{2}$$

$$q_{biqp} = \frac{R_1\sqrt{\omega_u R_2 (C_1 + C_2 + C_3)}}{R_1 + R_2 - \omega_u R_1 R_2 C_2} \tag{3}$$

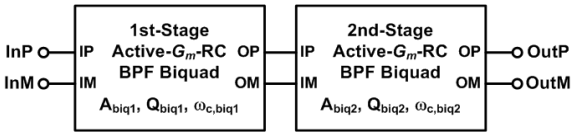


Fig. 2. Block diagram of a 4th-order active gm-RC BPF.

$$\omega_{c,biqp}^2 = \frac{\omega_u}{R_2(C_1 + C_2 + C_3)} \tag{4}$$

From the above equations, it is clarified that the quality factor q_{biqp} is enhanced according to the increase of C_2, and all transfer function parameters can be individually optimized since A_{biqp} and $\omega_{c,biqp}$ are independent of C_3 and R_1, respectively. A_{biqp} and q_{biqp} become process independent with a G_m adjusting circuit as is described in [3][4]; while circuit parameters related to $\omega_{c,biqp}$ are optimized with the ω_c and BW tuning scheme.

III. 4TH-ORDER ACTIVE- G_M-RC BPF DESIGN

As in the case of typical filter design, a higher-order active-G_m-RC BPF can be implemented with a series connection of biquads. In this subsection, we illustrate the design of higher-order active-G_m-RC BPFs with an example of a 4th-order BPF composed of the cascaded biquad structure as shown in Fig. 2. The transfer function of the total 4th-order BPF, $H(s)_{BPFtot}$, is given by

$$H(s)_{BPFtot} = H_{BPFbp1}(s) \cdot H_{BPFbp2}(s)$$

$$= \frac{A_{biqp1} \cdot \frac{\omega_{c,biqp1}}{q_{biqp1}} s}{s^2 + \frac{\omega_{c,biq1}}{q_{biqp1}} s + \omega_{c,biq1}^2} \cdot \frac{A_{biqp2} \cdot \frac{\omega_{c,biq2}}{q_{biqp2}} s}{s^2 + \frac{\omega_{c,biq2}}{q_{biqp2}} s + \omega_{c,biq2}^2} \tag{5}$$

where $H_{BPFbq1}(s)$ and $H_{BPFbq2}(s)$ are transfer functions of the proposed 1st-stage and 2nd-stage active-G_m-RC BPF biquad, respectively, and associated parameters, A_{biqpj}, q_{biqpj}, and $\omega_{c,biqpj}$, are identical to those of the proposed biquad represented from (2) to (4) where j indicates a stage number in the biquad chain.

The total gain at center frequency A_{tot}, bandwidth BW, and the center frequency ω_c of the 4th-order BPF are consequently given by

$$A_{tot} = \frac{A_{biqp1} A_{biqp2}}{q_c^2} \cdot \left(\frac{\omega_c}{BW} \right)^2 \tag{6}$$

$$BW = \omega_{c,biqp2} - \omega_{c,biqp1} \tag{7}$$

$$\omega_c = \sqrt{\omega_{c,biqp1} \omega_{c,biqp2}} \tag{8}$$

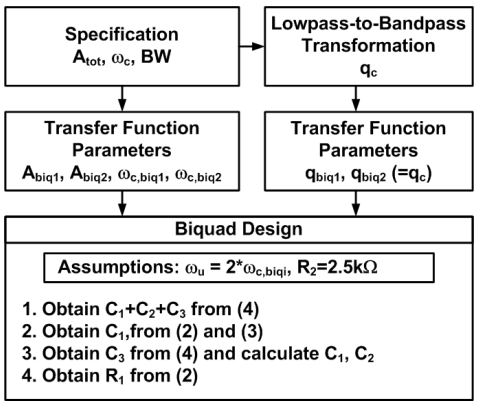


Fig. 3. Entire block diagram of the proposed 4-order ative-gm-RC BPF with automatic analog/digital tuning circuit.

TABLE I
Summary of target specification, parameters

Target Specification		
A_{tot}	1.0	
ω_c [Mrad/sec]	377 (60MHz)	
BW [Mrad/sec]	50.3 (8MHz)	
q_c	10.6	
Transfer Function Parameters	1st-Stage	2nd-Stage
A_{biqj}	1.42	1.42
q_{biqj}	10.6	10.6
$\omega_{c,biqj}$ [Mrad/sec]	359	395
Circuit Parameters	1st-Stage	2nd-Stage
R_1 [kΩ]	2.1	2.1
R_2 [kΩ]	2.5	2.5
C_1 [pF]	0.15	0.13
C_2 [pF]	1.06	1.06
C_3 [pF]	1.21	0.66

$$\omega_{c,biqpj} = \sqrt{\frac{\omega_{uj}}{R_{2j}\left(C_{1j} + C_{2j} + C_{3j}\right)}} \tag{9}$$

where q_c in (6) is the common quality factor set equal to q_{biqp1} and q_{biqp2}. The quality factor of the entire BPF, Q_0, corresponds to ω_c/BW, and the relationship between q_c and Q_0 can be obtained from the lowpass-to-bandpass transformation in filter design theory [5], and is given by

$$q_c^2 = \frac{Q_0}{a}\left[\left(\frac{2Q_0}{a} + \frac{b}{2aQ_0} \right) + \sqrt{\left(\frac{2Q_0}{a} + \frac{b}{2aQ_0} \right)^2 - 1} \right] \tag{10}$$

where a and b are normalized coefficients of a prototype LPF, and its transfer function is represented as

$$H_{LPFp}(s) = \frac{1}{s^2 + as + b} \tag{11}$$

For the Butterwoth approximation, a and b are set to 1.41 and 1, respectively. From the equation (6), it is clarified that the total gain at center frequency A_{tot} is not a

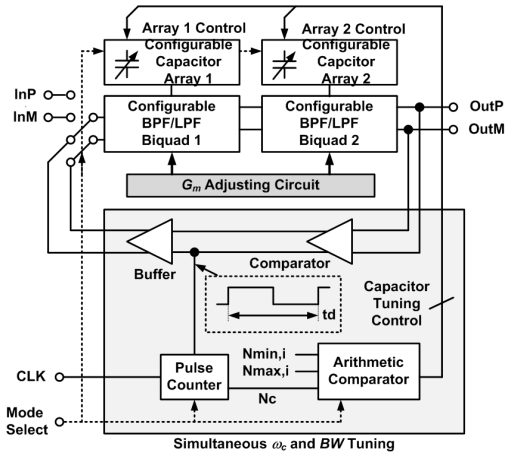


Fig. 4. Entire block diagram of the proposed 4-order ative-gm-RC BPF with automatic analog/digital tuning circuit.

simple product of A_{biqp1} and A_{biqp2}, but is a function of q_c, BW, and ω_c.

In consequence of the described analyses, the design flow of 4th-order active-G_m-RC BPF is proposed as is presented in Fig. 3. The target specification of the on-chip equivalent SAW filter application as well as biquad transfer function and circuit parameters derived from the proposed design flow are summarized in Table 1.

IV. Combined Analog-Digital Tuning System

The entire block diagram of the proposed 4th-order active-G_m-RC BPF featuring a combined analog-digital tuning circuits is shown in Fig. 4. The filter core consists of two configurable BPF/LPF biquads in order to alternate modes for filter or tuning operation, and the tuning of 4th-order BPF characteristics, represented with A_{tot}, ω_c, and BW, can be achieved through the tuning of biquad transfer function parameters indicated from (6) to (8). The analog tuning block is a G_m adjusting circuit shared by each biquad. This circuit acts based on the constant-G_m biasing technique, and eliminates the process dependence of A_{biqj} and q_{biqpi} as is illustrated in [3][4].

With the use of G_m adjusting circuit, the transfer function of OTA becomes

$$A(s) \approx \frac{\omega_u}{s} = \frac{G_m}{sC_L} = \frac{1}{sk_g R_b C_L} \tag{12}$$

The process variation of filter characteristics can be evaluated by replacing the resistor R_i and capacitor C_i in the biquad transfer function parameters with sheet resistance, R_s, unit capacitance, C_u, and an associated geometric factors, α_i and β_i. The most significant feature introduced by the G_m adjusting circuit is that $\omega_u R_2 C_2$ commonly used in (2) and (3) becomes independent from PVT variations, that is represented with

$$\omega_u R_2 C_2 = \frac{R_2 C_2}{sk_g R_b C_L} = \frac{\alpha_2 \beta_2}{sk_g R_b \beta_L} = K \tag{13}$$

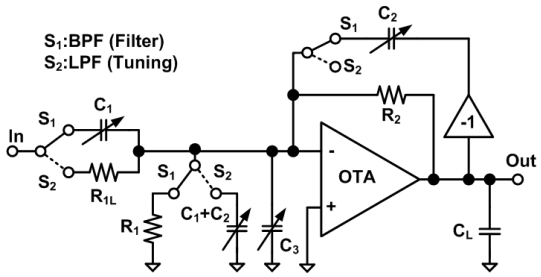


Fig. 5. Proposed configurable active-G_m-RC BPF/LPF biquad.

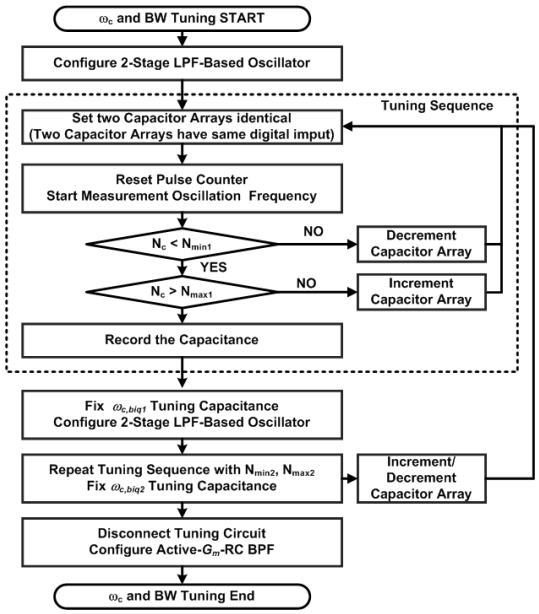


Fig. 6. Flowchart of ω_c and BW Tuning.

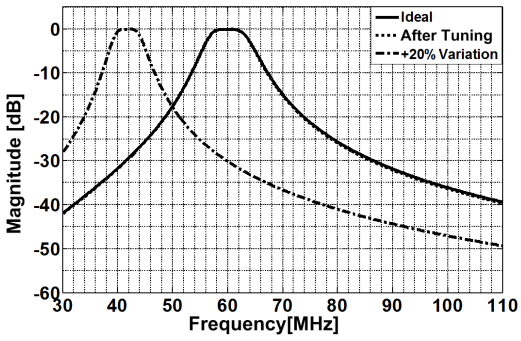
As a result, transfer function parameters of the proposed active-G_m-RC BPF biquad, A_{biqp} and q_{biqp}, are given by

$$A_{biqp} = K\left(\frac{\beta_1}{\beta_2}\right)\Big/\left(1+\frac{\alpha_2}{\alpha_1}-K\right) \tag{14}$$

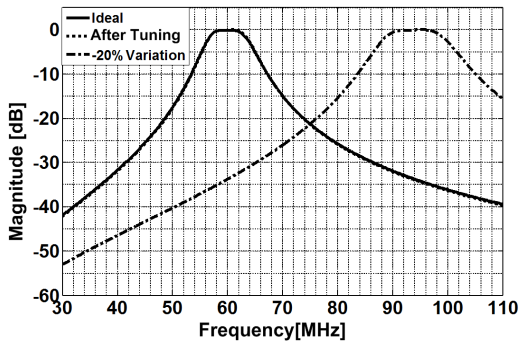
$$q_{biqp} = \sqrt{K\left(\frac{\beta_1+\beta_2+\beta_3}{\beta_2}\right)}\Big/\left(1+\frac{\alpha_1}{\alpha_2}-K\right) \tag{15}$$

From these equations it is concluded that biquad transfer function parameters, A_{biqp} and q_{biqp}, are functions of process independent geometric ratio, and have immunity to PVT variations thanks to the G_m adjusting circuit.

The tuning parameters related to ω_c and BW are $\omega_{c,biq1}$ and $\omega_{c,biq2}$, that are sequentially tuned with the digital-based tuning circuits organized with a configurable active-G_m-RC BPF/LPF biquad and the associated frequency detection circuit as shown in Fig. 4. The output frequency of the LPF-based oscillator is ideally defined as the equivalent frequency that the phase shift of the LPF becomes 180° in its frequency response. From the analysis of the LPF transfer function, the phase shift of two identical biquads $\Delta\theta$ is obtained as



(a)



(b)

Fig. 7. Simulated waveform of active-G_m-RC BPF .

$$\Delta\theta = -2\tan^{-1}\left(\frac{\omega\omega_{LPF}/q_{LPF}}{\omega^2-\omega_{LPF}^2}\right) \tag{16}$$

where q_{LPF} and ω_{LPF} are quality factor and pole frequency of the configured active-G_m-RC LPF. This equation indicates that the oscillation occurs at the LPF pole frequency ω_{LPF} with no regard to the quality factor q_{LPF}.

The resulted schematic diagram of the configurable active-G_m-RC BPF/LPF biquad is presented in Fig. 5 with an equivalent single-ended form expression. With the LPF-based oscillator, $\omega_{c,biqj}$ is tuned within a specified error according to the flowchart shown in Fig. 6.

The target specification of the on-chip equivalent SAW filter application as well as biquad transfer function and circuit parameters derived from the proposed design flow are summarized in Table I. The values of $\omega_{c,biqpj}$ needs to be slightly modified by taking account of lowpass-to-bandpass transformation[5].

In the tuning mode, the capacitor network is first set to the midpoint value to minimize the tuning time. After the configuration from the active-G_m-RC BPF to LPF is completed, the pulse counter starts to measure the oscillation frequency. If the counted pulse number N_c is smaller than N_{minj}, the capacitor value is decreased to increase the oscillation frequency; on the other hand, if the N_c is larger than N_{maxj}, the capacitor value is increased in the same manner. The digital block shown in Fig. 3 is fairly simple, the number of logic gates necessary for the counter, arithmetic comparator, and control block is less than 1000 [6].

V. Simulation Result

Behavior level simulation results on the frequency response of the 4th-order active-G_m-RC BPF are presented Fig. 7, where in both plots the ideal response without process variation is presented as a reference. Fig. 7 (a) presents responses before and after tuning for the case of +20% process variation in both resistors and capacitors; while Fig. 7 (b) presents the case for -20% process variation. Since the difference between the reference and after tuning result, caused by the capacitor only tuning, is negligible, we can conclude that unacceptable response before tuning can be properly calibrated to the reference response.

Estimated power consumptions with regard to input referred noise voltage are compared between active-RC and active-G_m-RC BPF biquads. The active-RC BPF biquad consumes 23.5mW with the OTA unity-gain frequency of around 2π*3GHz and the equivalent input referred noise of 450μV [1]. For the active-G_m- RC BPF biquad, power consumption is derived as 3.0mW with 120MHz unity-gain bandwidth and 120μV input referred noise as is described in the accompanied implementation paper. The resulted power reduction factor in active-G_m-RC over active-RC BPF biquad for the equivalent input referred noise condition is calculated as 18.8.

Acknowledgement

The authors wish to acknowledge the assistance and support of VLSI Design and Education Center (VDEC), The University of Tokyo with the collaboration with Synopsys Corporation.

References

[1] T.-Y. Lo, C.-C. Hsiao, K.-W. Hsueh and H.-S. Li, "A 1-V 60MHz bandpass filter with quality-factor calibration in 65nm CMOS," Proc. of IEEE Asian Solid-Sate Circuit Conference, pp. 53-56, 2009.

[2] B. Nauta, "A CMOS transconductance-C filter technique for very high frequencies", IEEE J. Solid-State Circuits, vol.27, no.2, pp.142-153, February 1992.

[3] V. Giannini, J. Craninckx and A. Baschirotto, Baseband Analog Circuits for Software Defined Radio, Springer Dordrecht, 2008.

[4] S. D'Amico, V. Giannini and A. Baschirotto, "A 4th-Order active-G_m-RC reconfigurable (UMTS/ WLAN) filter," IEEE J. Solid-State Circuits, vol.41, no.7, pp. 1630-1637, July 2006.

[5] R. Schaumann and M. Van Valkenburg, Design of Analog Filters, Oxford University Press New York, 2001.

[6] H. Kondo, M. Sawada, N. Murakami and S. Masui, "Design of complex BPF with automatic digital tuning circuit for low-IF receivers," IEICE Trans. Electron., vol. E92-C, no. 10, pp. 1304-1310, October 2009.

Dimension Reduction Based Direction of Arrive Estimation with an Uncalibrated Distributed Electromagnetic Component Sensor Array

Hui Chen, Qun Wan
Department of Electrical Engineering
UESTC
Chengdu, China
chenhui123@uestc.edu.cn

Yulin Liu
DSP Lab
Chongqing Communication College
Chongqin, China

***Abstract*—In this paper, we propose a dimension reduction based direction of arrival (DOA) estimation approach using an array of spatially distributed electric and magnetic component sensors, and the array comprises sensors that are uncalibrated with each other. The proposed method is compared with relevant Cramer Rao bound (CRB) and its computational complexity is analyzed. It is shown that the dimension reduction based DOA method can reduce the computational loads significantly, and enhance estimation accuracy for low signal noise ratio (SNR) condition. Computer simulations are conducted to illustrate the effectivity of the proposed estimators.**

Keywords-Direction of arrival estimation; Distributed Electromagnetic Component Sensor Array; Dimension reduction

I. INTRODUCTION

Most existing direction of arrival (DOA) estimation methods employ scalar sensor arrays in which the output of each sensor is a scalar corresponding to, for example, the pressure in the acoustic case or a scalar function of the electric filed in the electromagnetic case. Whereas the main advantage of the vector sensor array is that it makes use of all available electromagnetic information and should outperform the scalar sensor arrays in accuracy of DOA estimation. So the problem of estimating electromagnetic wave parameters using antenna arrays with diversely polarized elements is important in many applications. In addition to DOA [1], using the polarization information enables to improve the performance of active sensing systems such as radar [2-3] and increases the capacity of communication systems. In radar, polarimetric scattering information is useful for discriminating the targets' features such as geometrical structure, shape, orientation and so on. Besides, distributed electromagnetic component sensor array (DEMCA)'s structure enables it to have many advantages [14]. And we know the calibration of the whole array may be poor due to completely unknown or imprecisely known inter-sensor displacements, imperfect time synchronization of different sensors, unknown channel mismatches between some sensors that are located far away from each other, or a combination of the above mentioned effects. So the problem of DOA estimation using uncalibrated sensor array is important in practical applications.

This work was supported in part by the National Natural Science Foundation of China under grant 60772146, the Key Project of Chinese Ministry of Education under grant 109139, and by the Special Fund of Central Colleges Basic Scientific Research Operating Expenses under grant E022050205.

The conventional subspace algorithm such as multiple signal classification (MUSIC) cannot be used in the above mentioned situations as they are very sensitive even to very small array manifold model errors [4]. Normally, full offline calibration of the whole array is an existing solution to the above mentioned problem, which maybe an extremely complicated task [5]. Several self-calibration of solutions have been also proposed to enable joint calibration of the array and estimation of the source DOAs [6–10], but their computational complexity are very high, and performance will be severely degenerated in the condition of large sensor position errors [10]. The authors [11, 12] proposed a rank reduction estimator (RARE) method, and Chong-Meng Samson See and Arye Nehorai extend the application of RARE from a scalar sensor array to DEMCA [13], whereas it suffers high computation complexity.

In this paper, we propose a dimension reduction based subspace approach to obtain better estimation performance of multiple sources using an uncalibrated DEMCA. Spectral cost function is addressed using subspace orthogonal firstly. Then we introduce a simplified 2-dimension spectral cost function by combining multidimensional parameter, which reduces the computational loads significantly without degenerating performance. In addition, to verify the performance of our proposed method, Cramer Rao bound (CRB) is computed and compared with ours. Simulation results illustrate the validity of the proposed method finally.

The paper is organized as follows. In Section II, the measurement model of the DEMCA is introduced. Section III presents our proposed method. Then, computer simulations are conducted, and the results are given in Section IV. Finally, Section V gives some concluding remarks for the proposed method.

II. MEASUREMNET MODEL

The measurement model of a compact 6-dimension electromagnetic vector sensor (EMVS) is given by [1],

$$\begin{bmatrix}\mathbf{y}_E(t)\\ \mathbf{y}_H(t)\end{bmatrix}=\begin{bmatrix}\boldsymbol{I}_3\\ \mathbf{U}\end{bmatrix}\mathbf{VQw}s(t)+\begin{bmatrix}\mathbf{e}_E(t)\\ \mathbf{e}_H(t)\end{bmatrix}, \qquad (1)$$

where $\boldsymbol{I}_3$ is the third order identity matrix, and

$$\mathbf{U} \triangleq \begin{bmatrix} 0 & -\mathrm{u}_3 & \mathrm{u}_2 \\ \mathrm{u}_3 & 0 & -\mathrm{u}_1 \\ -\mathrm{u}_2 & \mathrm{u}_1 & 0 \end{bmatrix}. \tag{2}$$

The vector $\mathbf{u} = [\mathrm{u}_1, \mathrm{u}_2, \mathrm{u}_3] = [\cos\theta\cos\phi, \sin\theta\cos\phi, \sin\phi]^T$ indicates the unit direction vector from sensor to source, the matrix $\mathbf{V}$ is given by

$$\mathbf{V} = \begin{bmatrix} -\sin\theta & -\cos\theta\sin\phi \\ \cos\theta & -\sin\theta\sin\phi \\ 0 & \cos\phi \end{bmatrix}, \tag{3}$$

the wave polarization is defined by

$$\mathbf{Q} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \tag{4}$$

and

$$\mathbf{w} = \begin{bmatrix} \cos\beta \\ j\sin\beta \end{bmatrix}. \tag{5}$$

where θ, ϕ, α and β are the azimuth, elevation, polarized ellipse's orientation and eccentricity angle respectively.

Assuming that the signal sources are narrowband, the measurement model of the DEMCA in multiple sources environment becomes [14]

$$\underbrace{\begin{bmatrix} \mathbf{y}_E(t) \\ \mathbf{y}_H(t) \end{bmatrix}}_{\mathbf{y}(t)} = \sum_{k=1}^{K} \mathbf{a}(\boldsymbol{\theta}^{(k)}) s_k(t) + \underbrace{\begin{bmatrix} \mathbf{e}_E(t) \\ \mathbf{e}_H(t) \end{bmatrix}}_{\mathbf{n}(t)}, \tag{6}$$

$$\mathbf{a}(\boldsymbol{\theta}^{(k)}) = \underbrace{\boldsymbol{\Gamma}(\theta^k, \phi^k)\boldsymbol{\Omega}\begin{bmatrix} \boldsymbol{I}_3 \\ \mathbf{U}_k \end{bmatrix}\mathbf{V}_k}_{\mathbf{b}(\theta^k, \phi^k)} \underbrace{\mathbf{Q}_k\mathbf{w}_k}_{\boldsymbol{F}(\mathbf{Q}_k\mathrm{w}_k)} = \mathbf{b}(\theta^k, \phi^k)\boldsymbol{F}(\mathbf{Q}_k\mathrm{w}_k) \ , \tag{7}$$

where $\boldsymbol{\theta}^{(k)} = (\theta^k, \phi^k, \alpha^{(k)}, \beta^{(k)})$ denotes the direction and polarization parameters of the k-th source signal. $\boldsymbol{\Gamma}(\theta, \phi)$ is an $N \times N$ (N is the number of component sensors) diagonal matrix whose n-th diagonal entry is given by $\left[\boldsymbol{\Gamma}(\theta, \phi)\right]_{nn} = e^{j2\pi \boldsymbol{q}_n^T \mathbf{u}/\lambda}$, for $n = 1, ..., N$. This matrix provides the phase shift between the vector sensor centre and position $\boldsymbol{q}_n$ of the n-th element of the vector sensor. $\boldsymbol{\Omega}$ is an $N \times 6$ selection matrix with elements of 1 or 0 to pick out a choice from 6 components of the electromagnetic vector sensor. For example, $\boldsymbol{\Omega} = \boldsymbol{I}_6$ when DEMCA is placed as shown in Fig. 1.

Unlike the compact EMVS, DEMCA generalizes the vector sensor array and allows the differential delay measurements resulting from diverse placement of the component sensors and electromagnetic field measurements to be jointly exploited in estimating the source parameters. Given both the complete electromagnetic and spatial information, better parameter estimation with a smaller aperture array can be expected over a wide frequency range as compared to either a vector sensor or a scalar array.

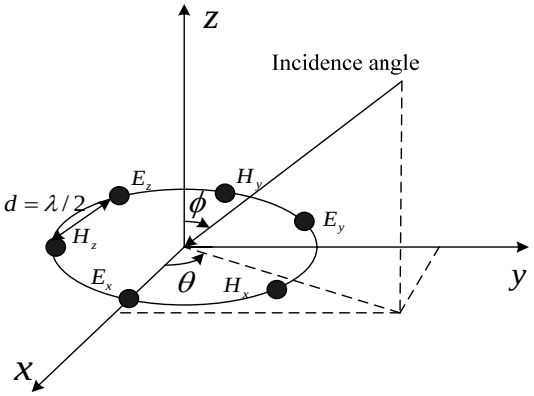


Fig.1. Distributed electromagnetic component sensor array.

In the following, the operators $(\cdot)^H$, $(\cdot)^T$ and $\mathrm{E}\{\cdot\}$ denote the Hermitian transpose, transpose, and statistical expectation respectively. The symbol $\mathrm{diag}\{z_1, z_2\}$ represents a diagonal matrix with diagonal entries z_1, z_2.

Considering the array error, we can model array snapshot compactly in matrix form with $\mathbf{a}_F(\boldsymbol{\theta})$,

$$\mathbf{Y}(t) = \mathbf{A}(\boldsymbol{\Theta})\mathbf{s}(t) + \mathbf{e}(t), \ \ t = 1, ..., T \ , \tag{8}$$

where $\mathbf{A}(\Theta) = [\mathbf{a}_F(\boldsymbol{\theta}^{(1)}), \mathbf{a}_F(\boldsymbol{\theta}^{(2)}), ..., \mathbf{a}_F(\boldsymbol{\theta}^{(K)})]$, $\mathbf{a}_F(\boldsymbol{\theta}^{(i)}) = \mathbf{C}\mathbf{a}(\boldsymbol{\theta}^{(i)})$ $(i = 1, ..., K)$, diagonal matrix $\mathbf{C} = \mathrm{diag}(\mathbf{C}_{1,1}, \mathbf{C}_{2,2}, \mathbf{C}_{3,3}, \mathbf{C}_{4,4}, \mathbf{C}_{5,5}, \mathbf{C}_{6,6})$, which contains all types of error information, $\boldsymbol{\Theta} = [\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, ..., \boldsymbol{\theta}^{(K)}]$, $\mathbf{s}(t) = [s_1(t), s_2(t), ..., s_K(t)]^T$.

With $\mathbf{Y}(t)$, the sample estimate of the array covariance matrix

$$\boldsymbol{R} \triangleq \mathrm{E}\{\mathbf{Y}(t)\mathbf{Y}^H(t)\} = \mathbf{A}(\Theta)\boldsymbol{R}_s\mathbf{A}(\Theta)^H + \sigma^2\boldsymbol{I} \tag{9}$$

is given by

$$\hat{\boldsymbol{R}} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{Y}(t)\mathbf{Y}^H(t) \ . \tag{10}$$

where $\boldsymbol{R}_s \triangleq \mathrm{E}\{\mathrm{s}(t)\mathrm{s}^H(t)\}$ is the source covariance matrix, σ^2 is the identical variance in each sensor.

When $K < 6$, the eigen-decomposition of the matrices (9) and (10) can be expressed in the form

$$\boldsymbol{R} = \boldsymbol{E}_s\boldsymbol{\Lambda}_s\boldsymbol{E}_s^H + \boldsymbol{E}_N\boldsymbol{\Lambda}_N\boldsymbol{E}_N^H \ , \tag{11}$$

$$\hat{\boldsymbol{R}} = \hat{\boldsymbol{E}}_s\hat{\boldsymbol{\Lambda}}_s\hat{\boldsymbol{E}}_s^H + \hat{\boldsymbol{E}}_N\hat{\boldsymbol{\Lambda}}_N\hat{\boldsymbol{E}}_N^H \ . \tag{12}$$

where the $K \times K$ diagonal matrices $\boldsymbol{\Lambda}_s$ and $\hat{\boldsymbol{\Lambda}}_s$ contain the K signal subspace eigenvalues of $\boldsymbol{R}$ and $\hat{\boldsymbol{R}}$ respectively, and the

$(6-K)\times(6-K)$ diagonal matrices $\boldsymbol{\Lambda}_N$ and $\hat{\boldsymbol{\Lambda}}_N$ contain the $6-K$ noise subspace eigenvalues of $\boldsymbol{R}$ and $\hat{\boldsymbol{R}}$ respectively. In turn, $\boldsymbol{E}_s$ and $\hat{\boldsymbol{E}}_s$ are $6\times K$ matrices whose columns are the signal subspace eigenvectors corresponding to the K largest eigenvalues of $\boldsymbol{R}$ and $\hat{\boldsymbol{R}}$ respectively, while $\boldsymbol{E}_N$ and $\hat{\boldsymbol{E}}_N$ are $6\times(6-K)$ matrices whose columns are the noise subspace eigenvectors corresponding to the $6-K$ smallest eigenvalues of $\boldsymbol{R}$ and $\hat{\boldsymbol{R}}$ respectively.

III. Direction of Arrive Estimation

Now we consider the well known spectral MUSIC algorithm which estimates the signal DOAs from the K highest peaks of the following function [15]

$$f(\boldsymbol{\theta},\mathbf{C})=\frac{1}{\mathbf{a}_F^H(\boldsymbol{\theta},\mathbf{C})\hat{E}_N\hat{E}_N^H\mathbf{a}_F(\boldsymbol{\theta},\mathbf{C})}. \tag{13}$$

In the idea case of exactly known $\boldsymbol{R}$, the DOAs can be found from the equation

$$\mathbf{a}_F^H(\boldsymbol{\theta},\mathbf{C})\hat{E}_N\hat{E}_N^H\mathbf{a}_F(\boldsymbol{\theta},\mathbf{C})=0. \tag{14}$$

To find the K highest peaks of (13), we have to use an exhaustive multidimensional search with respect to θ, ϕ, α and β, which becomes totally impractical. To overcome this problem, we can rewrite the equation (13) as

$$\begin{aligned} f(\boldsymbol{\theta},\mathbf{h}) &= \frac{1}{\left\|\mathbf{a}_F^H(\boldsymbol{\theta})\boldsymbol{E}_N\right\|^2}=\frac{1}{\left\|\boldsymbol{E}_N^H\mathbf{a}_F(\boldsymbol{\theta})\right\|^2} \\ &= \frac{1}{\left\|\boldsymbol{E}_N^H\mathbf{b}(\theta,\phi)\underbrace{\breve{\boldsymbol{C}}\breve{\boldsymbol{F}}(\mathbf{Qw})\mathbf{1}_6}_{\substack{\text{polarization and}\\ \text{error information}}}\right\|^2} \\ &= \frac{1}{\left[\underbrace{\breve{\boldsymbol{C}}\breve{\boldsymbol{F}}(\mathbf{Qw})\mathbf{1}_6}_{\substack{\text{polarization and}\\ \text{error information}}}\right]^H\mathbf{b}^H(\theta,\phi)\boldsymbol{E}_N\boldsymbol{E}_N^H\mathbf{b}(\theta,\phi)\left[\underbrace{\breve{\boldsymbol{C}}\breve{\boldsymbol{F}}(\mathbf{Qw})\mathbf{1}_6}_{\substack{\text{polarization and}\\ \text{error information}}}\right]} \\ &= \frac{1}{\boldsymbol{g}^{\boldsymbol{H}}\boldsymbol{C}\,\boldsymbol{g}} \end{aligned} \tag{15}$$

where $\breve{\boldsymbol{C}}^{2\times12}=\begin{bmatrix}\mathbf{C}_{1,1} & 0 & \ddots & 0 & \mathbf{C}_{6,6} & 0\\ 0 & \mathbf{C}_{1,1} & 0 & \ddots & 0 & \mathbf{C}_{6,6}\end{bmatrix}$, block diagonal matrix

$$\breve{\boldsymbol{F}}(\mathbf{Qw})^{\in12\times6}=\begin{bmatrix}(\boldsymbol{F}(\mathbf{Qw}))_{1,1} & 0 & 0\\ (\boldsymbol{F}(\mathbf{Qw}))_{2,1} & 0 & 0\\ 0 & \vdots & 0\\ 0 & \vdots & 0\\ 0 & 0 & (\boldsymbol{F}(\mathbf{Qw}))_{1,1}\\ 0 & 0 & (\boldsymbol{F}(\mathbf{Qw}))_{2,1}\end{bmatrix},$$ and

$\mathbf{1}_6=[1,1,1,1,1,1]^T$, $\boldsymbol{g}=\breve{\boldsymbol{C}}\breve{\boldsymbol{F}}(\mathbf{Qw})\mathbf{1}_6$ and $\boldsymbol{\Psi}=\mathbf{b}^H(\theta,\phi)\boldsymbol{E}_N\boldsymbol{E}_N^H\mathbf{b}(\theta,\phi)$. (15) obviously shows that the polarization information and error information is contained in vector parameter $\boldsymbol{g}$ only, so the matrix $\boldsymbol{\Psi}$ is independent of polarization and error information. Besides, the sample matrix $\hat{\boldsymbol{\Psi}}=\mathbf{b}^H(\theta,\phi)\hat{\boldsymbol{E}}_N\hat{\boldsymbol{E}}_N^H\mathbf{b}(\theta,\phi)$ is used instead of $\boldsymbol{\Psi}$ in practice. Therefore, the signal DOAs can be found from the K highest peaks of the following spectral function:

$$f_1(\boldsymbol{\theta},\mathbf{C})=\frac{1}{\lambda_{\min}\{\hat{\boldsymbol{\Psi}}\}}, \tag{16}$$

where $\lambda_{\min}\{\cdot\}$ is the operator that returns the smallest eigenvalue of a Hermitian matrix.

We still can analysis this problem in aspect of reduction of rank [12], the determinant of $\hat{\boldsymbol{\Psi}}$ as well as its smallest eigenvalue will tend to have a minimum when $\boldsymbol{\theta}$ coincides with one of the signal directions $\{\boldsymbol{\theta}_l\}_{k=1}^K$. Therefore another alternative spectral function can be used

$$f_2(\boldsymbol{\theta},\mathbf{C})=\frac{1}{\det\{\hat{\boldsymbol{\Psi}}\}}. \tag{17}$$

Besides, due to space limited for this article, the uniqueness of the signal DOA estimates obtained from the reduction criterion (16) in the case $T\to\infty$ and the CRB derivation expression will be omitted.

IV. Simulation results and Discussions

In this section, we present experimental results for our proposed scheme using an uncalibrated DEMCA. First we show our simulation results by plotting spatial spectra of signal sources, and then we compare the performance obtained by our approach with relevant CRB under different signal noise ratio (SNR). Throughout our simulations, we assume the signal sources impinging on an uncalibrated DEMCA, as shown Fig. 1, is made of x, y and z electric and magnetic component sensors arranged as a uniformly spaced circular array of half wavelength inter-sensor spacing. And $T=100$ independent snapshots are used to estimate the array covariance matrix in each simulation example. For simplicity, we assume that all signal sources are of equal power σ_S^2, and input SNR is defined as $10\log_{10}(\sigma_S^2/\sigma_n^2)$.

In the first simulation, we suppose two uncorrelated sources impinging on the uncalibrated DEMCA from the DOAs $\boldsymbol{\Theta}^{(1)}=[20^\circ,40^\circ,45^\circ,-5^\circ]$ and $\boldsymbol{\Theta}^{(2)}=[50^\circ,-30^\circ,60^\circ,-60^\circ]$, the array error matrix $\mathbf{C}$ = diag(1.05 + j*0.05,1.02

$+j*0.01, 1.02+j*0.05, 1.02+j*0.03, 0.98, 1.01+j*0.03)$, and the SNR is fixed at 30dB. Our proposed method determines the DOA accurately as indicted by the contour plot in Fig. 2, and reduces the computational complexity without considering polarization parameters.

To compare the performance obtained by our approach with CRB under different SNR, we define the root mean square error (RMSE) of the DOA estimates from 100 Monte Carlo trials as

$$\mathrm{RMSE}=\sqrt{\sum_{n=1}^{100}\sum_{k=1}^{N_s}((\hat{\theta}_k(n)-\theta_k)^2+(\hat{\phi}_k(n)-\phi_k)^2)\Big/(100N_s)}\ . \quad (18)$$

where $\hat{\theta}_k(n)$, $\hat{\phi}_k(n)$ are the estimation of θ_k and ϕ_k for the n th Monte Carlo trial respectively , and N_s is the number of all the signals. In the second simulation, we consider the same situation as that mentioned in the first simulation, and the results are averaged over 100 Monte Carlo simulations. Fig. 3 shows the RMSE of the DOA estimates versus input SNR. From the simulation results, it is clear that the RMSE estimators of ours signals, except low SNR, can meet CRB.

V. CONCLUSION

In this paper, an efficient dimension reduction based DOA estimation approach using an uncalibrated DEMCA has been proposed, which is also straight suit for uncalibrated vector array and multiple DEMCAs. And the example with combinative effect of various types of imperfections has been presented, which illustrates the proposed method has significant accuracy and low computational complexity advantage.

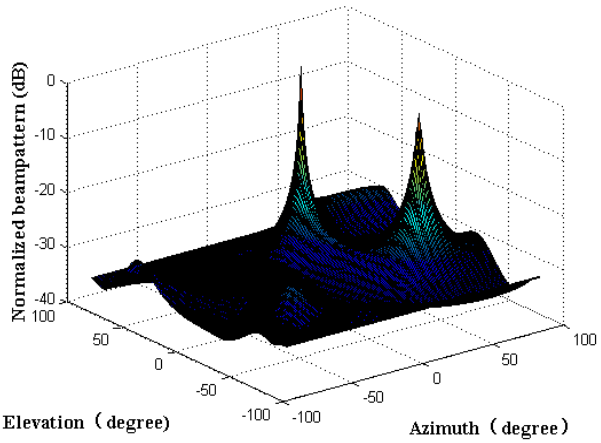


Fig. 2. Normalized beampattern using our proposed approach (16).

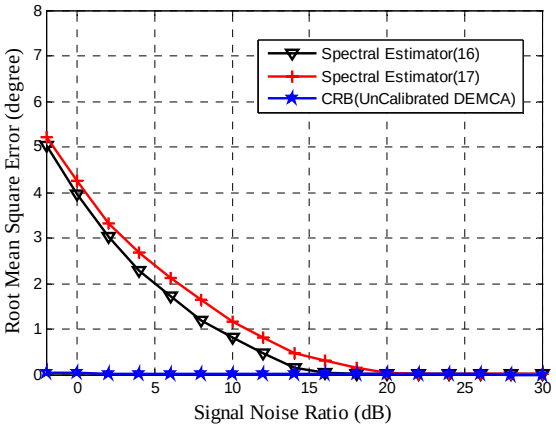


Fig. 3. RMSE of the DOA estimates versus input SNR.

REFERENCES

[1] A. Nehorai, E. Paldi, "Vector Sensor Array Processing for Electromagnetic Source Localization," IEEE Trans. Signal Process. Vol.42, No.2, pp.376-398, February 1994.

[2] S. M. Sowelam and A. H. Tewfik, "Waveform Selection in Radar Target Classification," IEEE Trans. Inf. Theory, Vol. 46, No.3, pp. 1014-1029, May 2000.

[3] M. Hurtado, J. J. Xiao and A. Nehorai, "Target Estimation, Detection, and Tracking: A Look at Adaptive Polarimetric Design," IEEE Signal Processing Magazine, Vol. 26, No.1, pp.42-52, January 2009.

[4] B. Friedlander, "A Sensitivity Analysis of the MUSIC Algorithm," IEEE Trans. Acoustics, Speech, and Signal Process, Vol. 38, No. 10, pp. 1740–1751, October 1990.

[5] B. Porat, B. Friedlander, "Accuracy Requirements in off-line Array Calibration," IEEE Trans. Aerospace and Electronic Systems, Vol. 33, No. 2, pp. 545–556, April 1997.

[6] Y. Rockah, and P. M. Schultheiss, "Array Shape Calibration Using Sources in Unknown Locations-Part 1: Far-field Sources," IEEE Trans. Acoustics, Speech, and Signal Process, Vol. 35, No. 3, pp. 286–299, March 1987.

[7] A. J.Weiss, B. Friedlander, "Self-calibration in High Resolution Array Processing," in Advances in Spectrum Estimation and Array Processing, S. Haykin, Ed. Englewood Cliffs: Prentice-Hall, Vol. II, 1991.

[8] B. C. Ng, C. M. S. See, "Sensor-array Calibration Using a Maximum-likelihood Approach," IEEE Trans. Antennas and Propaga, Vol. 44, No. 6, pp. 827–835, June 1996.

[9] M. Viberg, A. L. Swindlehurst, "A Bayesian Approach to Auto-calibration for Parametric Array Processing," IEEE Trans. Signal Process, Vol. 42, No. 12, pp. 3495–3507, December 1994.

[10] B. P. Flanagan, K. L. Bell, "Improved Array Self Calibration with Large Sensor Position Errors for Closely Spaced Sources," in Proc. Sensor Array Multichannel Signal Processing Workshop, Cambridge, MA, pp. 484–488, March 2000.

[11] M. Pesavento, A. B. Gershman, and K. M. Wong, "Direction Finding in Partly-calibrated Sensor Arrays Composed of Multiple Subarrays," IEEE Trans. Signal Process, Vol. 50, No. 9, pp.2103–2115, September 2002.

[12] C.M.S. See, A.B. Gershman, "Direction of Arrival Estimation in Partly Calibrated Subarray-based Sensor Arrays," IEEE Trans. Signal Process, Vol. 52, No.2, pp. 329–338, February 2004.

[13] C.M.S. See, A, Nehorai, "Source Localization with Partially Calibrated Distributed Electromagnetic Component Sensor Array," 2003 IEEE Workshop on Statistical Signal Processing, pp.458-461, 2003.

[14] C.M.S.See, A.Nehorai, "Distributed Electromagnetic Component Sensor Array," 7th Annual Adaptive Sensor Array Processing Workshop, March 1999.

[15] R.O.Schmidt, "Multiple Emitter Location and Signal Parameter Estimation," IEEE Trans. Antennas. Propaga, Vol. 34, No. 3, pp.276-280, 1986.

Enhancement of driver-dfficiency and the luminance uniformity in the light-emitting diode back-light units using RLC regulation and nubber circuits

Myouong-Sung Moon
Department of Electronic
LED-IT Fusion Technology Research Center
Gyeongbuk 712-749, KOREA

Ja-Soon Jang
Department of Electronic
LED-IT Fusion Technology Research Center
Gyeongbuk 712-749, KOREA
jsjang@ynu.ac.kr

Abstract— **In order to realize low power consumption, electrical system safety and the luminance uniformity of LED BLU (Back-Light Unit) system, new circuits with RLC (resistance-inductor-capacitance) regulation and snubber circuits have been proposed. From the simulation and experiment results based on our proposed model, the ripple power of the RLC regulation-embedded BLU circuit shows a dramatic reduction by more than 92 % as compared to the normal BLU and the overshoot voltage of the snubber circuit-embedded BLU circuit shows a dramatic reduction by more than 90 % as compared to the normal BLU, indicating that the proposed compensating-circuit is very promising for the realization of high-electrical efficiency and stable luminance characteristic BLU circuits.**

Keywords-Boost converter, Output ripple, Overshoot, RLC regualtion circuit, Snubber circuit

I. Introduction

Thin-Film-Transistor Liquid-Crystal-Displays (TFT-LCDs) are one of the most popular display devices scaling from small to large size. In particular, high-efficiency Back-Light Unit (BLU) circuits are of technological importance to operate TFT-LCDs having uniform luminance and local dimming control of the light-emitting diode (LED) BLU.

Very recently, 40-inch-scaled LED BLU TV with a side-emitting type has been successfully commercialized by SAMSUNG LED and the technological interest has been being increased so much. However, the side-emitting typed LED BLU is believed to be suitable up to 50-inch-scaled one, indicating that for more than 60-inch-scaled one, direct bottom-up-type-based BLUs are expected to use as a representative model because this can do a local dimming control. So as to realize high-performance LED BLU TV with larger than 60-inch length, high-efficiency BLU electric circuits are essential.

The block diagram of basic type boost converter is shown in Fig. 1. The boost converter is a high efficiency step-up DC/DC switching converter. The converter uses a transistor switch, typically a MOSFET, to pulse width modulate the voltage into an inductor. [2] The pulse-width modulation (PWM) boost converter that has been widely used as the driving circuit as a simple structure and its inductor current is equal to the input current. Accordingly, the input current is easy to control. [3] But by components of boost converter, boost converter is an unstable operation such as an output ripple and an overshoot phenomenon. In case of overshoot phenomenon, the power source of a system is under great stress when the system powers on. The operating mode of a Boost converter is usually divided into Continuous Conduction Mode (CCM) and Discontinuous Conduction Mode (DCM). In a boost converter, CCM is the preferred mode of operation for high-power applications when compared to the DCM, because CCM has lower conduction loss and smaller current stress on the semiconductor devices. However, the large reverse-recovery current of the output rectifier in a CCM boost converter causes power dissipation and interference problems. This happens at instant the switching elements is turned on, forcing the output rectifier to be suddenly reverse biased. The large reverse-recovery current spike of the rectifier also flows into switching elements. This cross conduction of current may cause excessive power dissipation and severe electromagnetic interference (EMI). In case of an output ripple, the operating mode of a Boost converter is usually divided into on and off operations such as a Fig. 2. Throughout a Fig. 3, we can know that the off operation of boost converter equate to a RLC circuit. By an inductance and a capacitance, impedance phase is happened and generate a ripple phenomenon. Many experiment results indicate that the output ripple voltage of a Boost converter in CCM increases with the decreasing of the inductance in the case of small inductance. For deceasing unstable electrical characteristic, simple methods are that capacitor or inductor values changed. But these methods have weakness such as overflow current by inrush current and slow response speed. [4] Also, the unstable luminance of LEDs is happened by unstable electrical characteristic, because the luminance of LEDs is related to the driving current. [5] Therefore to obtain sufficient electrical and optical characteristic, constant electrical properties must be supplied to each of the LEDs. [6] [7] [8] In this work, we

propose a model based on a RLC regulation circuit and a snubber circuit supported BLU circuits such as a Fig. 4. It was found that the ripple and overshoot phenomenon of the proposed system are significantly reduced as compared to the normal system. And we checked the relationship between upgrade electrical characteristics and luminance uniformity.

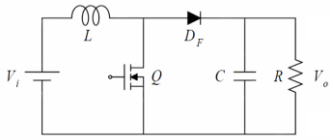
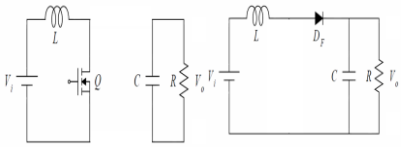


Fig. 1 Basic type boost converter



(a) Output voltage (b) Output current

Fig. 2 on and off operation of basic type boost converter

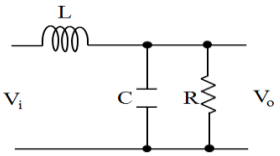


Fig. 3 Basic type boost converter

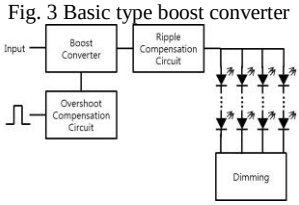
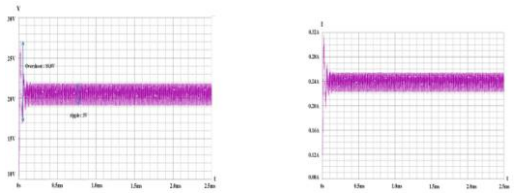


Fig. 4 Proposed LED driver

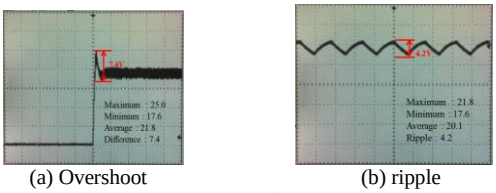
II. Operation principle

As we mentioned, a Boost converter is happened to unstable electrical and optical characteristics. Simulation and experiment results of basic type boost converter are shown in Fig. 5, 6 and 7. Throughout simulation and experiment results, we could know unstable electrical characteristic such as overshoot and ripple phenomenon of output and we could also know that the luminance of LEDs is not kept the uniformity by unstable electrical characteristic.



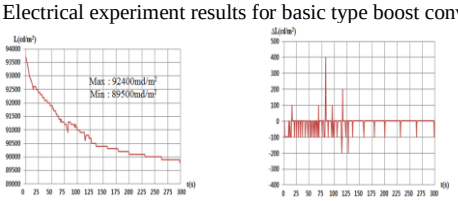
(a) Output voltage (b) Output current

Fig. 5 Simulation results for basic type boost converter



(a) Overshoot (b) ripple

Fig. 6 Electrical experiment results for basic type boost converter



(a) Intensity (b) Uniformity

Fig. 7 Optical experiment results for basic type boost converter

A. *RLC regulation circuit*

As we mentioned, the basic type converter is happened to output ripple. For decrease of output ripple phenomenon, we adopted the RLC circuit. The noise and unwanted ripple factors (which frequently happen in the analog circuits) can be removed by the use of RLC regulation circuits.

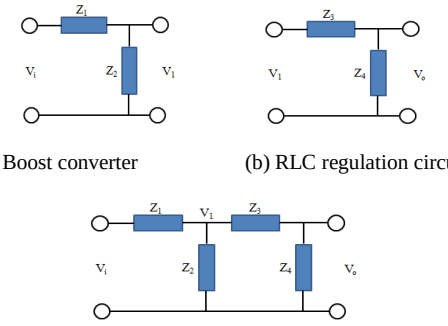
For input voltage and current of LEDs, we proposed a RLC regulation circuit. A RLC circuit is an electric circuit consisting of resistor, inductor and capacitor. RLC regulation circuits have been widely used to design time dependence of transfer function of the system because of their basic roles which can adjust filtration of frequency and system gains. Thus, the noise and unwanted ripple factors (which frequently happen in the analog circuits) can be removed by the use of RLC regulation circuits.

As we mentioned earlier, we could know that structure of boost converter can equate to a RLC regulation circuit. Therefore all two structures are happened to impedance phase. A basic RLC regulation circuit does not function as a compensation circuit for decreasing ripple. To decrease output ripple that is caused by impedance phase, we wanted to offset impedance phase of two structures. So we used the impedance matching method between a boost converter and a RLC circuit.

B. *Impedance matching Method*

Impedance is the combination of resistance and inductive reactance in a circuit that opposes current. Impedance matching is the practice of designing the input impedance of an electrical load or the output impedance of its corresponding signal source in order to maximize the power transfer or minimize reflections from the load. One application of a transformer is to match the impedance of a voltage source to the impedance of a load. [9] A Fig. 8 is impedance matching block diagram between a boost converter and proposed RLC circuit. To decrease output voltage ripple and to remove weakness of inrush current overflow and slow response speed by increase of capacitor value, addition between impedance phase of boost converter and impedance phase of RLC circuit have to be zero. Impedance matching enables maximum power to be delivered from the source to the load.

Identify applicable sponsor/s here. If no sponsors, delete this text box. (sponsors)



(a) Boost converter (b) RLC regulation circuit

(C) Boost converter with a RLC regulation circuit

Fig. 8 Block diagram

C. *Snubber circuit*

As we mentioned, the basic type converter is happened to overshoot phenomenon. This may cause excessive power dissipation and severe electromagnetic interference (EMI). To overcome these problems, various passive and active snubbers are proposed to reduce the current turn-off rate of rectifiers. It is expected to have a low impedance to charge all the capacitive loads during the transient phase of the start-up process until the system reaches a steady state. A snubber is a device used to suppress ("snub") voltage transients in electrical systems, pressure transients in fluid systems, or excess force or rapid movement in mechanical systems. The basic intent of a snubber is to absorb energy from the reactive elements in the circuit. The benefits of this may include circuit damping, controlling the rate of change of voltage or current or clamping voltage overshoot. In performing these functions a snubber limits the amount of stress which the switch must endure and this increases the reliability of the switch. We used modified RCD snubber type. A Fig. 9 is a snubber circuit that we modified a RCD snubber circuit. As aforementioned, overshoot output voltage is happened during switch on and off conditions. So difference of overshoot voltage and steady state voltage of drain part is decreased by resistor R_2. A capacitor is charged a stable pulse voltage during switch off condition such as Fig. 10 (a), and is discharged by resistor R_1 such as Fig. 10 (b). So, a drain voltage state is stabilized.

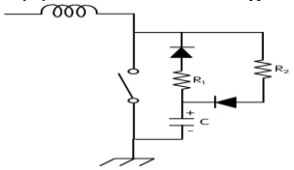
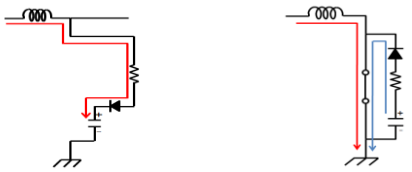


Fig. 9 Snubber circuit



(a) Off operation of a snubber circuit (b) On operation of a snubber circuit

Fig. 10 Operation of a Snubber circuit

III. Simulation and experiment results

As we mentioned, a Boost converter is happened to unstable electrical and optical characteristics. Simulation and experiment results of basic type boost converter are shown in Fig. 5, 6 and 7. Throughout simulation and experiment results, we could know unstable electrical characteristic such as overshoot and ripple phenomenon of output and we could also know that the luminance of LEDs is not kept the uniformity by unstable electrical characteristic.

A. *Design*

The proposed whole circuits consist of the boost converter, RLC regulation circuit and snubber circuit. In this circuit, a boost converter is selected as a main converter because of good efficiency and easy control of the input current. In addition, PWM input signals of the converter were used function generator signals, because of switch stress. In general, PWM input signals of the converter were made by a comparator that it was compared with both dimming signals and LED output ones. By a comparator, input signals of MOSFET gate were changed dynamic. So, stress of MOSFET gate was increased. Optimized values for each parameter of boost converter were obtained by steps of Fig. 11. For relationship between electrical characteristic and optical characteristic, we also operated optical experiment such as a Fig. 12. And input conditions are;

- DC-link voltage : 12V
- LED driver current : 240mA
- LED driver minimum current : 220mA
- LED driver maximum current : 260mA
- Switching frequency : 200kHz
- LEDs array : 4×4
- LED driver voltage : 20.5V
- LED driver ripple voltage: 4.0V

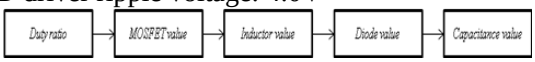


Fig. 11 This step is for calculated parameter values

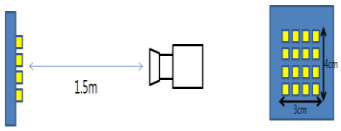
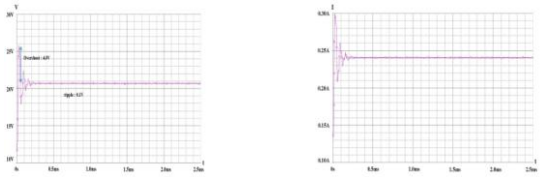


Fig. 12 Optical experiment environment

B. *Simulation and experiment*

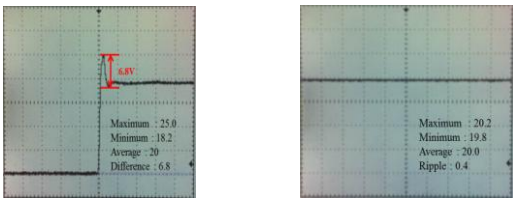
A Fig. 5 is simulation results of basic type converter, and a Fig.6 and a Fig. 7 are electrical and optical experiment results of basic type converter. Throughout two figures, we could know an overshoot and an output ripple phenomenon. As we mentioned, we adapted a RLC regulation circuit with an impedance matching method for removal output ripple phenomenon. Using calculated composition elements, we obtained simulation and experiment results such as Fig 13, 14 and 15. Certainly, an output ripple is decreased. But an overshoot phenomenon is yet happened and although luminance intensity is reduced, luminance uniformity of LEDs

is also not changed. Therefore unstable luminance phenomenon is that an overshoot phenomenon is more relative than an output ripple phenomenon. For removal overshoot phenomenon that is happened during switch on and off operations of boost converter, we adopted a snubber circuit. Throughout relation equations, we calculated composition elements. Using calculated composition elements, we could obtain simulation and experiment results such as Fig 16, 17 and 18. To know relation between overshoot phenomenon and luminance of LEDs, we did an overshoot change experiment. So, we obtained Fig. 19 and 20. Throughout Fig. 19 and 20, we could know that reduction of overshoot phenomenon can lead stable luminance of LEDs. As we mentioned, we could know that the luminance of LEDs is related to the amount of driving current. Therefore, stabilization of luminance can be performed by controlling the amount of flow of driving current.



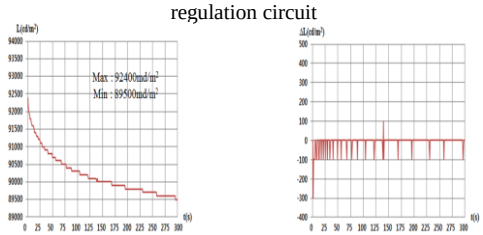
a) Output voltage (b) Output current

Fig. 13 Simulation results for boost converter with a RLC regulation circuit



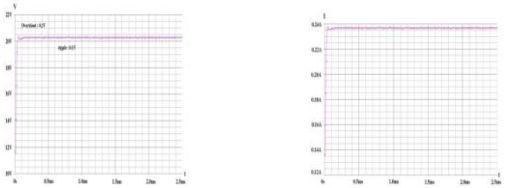
(a) Overshoot (b) ripple

Fig. 14 Electrical experiment results for boost converter with a RLC regulation circuit



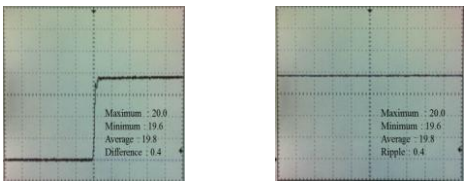
(a) Intensity (b) Uniformity

Fig. 15 Optical experiment results for boost converter with a RLC regulation circuit



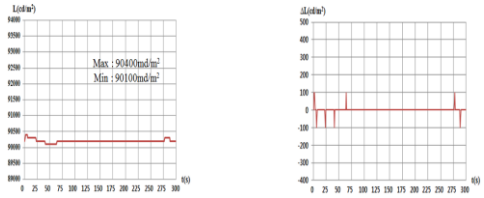
a) Output voltage (b) Output current

Fig. 16 Simulation results for boost converter with a RLC regulation circuit and a Snubber circuit



(a) Overshoot (b) ripple

Fig. 17 Electrical experiment results for boost converter with a RLC regulation circuit and a Snubber circuit



(a) Intensity (b) Uniformity

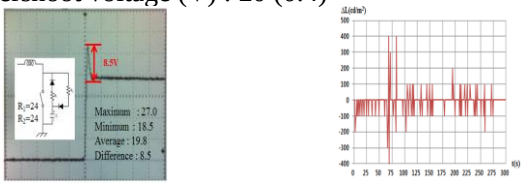
Fig. 18 Optical experiment results for boost converter with a RLC regulation circuit and a Snubber circuit

Our experiment results for proposed models are given as below and Table. 1:

A. basic type
1) Ripple voltage (V) : 17.6~21.8 (4.2)
2) Average voltage (V) : 20.1
3) Ripple current (mA) : 41.3~71.1 (29.8)
4) Average current (mA) : 57.5
5) Output power (W) : 0.7269~1.55 (0.8812)
6) Average power (W) : 1.156
7) Overshoot voltage (V) : 25 (7.4)

B. added LC circuit type
1) Ripple voltage (V) : 19.8~20.2 (0.4)
2) Average voltage (V) : 20
3) Ripple current (mA) : 54.54~57.85 (3.31)
4) Average current (mA) : 57.02
5) Output power (W) : 1.0799~1.1686 (0.0887)
6) Average power (W) : 1.1404
7) Overshoot voltage (V) : 25 (6.8)

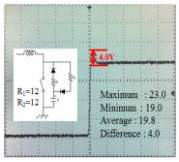
C. added LC circuit and snubber circuit type
1) Ripple voltage (V) : 19.6~20 (0.4)
2) Average voltage (V) : 19.8
3) Ripple current (mA) : 54.54~56.20 (1.66)
4) Average current (mA) : 55.37
5) Output power (W) : 1.069~1.124 (0.055)
6) Average power (W) : 1.0963
7) Overshoot voltage (V) : 20 (0.4)

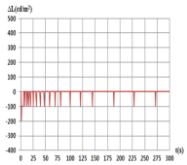


(a) Electrical characteristic (b) Optical characteristic

Fig. 19 Electrical and optical characteristics of proposed circuit by overshoot

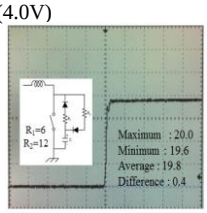
change (8.5V)

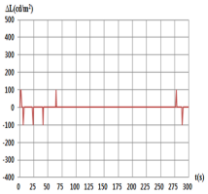




a) Electrical characteristic (b) Optical characteristic

Fig. 20 Electrical and optical characteristics of proposed circuit by overshoot change (4.0V)





(a) Electrical characteristic (b) Optical characteristic

Fig. 21 Electrical and optical characteristics of proposed circuit by overshoot change (0.4V)

TABLE I SUMMARY OF EXPERIMENT RESULTS FOR THE BLU CIRCUITS WITHOUT A REGULATION, WITH A RLC REGULATION CIRCUIT, AND WITH A SNUBBER CIRCUIT, RESPECTIVELY

	Basic type	Added LC circuit type	Added LC circuit and snubber circuit type
Ripple voltage	3.80	↓0.4V(89.474%)	↓0.4V(89.474%)
Ripple current	32mA	↓4mA(80%)	↓4mA(80%)
Ripple power	881.2mW	↓104.4mW(88.153%)	↓102.8mW(88.334%)
Difference (Overshoot voltage and undershoot voltage)	8V	↓6.8V(15%)	↓0.4V(95%)

IV. Conclusions

The proposed LED BLU circuit is consisted of boost converter, RLC regulation circuit and a snubber circuit for an output ripple and an overshoot phenomenon. To decrease output ripple, we selected a RLC regulation circuit with impedance matching method. And to decrease output overshoot voltage, we selected a snubber circuit. In particular, we found each parameter by circuit theories to take uniform input current and voltages of the 4 LED arrays. Throughout comparative analyses with a basic boost converter, we confirmed that the proposed LED BLU circuit has better output electrical characteristics and luminance uniformity than the basic boost converter.

References

[1] Won-Sik Oh, Kyu-Min Cho, Daeyoun Cho, Gun-Woo Moon, ByungChoon Yang, and Taeseok Jang, “A Novel Adaptive Dimming Technique with X-Y Channels for LED Backlight system of LCD TVs,” IEEE 2008

Daoshen Chen and Jih-Sheng Lai, “Design Consideration for Power Factor Correction Boost Converter Operating at the Boundary of Continuous Conduction Mode and Discontinuous Conduction Mode,” IEEE Applied Power Electronics Conference, Mar. 1993, pp. 267-273.

Liu Jian, Liu Shulin, Wang Zhao an. “Stability of Single Stage Power Factor Correction DCM Intergrated Converters,” Acta Electronica Sinica, 1999, 27(10):88-92. in Chinese.

Hyosung Kim, “Compensation Voltage Control in Dynamic Voltage Restorers by Use of Feed Forward and State Feedback Scheme,” IEEE Trans. Ind. Electron., vol. 20, no. 5, SEP 2005

Masahiro Nishikawat, Yoichi Ishizukat, Hirofumi Matsuot and Koichi Shigematsutt, “An LED Drive Circuit with Constant-Output- Current Control and Constant-Luminance Control,” Telecommunications Energy Conference, 2006. INTELEC '06. 28th Annual International, Sept. 2006

B.T. Irving, Y. Jang and M.M. Jovanovic, “A Comparative Study of Soft-Switched CCM Boost Rectifiers and Interleaved Variable-Frequency DCM Boost Rectifier,” IEEE Applied Power Electronics Conf., Mar. 2000, pp. 171-177.

Zhansong Zhang, Xuansan Cai, “Theory and Design of Switching power supply, Publishing house of electronics industry,” Beijing, 2004. in Chinese.

Masahiro Nishikawa, Yoichi Ishizuka, Hirofumi Matsuo and Koichi Shigematsu, “An LED Drive Circuit with Constant-Output-Current Control and Constant-Luminance Control,” IEEE 2006

A. Giovanni Beccuti, Georgios Papafotiou and Manfred Morari, “Optimal Control of the Boost dc-dc Converter,” 44th IEEE Conference on Decision and Control, and the European Control Conference 2005 Seville, Spain, December 12-15, 2005

L-Coupling Noise Countermeasure for LTE-Class Receiver RFIC by Using Magnetic Thin-Film

Wataru Kodate[1], Yasushi Endo[1], Masahiro Yamaguchi[1,2]
[1] Department of Electrical and Communication Engineering, Graduate School of Engineering, Tohoku University, Japan
[2] New Industry Creation Hatchery Center, Tohoku University, Japan
{kodate, endo, yamaguti} @ecei.tohoku.ac.jp

Abstract—**Performance of magnetic thin-film is evaluated as a candidate for on-chip noise suppressor. Miniature loops coils implemented on a 90 nm design test element group (TEG) chip are used as the noise transmitter and receiver. Differential transmission parameter S_{dd21} is evaluated as a measure of inductive noise intra-decoupling. Magnetic film is useful for frequency selective noise decoupler. The coupling is supported most at the ferromagnetic resonance (FMR) frequency, which is deviated from its intrinsic frequency and the degree of deviation can be calculated by using characteristic length. These results demonstrate that the proposed method is useful to characterize the IC chip level noise suppressor in the GHz range.**

Keywords-Long term evolution; intra-decoupling; on-chip; magnetic thin-film; ferromagnetic resonance

I. INTRODUCTION

Long Term Evolution (LTE) is a latest standard in use for wireless communication, which has a peak data rate over 300 Mbps in the downlink [1]. For LTE-class receiver radio frequency integrated circuit (RFIC), a design rule for semiconductors jumps into 65 nm, which is three generation finer than before, and the internal data transfer rate can be as fast as more than 1 Gbps. Accordingly fast switching noise from digital circuits degrades the performance of integrated analog circuits on the RFIC chip.

The noise propagation mechanisms on the chip are either of resistive (R), capacitive (C) or inductive (L) -couplings [2]. The countermeasure for the L-coupling is newly required from the LTE-era RFICs. Soft magnetic thin-films with high permeability are useful to suppress the noise above a certain frequency through ferromagnetic resonance (FMR) and Joule losses [3]. The magnetic thin film can also electromagnetically shield the noise not to radiate for external space [4]. Therefore, it is expected that the magnetic thin-films should be effective in the suppression of L-coupling in the RFIC chips. To begin with, it is necessary to develop a method to evaluate the performance of such a noise suppressor.

Besides, there is a recent IEC standard to evaluate sheet type ferromagnetic noise suppressor [5], where the methodology to evaluate the electromagnetic coupling between the electrically ground-ended two coils with diameter of 3 mm are precisely defined. However, the size of IC chips is often less than 3 mm and differential structure is preferable for smaller and faster chip operation. Therefore, new measurement methods for L-decoupling on the basis of the IEC standard to evaluate the sheet type ferromagnetic noise suppressor are proposed. This paper shows the on-chip measurements of the L-coupling and its countermeasure by using the magnetic thin-film.

II. EVALUATION MEATHODS FOR MAGNETIC THIN-FILM

Fig. 1 shows the L-coupling pass on an IC chip and the way to countermeasure it by using magnetic thin-film. One coil is applied RF currents and generates RF magnetic flux. This coil is regarded as a noise aggressor, such as digital circuits. Flux voltage received by the other coil is measured for the cases with and without the magnetic thin-film on the chip. This coil is regarded as a noise victim, such as analog circuits. The magnetic thin-film is evaluated by how the coupling between the coils is reduced by the noise suppressor. In Fig. 1 (a), intra-decoupling means the suppression of the coupling between two loop coils that are placed on the same side of the magnetic thin-film. On the other hand, in Fig. 1 (b), inter-decoupling means the suppression of the coupling between two loop coils that are placed on the other side of the magnetic thin-film. In this paper, the intra-decoupling is discussed.

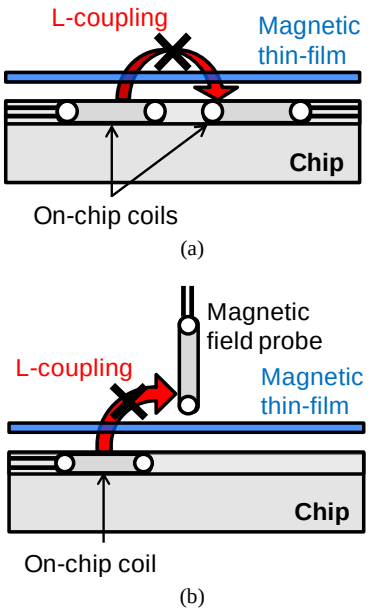


Fig. 1. L-coupling pass on an IC chip and the way to countermeasure it by using magnetic thin-film. The magnetic thin-film is evaluated by how the coupling between the coils is reduced by the noise suppressor. (a) Measurement method of intra-decoupling (b) Measurement method of inter-decoupling.

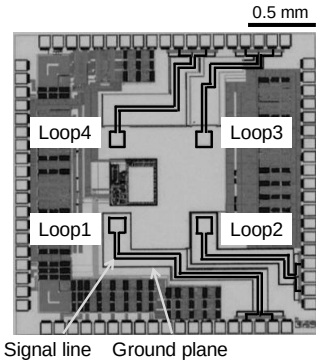


Fig. 2. A test element group chip implemented in CMOS 90 nm process.

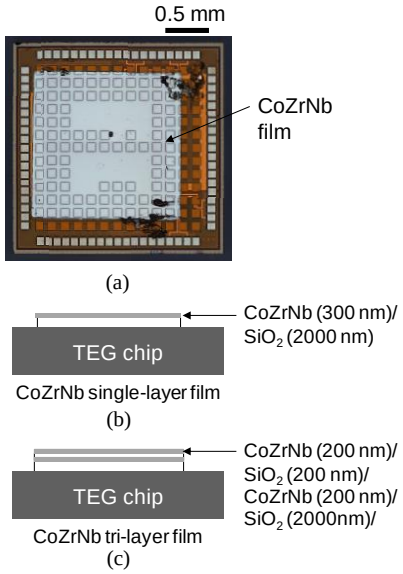


Fig. 3. TEG chip integrated with CoZrNb film. (a) Photograph of the chip. (b) Cross section with CoZrNb single-layer film. (c) Cross section with CoZrNb tri-layer film.

III. Integration of Films onto Chip

A. TEG Chip description

Fig. 2 shows a test element group (TEG) chip implemented in CMOS 90 nm process [6]. The chip is 2.5 mm × 2.5 mm in size and has four miniature differential loop coils. The size of the coils is 100 μm square and each coil is separated by 680 μm. The suppression of the coupling is evaluated by using these coils integrated with the magnetic thin-film. Long lead lines between the coil and pad may influence the measurements of the coupling between the coils, which is also discussed in this paper.

B. Magnetic Film

The $Co_{85}Zr_3Nb_{12}$ films were integrated by RF sputtering for experiments. They are with saturation magnetization M_s of 1.0 T, magnetic anisotropy field H_k of 1.0 kA/m, resistivity ρ of 1.2 × 10^{-8} Ωm, FMR frequency f_r of 1.2 GHz. Three test samples are prepared and integrated onto the surface of the TEG chip: CoZrNb single-layer film, CoZrNb/SiO_2/CoZrNb tri-layer film, and Cu single-layer film. Fig. 3 (a) show photograph of the chip integrated with the film. Hard axis of the 0.3 μm-thick single-layer CoZrNb film is in the right angle to the coupling direction between the two loop coils. The tri-layer CoZrNb (0.2 μm)/SiO_2 (0.2 μm)/CoZrNb (0.2 μm) film has easy axes orthogonally oriented to each other to obtain isotropic property in the film plane. Fig. 3 (b) and (c) explain the cross section of the single-layer and the tri-later films, respectively. The 0.3 μm-thick Cu film is integrated for comparison.

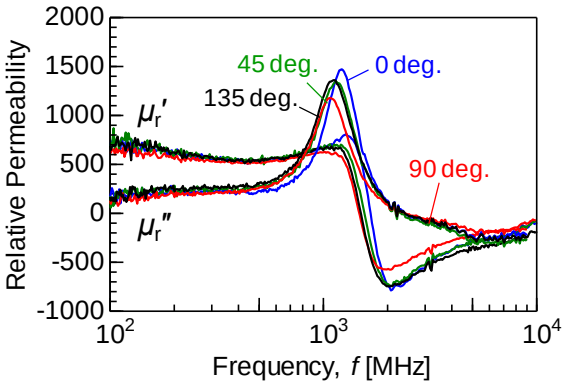


Fig. 4. Frequency profiles of relative permeability for CoZrNb tri-layer with variation of external magnetic field. μ_r' is real part of relative permeability and μ_r" is imaginary part.

Fig. 4 shows the measured frequency dispersion of complex permeability for the tri-layer film, showing that the film has isotropic permeability in the film plane. This is good for the IC chip level noise suppressor since the direction of noise current in the chip is usually random. Frequency profile of complex permeability for the CoZrNb single-layer is same as the 0 degree curves of the CoZrNb tri-layer.

IV. Intra-Decoupling Analysis

A. Measurement Setup

Suppression effect of the intra-coupling by the magnetic thin-film is evaluated by the differential transmission parameters of S-matrix between the coils. Loop1 and Loop4 in Fig. 1 are used in this work. Single-ended *S*-parameters are measured using two wafer probes (I40-D-GSGSG-100, Cascade Microtech Inc.) with four port network analyzer (N5230C, Agilent Technologies Inc.). The measured *S*-parameters are converted to differential *S*-parameters [7], [8] and applied for further discussion.

B. Experimental Results

Fig. 5 shows the differential transmission characteristics S_{dd21} between the two coils. S_{dd21} is proportional to the frequency for all the chips. A dip point of S_{dd21} is seen at 1.8 GHz for both CoZrNb single- and tri-layer films, suggesting that both films feel similar demagnetizing field to shift the FMR frequency from intrinsic value of 1.2 GHz to 1.8 GHz. The tri-layer film has another unknown dip at 1.4 GHz. The Cu film yielded 5 dB attenuation from 0.5 to 5 GHz. This means the magnetic film is useful for frequency selective noise decoupling.

C. Analysis of Noise Reduction Frequency

Some dip points for the CoZrNb single-layer film may be caused by the shift of the FMR frequency. It is reported that the loss peak frequency shifts toward those frequencies higher than

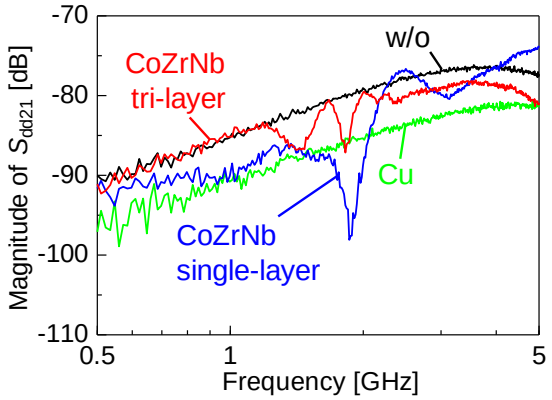


Fig. 5. Measurement result of transmission characteristic S_{dd21}.

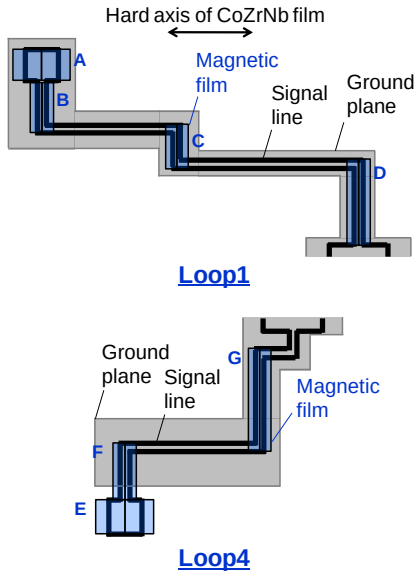


Fig. 6. Simulation model of Loop1 and Loop4 in Fig. 2. A is coil with GND. E is coil without GND. B, C, D, F, G are lead line with GND.

the intrinsic FMR frequency because of demagnetizing field [9]. Our idea is that the shift of the FMR frequency f_r by demagnetizing field is calculated by using characteristic length λ defined as the length where the magnetic flux along the magnetic film decreases down to $1/e$ of its original value [10].

Fig. 6 shows simulation model of Loop1 and Loop4 in Fig. 2. Magnetic flux B_x of the hard axis of the film is calculated by a three-dimensional FEM electromagnetic field simulator (HFSS Ver. 11.1.3, ANSYS Inc.). The measured seven areas are indicated as A-G in Fig. 6, which are coil with ground plane (A), coil without ground plane (E) and lead wires with ground planes (B, C, D, F, G).

Fig. 7 shows calculated magnetic flux B_x inside the film at areas A, E and F, respectively. The magnetic flux B_x is maximized on the signal lines. In case of the coil without ground plane (E), the magnetic flux B_x is three times higher than those of the rests. This is because the magnetic field generated by the ground plane negates magnetic field generated by the signal line. Each current of signal lines flows in the reverse direction and magnetic pole inside the film at the center of the parallel line (zero point) is reversed. The characteristic length λ from zero point to $1/e$ of magnetic flux B_x is regarded as effective film width w_s to calculate the demagnetizing factor.

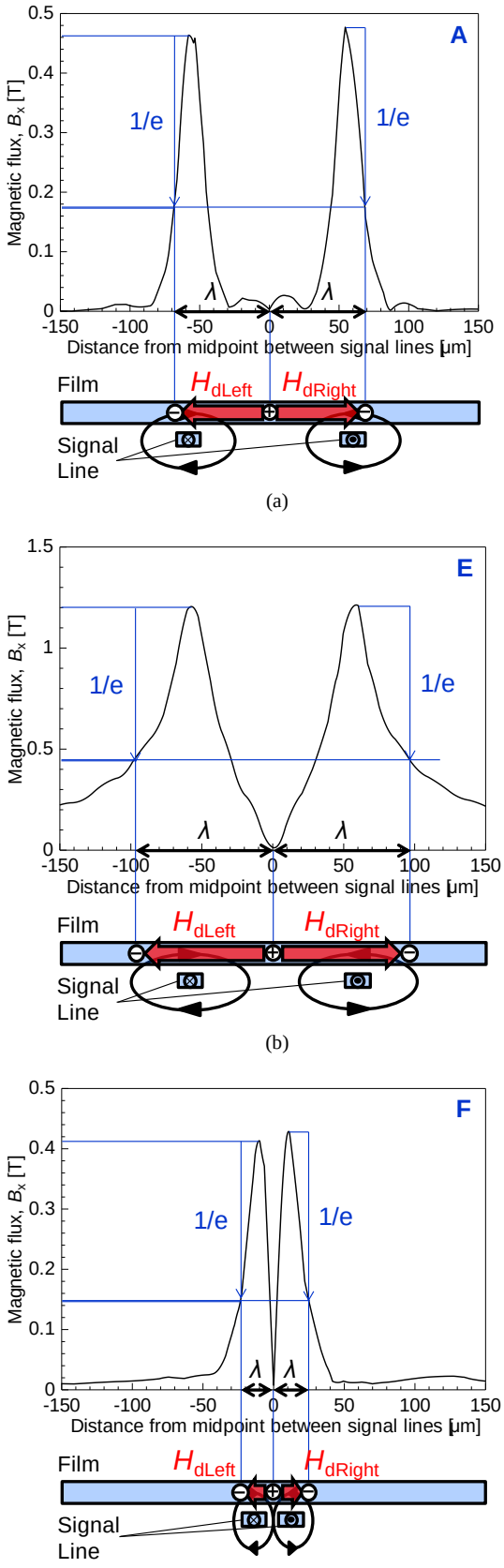


Fig. 7. Calculated magnetic flux B_x inside film (a) above coil with ground plane at area of A. (b) above coil without ground plane at area of E. (b) above lead with ground plane at area of F.

TABLE I. CALCULATED FMR FREQUENCY CONSIDERING DEMAGNETIZING FIELD

Measured area	Film configuration			Calculated results	
	Length l_m *[μm]*	*Width* w_s *[μm]*	*Thickness* t_m *[μm]*	*Demagnetizing factor* N_d	*FMR frequency* f_{rd} *[GHz]*
A	100	69	0.3	0.004	2.07
B	144	21	0.3	0.014	3.52
C	102	23	0.3	0.013	3.40
D	282	23	0.3	0.013	3.44
E	100	98	0.3	0.002	1.81
F	152	23	0.3	0.013	3.39
G	282	23	0.3	0.013	3.40

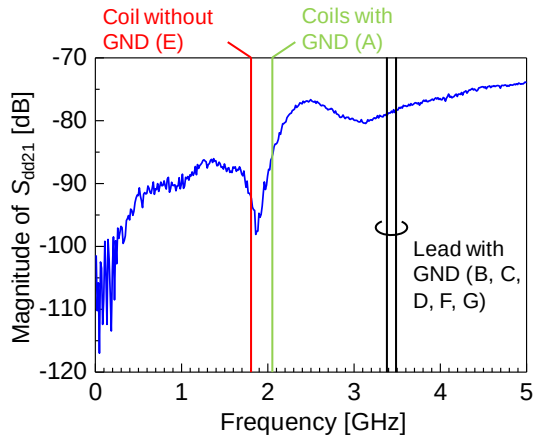


Fig. 8. Comparison of the calculated FMR frequency with measurement result of CoZrNb single-layer film in Fig. 5.

The demagnetizing factor N_d is calculated by approximating the magnetic film by a very flat ellipsoid whose axes are given by the film length l_m, film thickness t_m and the effective film width w_s [11]. Demagnetizing factor N_d is given by

$$N_d = \frac{t_m}{w_s}\frac{E-(w_s^2/l_m^2)K}{(1-w_s^2/l_m^2)} \quad (1)$$

where K and E are complete elliptic integrals. The FMR frequency f_{rd} considering the demagnetizing field is given by

$$f_{rd} = \frac{\gamma}{2\pi}\sqrt{\frac{M_s H_k + N_d M_s^2}{\mu_0}} \quad (2)$$

where γ is gyromagnetic ratio, μ_0 is permeability of vacuum.

Table 1 shows calculated FMR frequency considering the demagnetizing field of each area. In Fig. 8, these calculated results are compared with measurement result of the CoZrNb single-layer film shown in Fig. 5. Calculated results almost agreed with measurements, suggesting that some measured dips for the CoZrNb single-layer film are due to the demagnetizing field of the lead lines. It is also claimed that the shift of FMR frequency can be described by the characteristic length.

As described above, the lead lines between the coil and the pad have some effect on intra-decoupling performance. Therefore short lead lines are preferable. It also suggests to study line-to-line crosstalk. The next TEG chip is designed to take account of these considerations.

V. CONCLUSION

Measurement method of L-coupling for IC chip by magnetic thin-film is proposed. The measurement results demonstrate that the magnetic film is useful for frequency selective noise decoupler. The FMR frequency under demagnetizing field can be calculated by using characteristic length, and such FMR frequency gives the noise reduction frequency. For the coming TEG chip for intra-decoupling study, it is preferred to minimize lead line length between the coil and pads.

ACKNOWLEDGMENT

This work was supported in part by Development of Technical Examination Services Concerning Frequency Crowding from the Ministry of Internal Affairs and Communications of Japan. The authors are grateful to Dr. Norio Masuda, Netcomsec, Inc. (formerly NEC Co.) for the collaboration for the TEG chip. The authors are grateful to Mr. Yoshio Mitsuzuka for helping film deposition.

REFERENCES

[1] 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception (Release 8), 3GPP TS 36.101 V8.14.0, 2011.

[2] M. Yamaguchi, NICHe seminar and IEEE EMC Society Sendai Chapter Colloquim, CEWS-1-3, 2011.

[3] M. Yamaguchi, K. H. Kim, T. Kuribara, and K.-I. Arai, "Thin-film RF noise suppressor integrated in a transmission line," IEEE Trans. Magn.,vol. 40, no. 5, pp. 3183-3185, 2002.

[4] M. Yamaguchi, S. Koya, H. Torizuka, S. Aoyama, and S. Kawahito, "Shielded-loop type on-chip magnetic field probe to evaluate radiatedemission from thin-film noise suppressor," IEEE Trans. Magn., vol. 43, no. 6, pp. 2370-2372, 2007.

[5] International Electrotechnical Commission, "International Standard IEC62333-2 First Edition, Noise suppression sheet for digital devices and equipment – Part 2: Measuring methods," 2006.

[6] N. Masuda, "Development of On-chip Loop Coils for Evaluation of RF Noise Suppression Film," The 59th Electronic Components and Technology Conference (ECTC), pp. 815-818, 2009.

[7] D. E. Bockelman, W. R. Eisenstadt, "Combined Differential and Common Mode Scattering Parameters: Theory and Simulation," IEEE Trans. Microwave Theory Tech., vol. 43, pp. 1530-1539, 1995.

[8] W. Fan, Albert Lu, L. L. Wai, B. K. Lok, "Mixed-Mode S-Parameter Characterization of Differential Structure," Electronics Packaging Technology Conference, pp. 533-537, 2003.

[9] K. H. Kim, S. Ohnuma, and M. Yamaguchi, "RF Integrated Noise Suppressor Using Soft Magnetic Flms," IEEE Trans. Magn., vol. 40, no. 4, pp. 3031-3033, 2004.

[10] Jones, R., Jr, "Analysis of the efficiency and inductance of multiturn thin film magnetic recording head," IEEE Trans. Magn., vol. 14, no. 5, pp. 509-511, 1978.

[11] J. A. Osborn,"Demagnetizing Factors of the General Ellipsoid" Phys. Rev., vol. 67, pp. 351-357, 1945.

Multidimensional scaling-based passive emitter location from arrival of angular measurements

Xiaofeng Shen, Shuang Qin, Qun Wan, Wanlin Yang
University of Electronic Science and Technology of China
School of electronic engineering
Chengdu, People's Republic of China
E-mail: qinshuang.cc@gmail.com

Abstract—**In this paper, a novel multidimensional scaling (MDS) method based on arrival of angular (AOA) measurements is proposed for passive emitter location in wireless networks. Simulations are included to contrast the estimator performance with conventional MDS algorithm and least square algorithm designed for AOA measurements.**

Keywords: multidimensional scaling; arrival of angular (AOA); least square algorithm

I. INTRODUCTION

The emergence of Wireless Sensor Networks (WSNs) has facilitated our interaction with the physical environment. A WSN consists of a large number of distributed sensor nodes, which are generally inexpensive and resource constrained. The network is often configured such that the communication between the sensor nodes and the base stations requires multiple hops. Such a network topology can be traced back to the ancient defensive systems. Instead of using electronic sensors, in the past, beacon towers would send signals (e.g., beacon fires, flags, smoke and drums) upon the observation of enemy activity. The signals usually passed through several towers before reaching the command center. In contrast to this ancient system, modern WSNs require no or minimal human attendance.

Multidimensional scaling (MDS) analysis starts with one or more distance matrices (or similarity matrices) of points in a multidimensional space and finds a geometrical placement of the points in a low-dimensional space, usually 2-D or 3-D, where the distances between points resemble their original similarities [1].

Given the time-of-arrival (TOA) information, the location approach using MDS analysis has been verified to be more robust than the conventional algorithms in the condition of large range measurement noise, and received considerable interest. In the scene of cellular network MS localization, modified MDS algorithm is addressed in [2]. So and Chan [3] developed a MDS-based generalized subspace localization approach which allows any number of BSs. Qin *et al.* [4] have proposed a fast algorithm in minimum localization system. In the area of Parameters localization of coherently distributed sources, Guo has given a series of innovative research[5]-[7].

In this paper, we innovatively introduce multidimensional scaling algorithm into arrival of angular (AOA) measurements, and the rest of this paper is organized as follows. In Section 2-3, algorithm derivation and computer simulation are addressed. Conclusions are drawn in Section 4.

II. AOA BASED MULTIDIMENSIONAL SCALING ANALYSIS

In this paper, we consider the problem of mobile station (MS) location using AOA measurements at any BSs. Let $\mathbf{x}_0 = [x_0 \quad y_0]^T$ be the MS position to be determined, and BSs locate at $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_M \end{bmatrix}^T$, $\mathbf{y} = \begin{bmatrix} y_1 & y_2 & \cdots & y_M \end{bmatrix}^T$. We have position vector

$$\mathbf{p}_x = \mathbf{x} - x_0\mathbf{e} \tag{1}$$

$$\mathbf{p}_y = \mathbf{y} - y_0\mathbf{e} \tag{2}$$

and

$$\mathbf{P} = \left[\mathbf{p}_x, \mathbf{p}_y\right] \tag{3}$$

where $\mathbf{e}$ denotes the ones vector.

In AOA measurements, the measurements of angular $\varphi_i \ (i = 1, 2, \ldots, M)$ between MS and base stations (BSs) can be represented as

$$\mathbf{S} = \begin{bmatrix} \sin\varphi_1 & 0 & \cdots & 0 \\ 0 & \sin\varphi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sin\varphi_M \end{bmatrix} \tag{4}$$

$$\mathbf{C} = \begin{bmatrix} \cos\varphi_1 & 0 & \cdots & 0 \\ 0 & \cos\varphi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \cos\varphi_M \end{bmatrix} \tag{5}$$

constructing MDS symmetric matrix

$$\mathbf{Q} = \left[\cos\left(\theta_{ij}\right) \right] \tag{6}$$

where

$$\theta_{ij} = \pi - \left(\phi_{ij} + \phi_{ji}\right) \tag{7}$$

and

$$\phi_{ij} = \varphi_i - \gamma_{ij} \tag{8}$$

$$\phi_{ji} = \varphi_j - \gamma_{ji} \tag{9}$$

γ_{ij} and γ_{ji} are the direction angle between BSs.

According to the measurements of angular $\varphi_i \left(i = 1, 2, \ldots, M\right)$, we have

$$\mathbf{S}\mathbf{p}_x = \mathbf{C}\mathbf{p}_y \tag{10}$$

according to (1) and (2), we get

$$\mathbf{S}\mathbf{x} - x_0\mathbf{s} = \mathbf{C}\mathbf{y} - y_0\mathbf{c} \tag{11}$$

where $\mathbf{s} = \mathrm{diag}(\mathbf{S})$, $\mathbf{c} = \mathrm{diag}(\mathbf{C})$ is diagonal matrix.

Though matrix transformation we obtain

$$\begin{aligned} y_0 &= -\frac{1}{\mathbf{c}^T\mathbf{c}}\mathbf{c}^T\mathbf{S}\mathbf{x} + \frac{1}{\mathbf{c}^T\mathbf{c}}\mathbf{c}^T\mathbf{C}\mathbf{y} + \frac{1}{\mathbf{c}^T\mathbf{c}}x_0\mathbf{c}^T\mathbf{s} \\ &= \alpha_x + x_0\beta_y \end{aligned} \tag{12}$$

$$\begin{aligned} x_0 &= -\frac{1}{\mathbf{s}^T\mathbf{s}}\mathbf{s}^T\mathbf{C}\mathbf{y} + \frac{1}{\mathbf{s}^T\mathbf{s}}\mathbf{s}^T\mathbf{S}\mathbf{x} + \frac{1}{\mathbf{s}^T\mathbf{s}}y_0\mathbf{s}^T\mathbf{c} \\ &= \alpha_y + y_0\beta_y \end{aligned} \tag{13}$$

where

$$\alpha_x = -\frac{1}{\mathbf{c}^T\mathbf{c}}\mathbf{c}^T\mathbf{S}\mathbf{x} + \frac{1}{\mathbf{c}^T\mathbf{c}}\mathbf{c}^T\mathbf{C}\mathbf{y} \tag{14}$$

$$\alpha_y = -\frac{1}{\mathbf{s}^T\mathbf{s}}\mathbf{s}^T\mathbf{C}\mathbf{y} + \frac{1}{\mathbf{s}^T\mathbf{s}}\mathbf{s}^T\mathbf{S}\mathbf{x} \tag{15}$$

and

$$\beta_x = \frac{1}{\mathbf{c}^T\mathbf{c}}\mathbf{c}^T\mathbf{s} \tag{16}$$

$$\beta_y = \frac{1}{\mathbf{s}^T\mathbf{s}}\mathbf{s}^T\mathbf{c} \tag{17}$$

According to MDS symmetric matrix in (6), we have

$$\mathbf{P}\mathbf{P}^{\mathbf{T}} = \mathbf{R}\mathbf{Q}\mathbf{R} \tag{18}$$

where

$$\mathbf{R} = \begin{bmatrix} r_1^{-1} & 0 & \cdots & 0 \\ 0 & r_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_M^{-1} \end{bmatrix} \tag{19}$$

Then we have

$$\mathbf{Q} = \mathbf{R}^{-1}\mathbf{P}\mathbf{P}^{\mathbf{T}}\mathbf{R}^{-1} \tag{20}$$

Form (18), the MDS matrix $\mathbf{Q}$ is symmetric, moreover the rank of $\mathbf{Q}$ is 2.

Decomposing **Q** by eigenvalue factorization yields

$$\mathbf{Q} = [\mathbf{U}_s\mathbf{U}_n]\begin{bmatrix} \mathbf{\Sigma}_s & 0 \\ 0 & 0 \end{bmatrix}[\mathbf{U}_s\mathbf{U}_n]^T \tag{21}$$

Through orthogonal property, we have

$$\mathbf{U}_n^T\mathbf{R}^{-1}\mathbf{p}_x = \mathbf{0} \tag{22}$$

$$\mathbf{U}_n^T\mathbf{R}^{-1}\mathbf{p}_y = \mathbf{0} \tag{23}$$

from (22), (23) we obtain

$$\begin{cases} \mathbf{U}_n^T\mathbf{R}^{-1}\left(\mathbf{x} - x_0\mathbf{e}\right) = 0 \\ \mathbf{U}_n^T\mathbf{R}^{-1}\left(\mathbf{y} - y_0\mathbf{e}\right) = 0 \end{cases} \tag{24}$$

and

$$\begin{cases} \mathbf{U}_n^T\mathbf{X}\mathbf{r} = x_0\mathbf{U}_n^T\mathbf{r} \\ \mathbf{U}_n^T\mathbf{Y}\mathbf{r} = y_0\mathbf{U}_n^T\mathbf{r} \end{cases} \tag{25}$$

where $\mathbf{r} = \mathbf{diag}\left(\mathbf{R}^{-1}\right)$, $\mathbf{X} = \mathbf{diag}\left(\mathbf{x}\right)$, $\mathbf{Y} = \mathbf{diag}\left(\mathbf{y}\right)$ are diagonal matrix. Substituting (12), (13) into (24), (25) we have

$$\mathbf{U}_n^T\mathbf{Y}\mathbf{r} = y_0\mathbf{U}_n^T\mathbf{r} = \left(\alpha_x + x_0\beta_y\right)\mathbf{U}_n^T\mathbf{r} \tag{26}$$

$$\mathbf{U}_n^T\mathbf{X}\mathbf{r} = x_0\mathbf{U}_n^T\mathbf{r} = \left(\alpha_y + y_0\beta_y\right)\mathbf{U}_n^T\mathbf{r} \tag{27}$$

From (26), (27) we have

$$\begin{bmatrix} \mathbf{U}_n^T\mathbf{X} \\ \mathbf{U}_n^T\mathbf{Y} - \alpha_x\mathbf{U}_n^T \end{bmatrix}\mathbf{r} = x_0\begin{bmatrix} \mathbf{U}_n^T \\ \beta_x\mathbf{U}_n^T \end{bmatrix}\mathbf{r} \tag{28}$$

$$\begin{bmatrix} \mathbf{U}_n^T\mathbf{Y} \\ \mathbf{U}_n^T\mathbf{X} - \alpha_y\mathbf{U}_n^T \end{bmatrix}\mathbf{r} = y_0\begin{bmatrix} \mathbf{U}_n^T \\ \beta_y\mathbf{U}_n^T \end{bmatrix}\mathbf{r} \tag{29}$$

According to (28), (29), the MS location $\mathbf{x}_0 = [x_0 \quad y_0]^T$ can be estimated through generalized eigenvalue decomposition.

III. SIMULATION

Computer simulations have been carried out to evaluate the performance of the proposed algorithm. We compared MSPEs of the AOA measurements MDS estimator with least squares algorithm, classic MDS algorithm based on TOA measurements.

Each range error n_i is a zero-mean white Gaussian process with variance $\sigma_i = \varphi_i / SNR$, where SNR is the signal-to-noise ratio. Each result is an average of 10 000 independent runs. Starting with a mobile positioning scenario with the first three BSs of known coordinates (0, 0)m, (0, 6000)m, (6000, 6000)m, (6000, 0)m,(6000, -6000)m,(0, -6000)m,(-6000, -6000)m,(-6000, 0)m,(-6000, 6000)m.

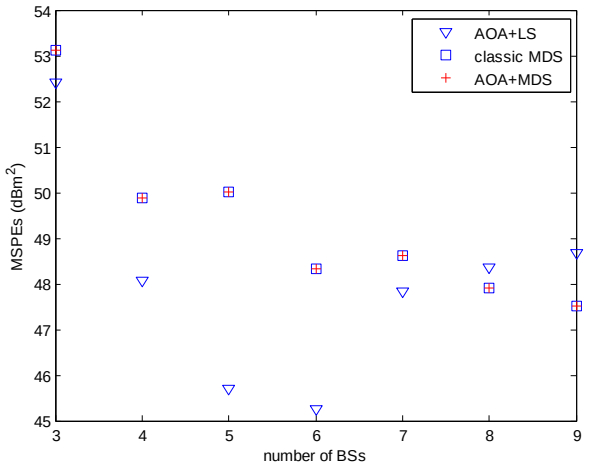


Fig. 1: MSPEs versus number of BSs for randomly distributed MS.

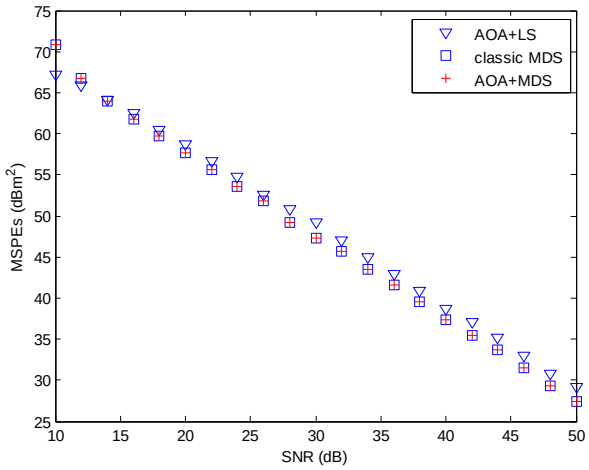


Fig. 4: MSPEs versus range error for randomly distributed MS.

the position of MS is uniformly distributed within the square bounded by (1000, 5000)m, (5000, 5000)m, (5000, -5000)m, (1000, -5000)m in each trial.

Fig.1 plots MSPEs versus number of BSs for randomly distributed MS. It is seen that the proposed algorithm has comparable performance with classic MDS algorithm even though they are based two different types of measurements. Classic MDS is based on TOA measurement; on the other hand, the proposed algorithm is based on AOA measurement.

Fig.2 plots MSPEs versus range error for randomly distributed MS. The number of BSs is 9, because from Fig.1 we can see that, when the number of BSs is added, the proposed algorithm is superior to LS algorithm.

IV. CONCLUSION

In this letter, we have addressed an AOA based MDS method for mobile positioning with at any BSs. Unlike the existing MDS algorithms depending on TOA measurements, the proposed algorithm introduced an new method to AO measurement location, which leads to a better performance when the number of BSs is enough.

REFERENCES

[1] I. Borg and P. Groenen, *Modern Multidimensional Scaling: Theory and Applications*. New York: Springer, 2005

[2] K. W. Cheung and H. C. So, "A multidimensional scaling framework for mobile location using time-of-arrival measurements," *IEEE Trans. Signal Processing*, vol. 53, no. 4, pp. 460-470, Apr. 2005

[3] H. C. So and K. W. Chan, "A generalized subspace approach for mobile positioning with time-of-arrival measurements," *IEEE Trans. Signal Processing*, vol. 55, no. 10, pp. 5103-5107, Oct. 2007

[4] S. Qin, Q. Wan, Z.-X. Cheng, A.-M. Huang, "Fast multidimensional scaling analysis for mobile positioning," IET Signal Processing, Volume 5, Issue 1, p.81-84 Feb. 20.

[5] Guo Xiansheng, Wan Qun, Wu Bin, Yang Wanlin. Parameters localisation of coherently distributed sources based on sparse signal representation. IET Radar Sonar & Navigation,2007,1(4):261-265(SCI:203AT,EI: 072710694090).

[6] Guo Xiansheng, Wan Qun, Shen Xiaofeng Dou Heng, Low-complexity parameters estimator for Multiple 2-D Domain Incoherently Distributed Sources, TURKISH JOURNAL OF ELECTRICAL ENGINEERING & COMPUTER SCIENCES, 2010, July. doi:10.3906/elk-0912-328.

[7] Guo Xiansheng, Wan Qun, Yang Wanlin, Lei Xuemei. Low-complexity decouple DOAs estimation algorithm for 2D coherentlydistributed sources. Sci China Ser F-Inf Sci,May 2009, vol. 52, no. 5, 835-842 (SCI: 000266094200010).

Implementation of Dimmable Visible Light Communication System

Seok-Ju Lee and Sung-Yoon Jung*
Dept. of EE, Yeungnam University / LED-IT Fusion Technology Research Center
Gyungsan si, Gyeongsangbuk-do, Republic of Korea
lsj0520@hotmail.com and syjung@ynu.ac.kr (*corresponding author)

***Abstract*—In this paper, we implement Dimmable Visible Light Communication (VLC) system based on VPPM (Variable Pulse Position Modulation) modulation scheme. VPPM is the modulation scheme proposed by IEEE 802.15 standard group, which can support illumination with dimming control and communication simultaneously. In addition, we propose DC bias elimination scheme in order to minimize the effect of external optical noises. As a result, 25%, 50%, and 75% brightness can be adjusted guaranteeing 1k bps data rate at 20cm Tx &Rx distance.**

Keywords-Dimming, visible light communication(VLC), variable pulse position modulation(VPPM), light emitting diode(LED).

I. Introduction

Rapid industrial development has increased the use of fossil fuels. Because greenhouse effect caused by carbon dioxide is growing, there are lots of efforts to solve this problem. The Kyoto Protocol, adopted by Conference of the Parties under the United Nations' Climate Change Convention, limits greenhouse gas emissions that cause global warming. As a consequence, many countries are developing Green IT technologies to reduce carbon emissions. Accordingly, LED lighting recently has attracted great attention as green IT technology because LED shows more electricity savings than incandescent lamp. Even though fluorescent lamps contain mercury, LED does not. LED's long life greater than 50,000 hours is another advantage than conventional lamps. Another benefit of LED is that it is a digital device. By turning 'ON' and 'OFF' LED, Visible Light Communication (VLC) which is considered as LED-IT convergence technology is possible.

VLC uses 350nm ~ 800nm wavelength for communication. It gives less pollution to environment and harmless to human body. Also, it does not occupy expensive frequency resources. Because LED lighting is recently considered for constructing green infrastructure, cost effective VLC services are also possible [1] [2].

In this paper, we implement dimmable VLC system based on VPPM modulation scheme. VPPM is the modulation scheme proposed by IEEE 802.15 standard group. To support illumination with dimming control and communication simultaneously, it uses binary PPM for communication and PWM for dimming control. In addition, our system minimizes the effect of external optical noises by providing DC bias elimination scheme. As a result, 25%, 50%, and 75% brightness can be adjusted guaranteeing 1k bps data rate at 20cm Tx & Rx distance. The rest of the paper is organized as follows. First, we describe the principle of VLC in Section II. Then, the implementation of the proposed VLC system is described in Section III. Next, we introduce the field test results in Section IV. Finally, conclusion will be presented in Section V.

II. Visible Light Communication

As shown in Fig. 1, the basic principle of VLC is to use the flicker of LED. LED's velocity for replacing electricity to light is about 30 nanometers to 250 nanometers. Therefore, fast switching can make VLC possible [3]. Using this principle, we can provide both lighting and communication functions, simultaneously. However, VLC should be designed to prevent Flickering phenomenon to human eyes. Flickering phenomenon is that the light changes are recognized by human eye, then losing the functionality as light. Therefore, IEEE 802.15.7 VLC standard has decided that Maximum Flickering Time Period (MFTP) is 5ms (200HZ) and the brightness between MFTP should be same [4].

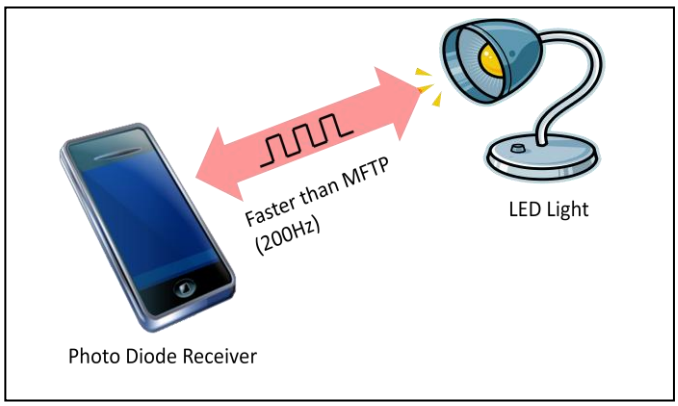


Fig. 1 Technology concept of visible wireless communication

For implementing dimmable VLC, our system uses VPPM modulation schemes. To support illumination with dimming control and communication simultaneously, it uses binary PPM for communication and pulse width for dimming control. Fig. 2 shows an operating example of the VPPM.

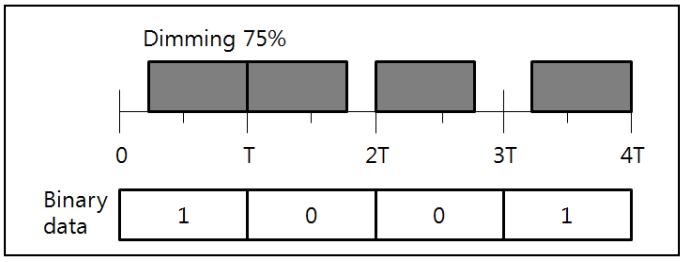


Fig. 2 Operating example of VPPM (75% dimming)

III. Implementation of Dimmable Visible Light Communication System

Fig. 3 shows a schematic diagram of the proposed VLC system using LED.

After creating the data from MCU, VPPM signal is generated through the LED driver. Then, it is transmitted through LED. After receiving VLC signal via PhotoDiode (PD), demodulation is performed after the DC bias removal and amplification. Finally, the transmitted data is displayed.

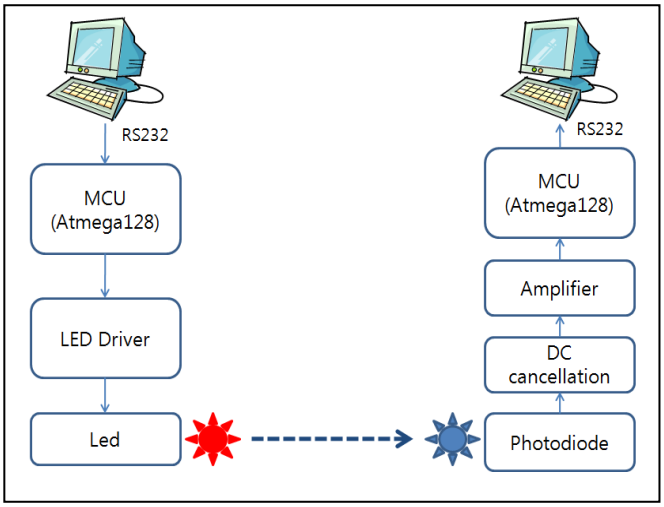


Fig. 3 Schematic diagram of the Visible Light Communication system

A. Transmitter

When we generate data through computer and transmit through RS232 protocol, MCU modulate them using VPPM and adjust it to the desired dimming level (25%, 50% and 75% dimming is possible in our system). Then, data is transmitted through LED. We use Atmega128 MCU and uses four 1W LED for light source. The layout of the transmitter is shown in Fig. 4. It is designed to control the switching of LEDs using Atmega128.

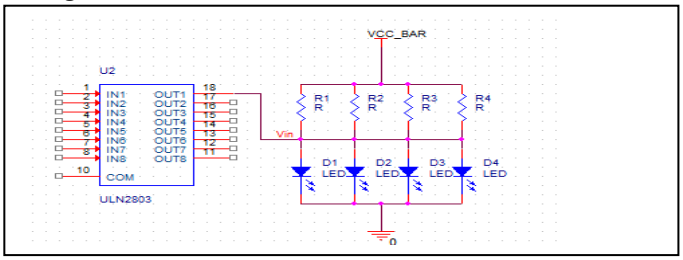


Fig. 4 Transmitter Layout

Fig. 5 shows a data frame structure of transmitting signal. If there are no data to transmit, dimming frame, which has predetermined dimming level, is sent waiting for data. When there are data to transmit, 'Start byte' of 8bit and payload is transmitted, sequentially. After data transfer is completed, the transmitter sends 'End byte'. Finally, dimming frame is sent again.

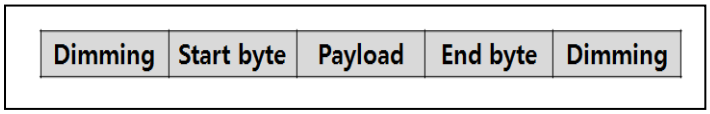


Fig. 5 Transmission frame structure

B. Receiver

Fig. 6 shows the layout of receiver circuit. Visible light received through PD is converted into electrical signals. We use BPW21R PD that represents maximum receiver sensitivity at a wavelength of 565nm. Then, we remove the DC bias and convert the signal to TTL level using the high pass filter. Finally, we amplify the signal ten times using OP-Amp to have max output greater than 2V.

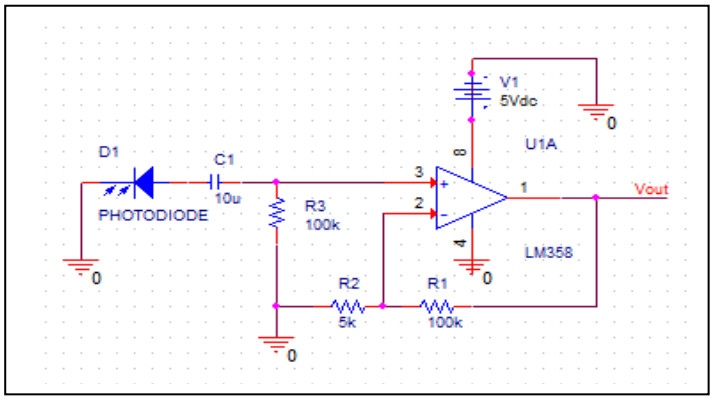
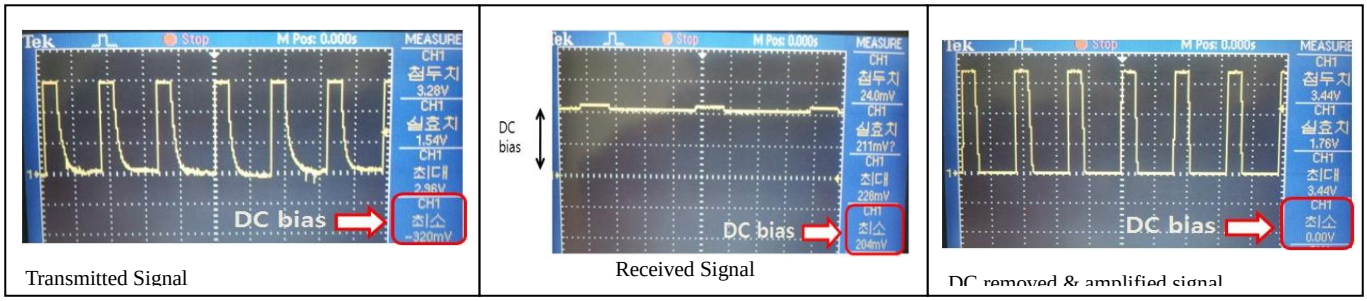


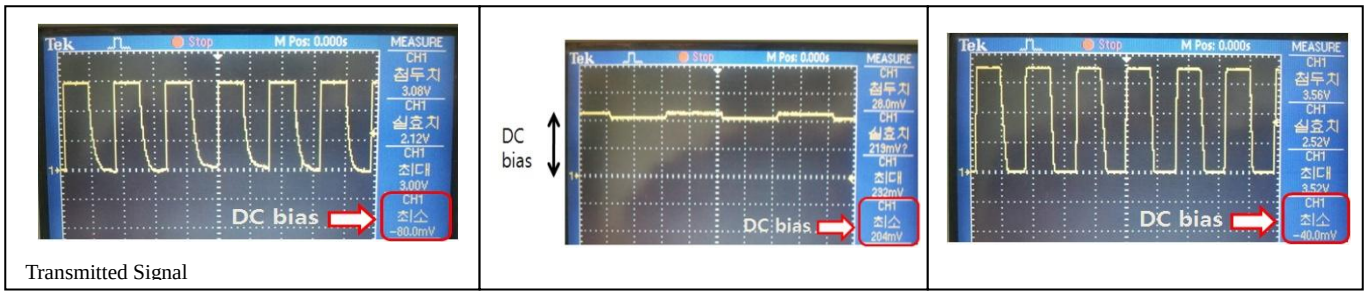
Fig. 6 Receiver Layout

Fig. 7 (A) - (C) shows examples of the transmitted and received signals according to dimming levels. The first figure shows the transmitted waveform. The second figure indicates received signal which contains DC bias due to environmental noise of ambient light. The last figure is DC removed and amplified signal.

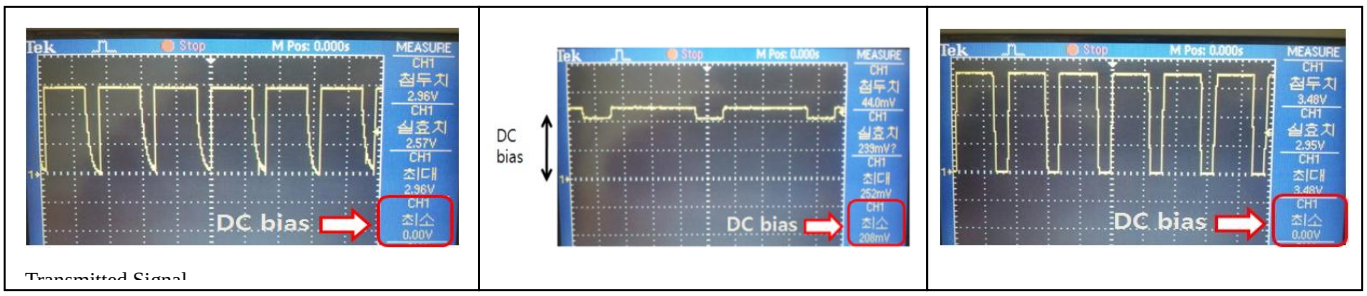
Fig. 8 shows the flow chart of the transmitter and receiver. As shown in Fig. 8 (a), initialization process is repeated until data come in. When transmitting data come, 'Start byte', data payload, and 'End byte' are modulated. Then, they are transmitted through LED. In Fig. 8 (b), the receiver detects 'Start byte' for synchronization. After that, output is displayed by demodulating payload information. If 'End byte' is finally received, the transmission procedure stops. For providing illumination functionality, dimming frame is transmitted when there are no data to transmit.



A. VPPM waveform with 25% dimming



B. VPPM waveform with 50% dimming



C. VPPM waveform with 75% dimming

Fig. 7 Transmission waveform for each dimming

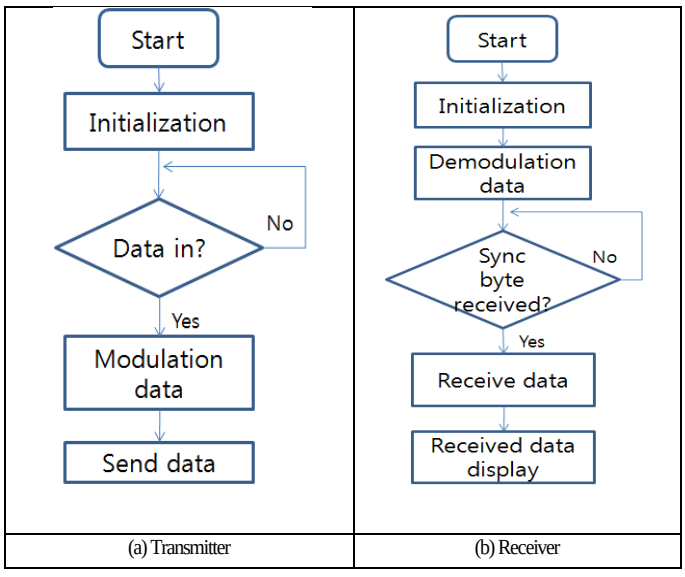


Fig. 8 Flow chart

IV. Field tests

Fig. 9 shows the implemented Visible Light Communication system. The distance between the transmitter and the receiver is about 20cm (Max. 40cm distance is possible). 1k bps data rate is possible at the baud rate of 9600kbps with dimming support.

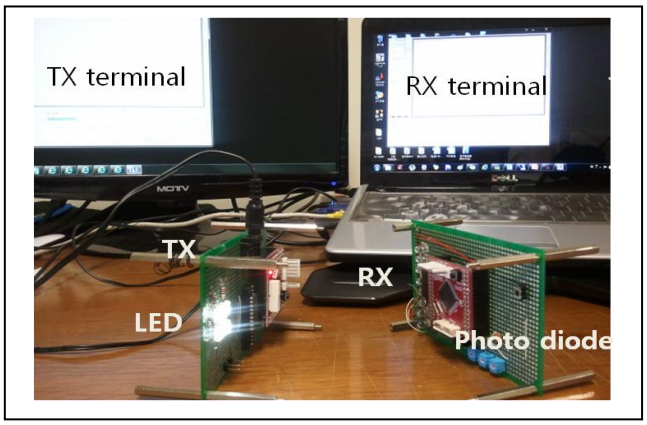


Fig. 9 Implementation of Visible Light Communication System

V. Conclusion

VLC is emerging LED-IT convergence technologies. Based on VPPM, we implemented dimmable VLC system to verify the operational possibility of VLC. In addition, our system tries to minimize the effect of external optical noises by providing DC bias removal scheme. Through field tests, 1kbps data transfer at 20cm distance is possible with 25%, 50%, 75% dimming support.

Acknowledgment

This research was conducted under the industrial infrastructure program for fundamental technologies (No. 10033630) which is funded by the Ministry of Knowledge Economy (MKE, Korea).

References

[1] T.G.Kang, "Wirelss Visible Light Communication application service model," ETRI, May 2009.

T.G.Kang, "The Convergence Technology of LED Illumination and Visible Light Communications," Electronics and Telecommunications Trends, vol.23, Oct 2008.

KOMINE T, NAKAGAWA M, "Fundamental analysis for visible-light communication system using LED lights," IEEE Trans. on Consumer Electron., vol. 50, no. 1, pp. 100–107, Feb.2004.

IEEE 802. 15. 7 Draft Task Group, Wireless MAC and PHY Specifications for Visible Light Wireless Personal Area Networks, IEEE 802. 15. 7 Draft D2, 2010.

Cost Function for Uniformly Distributed Source Signals Using Kullback-Leibler Divergence for Independent Component Analysis

Kota TANZAWA, Shunsuke KOSHITA, Masahide ABE and Masayuki KAWAMATA
Department of Electronic Engineering, Graduate School of Engineering,
Tohoku University, Sendai, Japan
Email:{tanzawa, koshita, masahide, kawamata}@mk.ecei.tohoku.ac.jp

***Abstract*— In this paper, we derive the cost function for uniformly distributed source signals using the Kullback-Leibler divergence for independent component analysis (ICA). To derive the cost function, we assume that two source signals are mutually independent and follow the uniform distribution. Under this assumption, we derive the probability density function (PDF) of an output signal using linear transformation of the PDF of the source signals. From the PDF of the output signal, we derive the cost function for the uniformly distributed source signals.**

***Index Terms*— Independent component analysis (ICA), cost function, Kullback-Leibler divergence, local minima.**

I. Introduction

Independent component analysis (ICA) is a method to separate source signals from observed signals without resorting to any prior knowledge [1] [2] [3]. We can separate the source signals from the observed signals by a linear transformation that transforms the observed signals to output signals that are matually independent. In order to find the linear transformation, we introduce a cost function which measures independency of the output signals. For example, the cost function using the Kullback-Leiblerdivergence is one of the most useful cost functions [4]. The cost function is minimized if and only if the output signals are mutually independent. That is, we can separate the source signals from the observed signals when the cost function is minimized.

Since the problem minimizing the cost function is a nonlinear optimization problem, it is difficult to solve it analytically. In general, the minimization problem is solved by using numerically iterative calculations, for example, gradient methods [5] [6]. Gradient methods might converge not global minimum but other local minima if the cost function has not only global minimum but also other local minima. However, no proof is available to show that the cost function has not only global minimum but also other local minima in ICA [1] [7]. It is necessary to examine the existence of the local minima except the global minimum of the cost function.

In this paper, to examine the existence of the local minima except the global minimum of the cost function, we derive the cost function for uniformly distributed source signals using the Kullback-Leibler divergence for ICA. To derive the cost function, we assume that two source signals are mutually independent and follow the uniform distribution. Under this assumption, we derive the probability density function (PDF) of the output signals using linear transformation of PDF of the source signals. From the PDF of the output signal, we derive the cost function for the uniformly distributed source signals.

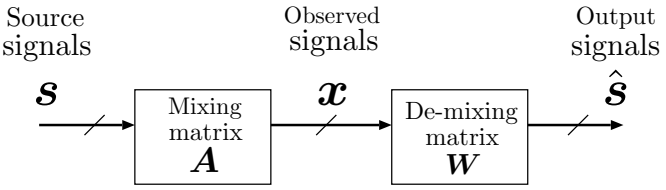


Fig. 1. Block diagram of blind source separation problem.

II. ICA and the Cost Function using Kullback-Leibler Divergence

This section introduces ICA and the cost function using the Kullback-Leibler divergence.

A. Independent Component Analysis (ICA)

Fig. 1 shows block diagram of blind source separation (BSS) problem. We assume that *source signals* $\boldsymbol{s} = [s_1, \cdots, s_m]^{\mathrm{T}}$ are mutually independent. The source signals $\boldsymbol{s}$ are linearly mixed and observed by m sensors. This process is modeled by an unknown nonsingular *mixing matrix* $\boldsymbol{A} \in \mathbb{R}^{m \times m}$ as follows:

$$\boldsymbol{x} = \boldsymbol{A}\boldsymbol{s} \tag{1}$$

where $\boldsymbol{x} = [x_1, \cdots, x_m]^{\mathrm{T}}$ are called *observed signals*. The goal of ICA is to separate the source signals $\boldsymbol{s}$ from the observed signals $\boldsymbol{x}$ without the knowledge of $\boldsymbol{A}$ nor $\boldsymbol{s}$. To achieve this task, we multiply $\boldsymbol{x}$ by a *de-mixing matrix* $\boldsymbol{W} \in \mathbb{R}^{m \times m}$ as follows:

$$\hat{\boldsymbol{s}} = \boldsymbol{W}\boldsymbol{x} \tag{2}$$

which gives *output signals* $\hat{\boldsymbol{s}} = [\hat{s}_1, \cdots, \hat{s}_m]^{\mathrm{T}}$, such that $\boldsymbol{W}$ makes $\hat{\boldsymbol{s}}$ mutually independent.

B. The Cost Function using Kullback-Leibler Divergence

To separate the source signals $\boldsymbol{s}$ from the observed signals $\boldsymbol{x}$, we find a de-mixing matrix $\boldsymbol{W}$ which makes the output signals

$\hat{\boldsymbol{s}}$ mutually independent. The output signals $\hat{\boldsymbol{s}}$ are mutually independnet if and only if

$$p(\hat{\boldsymbol{s}}) = \prod_{i=1}^{m} p_i(\hat{s}_i) \tag{3}$$

where $p_i(\hat{s}_i)$ is the PDF and $p(\hat{\boldsymbol{s}})$ is the joint PDF. To measure independency of $\hat{\boldsymbol{s}}$, we use the Kullback-Leibler divergence. The Kullback-Leibler divergence between two joint PDFs $f_1(\hat{\boldsymbol{s}})$ and $f_2(\hat{\boldsymbol{s}})$ is defined as

$$I = \int_{-\infty}^{+\infty} f_1(\hat{\boldsymbol{s}}) \log\left(\frac{f_1(\hat{\boldsymbol{s}})}{f_2(\hat{\boldsymbol{s}})}\right) d\hat{\boldsymbol{s}}. \tag{4}$$

The Kullback-Leibler divergence is a natural measure of deviation for two join PDFs. The Kullback-Leibler divergence I of Eq. (4) satisfies

$$I \geq 0 \tag{5}$$

with equality if and only if $f_1(\hat{\boldsymbol{s}}) = f_2(\hat{\boldsymbol{s}})$. Substituting $f_1(\hat{\boldsymbol{s}}) = p(\hat{\boldsymbol{s}})$ and $f_2(\hat{\boldsymbol{s}}) = \prod_{i=1}^{m} p_i(\hat{s}_i)$ into Eq. (4), we obtain

$$\begin{aligned} I = &\int_{-\infty}^{+\infty} q(\boldsymbol{x}) \log(q(\boldsymbol{x})) d\boldsymbol{x} - \log|\det(\boldsymbol{W})| \\ &- \sum_{i=1}^{m} \left(\int_{-\infty}^{+\infty} p_i(\hat{s}_i) \log(p_i(\hat{s}_i)) d\hat{s}_i \right) \end{aligned} \tag{6}$$

where $q(\boldsymbol{x})$ is the joint PDF. From Eq. (5), I is minimized only when Eq. (3) holds. Since the term $\int_{-\infty}^{+\infty} q(\boldsymbol{x}) \log(q(\boldsymbol{x})) d\boldsymbol{x}$ in Eq. (6) does not depend on $\boldsymbol{W}$, we omit the constant term from I of Eq. (6) and obtain the cost function $J(\boldsymbol{W})$ given by

$$\begin{aligned} J(\boldsymbol{W}) = &- \log|\det(\boldsymbol{W})| \\ &- \sum_{i=1}^{m} \left(\int_{-\infty}^{+\infty} p_i(\hat{s}_i) \log(p_i(\hat{s}_i)) d\hat{s}_i \right). \end{aligned} \tag{7}$$

Since Minimizing $J(\boldsymbol{W})$ is equal to minimizing I, from Eq. (5), $J(\boldsymbol{W})$ is minimized only when Eq. (3) holds. If we find $\boldsymbol{W}$ which minimizes the cost function $J(\boldsymbol{W})$, we can separate $\boldsymbol{s}$ from $\boldsymbol{x}$.

III. The Cost Function for Uniformly Distributed Source Signals

Since the problem minimizing $J(\boldsymbol{W})$ of Eq. (7) is a nonlinear optimization problem, it is difficult to solve it analytically. In general, the minimization problem is solved by using numerically iterative calculations, for example, gradient methods [5] [6]. Gradient methods might converge not global minimum but other local minima if the cost function has not only global minimum but also other local minima. However, no proof is available to show that the cost function has not only global minimum but also other local minima in ICA [1] [7]. It is necessary to examine the existence of the local minima except the global minimum of the cost function.

In this paper, to examine the existence of the local minima except the global minimum of the cost function, we derive the cost function for the uniformly distributed source signals using the Kullback-Leibler divergence for ICA. To derive the cost function, we assume that two source signals $\boldsymbol{s}$ are mutually independent and follow the uniform distribution $U(-1, 1)$. Under this assumption, we derive the PDF $p_i(\hat{s}_i)(i = 1, 2)$ using the linear transformation of the PDF of $\boldsymbol{s}$. From the PDF $p_i(\hat{s}_i)$, we express the term $\int_{-\infty}^{+\infty} p_i(\hat{s}_i) \log(p_i(\hat{s}_i)) d\hat{s}_i$ of Eq. (7) as the the equation of $\boldsymbol{G} = \boldsymbol{WA}$. From the term, we derive the cost function for the uniformly distributed source signals.

A. Derivation of the Probability Density Function of the Output Signal

From the assumption that two source signals $\boldsymbol{s}$ are mutually independent and follow the uniform distribution $U(-1, 1)$, we can write the joint PDF $r(\hat{\boldsymbol{s}})$ as

$$r(s_1, s_2) = \begin{cases} \frac{1}{4}, & |s_1| \leq 1 \text{ and } |s_2| \leq 1 \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

From Eqs. (1) and (2), we obtain

$$\hat{\boldsymbol{s}} = \boldsymbol{Wx} = \boldsymbol{WAs} = \boldsymbol{Gs} \tag{9}$$

where we define $\boldsymbol{G} = \boldsymbol{WA}$ to simplify the notation. Eq. (9) can also be written as

$$\hat{s}_i = g_{i1} s_1 + g_{i2} s_2, \;\; i = 1, 2 \tag{10}$$

where g_{ij} is the (i, j)-element of $\boldsymbol{G}$ and $j = 1, 2$. To obtain the PDF $p_i(\hat{s}_i)$ of $\hat{s}_i$, we use the linear transformation of the PDF of $\boldsymbol{s}$ [8]:

$$p_{is}(\hat{s}_i, s_2) = r(s_1, s_2) \left| g_{i1} \right|^{-1} \tag{11}$$

$$s_1 = -\frac{g_{i2}}{g_{i1}} s_2 + \frac{\hat{s}_i}{g_{i1}} \tag{12}$$

where $p_{is}(\hat{s}_i, s_2)$ is the joint PDF. From Eqs. (8), (11) and (12), we obtain

$$p_{is}(\hat{s}_i, s_2) = \begin{cases} \left| 4 g_{i1} \right|^{-1}, & \left| \frac{\hat{s}_i - g_{i2} s_2}{g_{i1}} \right| \leq 1 \text{ and } |s_2| \leq 1 \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

We integrate $p_{is}(\hat{s}_i, s_2)$ of Eq. (13) with respect to s_2 from $-\infty$ to $+\infty$ and obtain the PDF $p_i(\hat{s}_i)$ given by

$$p_i(\hat{s}_i) = \begin{cases} \alpha_i(\hat{s}_i + \beta_i), & -\beta_i \leq \hat{s}_i \leq -\gamma_i \\ (2 g_{iM})^{-1}, & -\gamma_i < \hat{s}_i < \gamma_i \\ \alpha_i(-\hat{s}_i + \beta_i), & \gamma_i \leq \hat{s}_i \leq \beta_i \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

where

$$\alpha_i = (4 g_{iM} g_{im})^{-1} \tag{15}$$

$$\beta_i = g_{iM} + g_{im} \tag{16}$$

$$\gamma_i = g_{iM} - g_{im} \tag{17}$$

$$g_{iM} = \max_j |g_{ij}|, \quad j = 1, 2 \tag{18}$$

$$g_{im} = \min_j |g_{ij}|. \tag{19}$$

B. Derivation of the Cost Function for Uniformly Distributed Source Signals

From the PDF $p_i(\hat{s}_i)$, we obtain the cost function for the uniformly distributed source signals. Substituting Eqs. (14), (15), (16) and (17) into the term $\int_{-\infty}^{+\infty} p_i(\hat{s}_i)\log(p_i(\hat{s}_i))d\hat{s}_i$ of Eq. (7), we obtain

$$\int_{-\infty}^{+\infty} p_i(\hat{s}_i)\log(p_i(\hat{s}_i))d\hat{s}_i = -\log(2g_{iM}) - \frac{g_{im}}{2g_{iM}}. \quad (20)$$

Substituing Eq. (20) into Eq. (7), we can derive the cost function $J(\boldsymbol{W})$ for the uniformly distributed source signals as

$$J(\boldsymbol{W}) = -\log|\det(\boldsymbol{W})| + \sum_{i=1}^{2}\left(\log(2g_{iM}) + \frac{g_{im}}{2g_{iM}}\right). \quad (21)$$

IV. Numerical Examples

In this section, we plot the cost function for the uniformly distributed source signals and visualize the contours of it. We consider the following three numerical examples:

Case 1: $\boldsymbol{A}$ is the identity matrix.
Case 2: $\boldsymbol{A}$ is a symmetric matrix.
Case 3: $\boldsymbol{A}$ is an asymmetric matrix.

To plot the cost function, we constrains the norm of the rows of the de-mixing matrix $\boldsymbol{W}$ to be unity:

$$\boldsymbol{W} = \begin{bmatrix} \cos\theta_1 & \sin\theta_1 \\ \cos\theta_2 & \sin\theta_2 \end{bmatrix} \quad (22)$$

Substituting Eq. (22) into $J(\boldsymbol{W})$ of Eq. (21), the cost function is expressed as

$$J(\theta_1,\theta_2) = -\log|\sin(\theta_2-\theta_1)| + \sum_{i=1}^{2}\left(\log(2g_{iM}) + \frac{g_{im}}{2g_{iM}}\right) \quad (23)$$

$$\boldsymbol{G} = \begin{bmatrix} a_{11}\cos\theta_1 + a_{21}\sin\theta_1 & a_{12}\cos\theta_1 + a_{22}\sin\theta_1 \\ a_{11}\cos\theta_2 + a_{21}\sin\theta_2 & a_{12}\cos\theta_2 + a_{22}\sin\theta_2 \end{bmatrix} \quad (24)$$

where a_{ij} is the (i,j)-element of $\boldsymbol{A}$. From the symmetric property of $J(\theta_1,\theta_2)$ of Eq. (23), we only show figures for

$$0 \le \theta_i < \pi, \ \ \theta_1 \le \theta_2. \quad (25)$$

A. Case 1: $\boldsymbol{A}$ is the Identity Matrix

The mixing matrix is

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (26)$$

Substituting Eq. (26) into $J(\theta_1,\theta_2)$ of Eq. (23), we plot $J(\theta_1,\theta_2)$ in Fig. 2.

Ideally, $\boldsymbol{W} = \boldsymbol{A}^{-1}$ minimizes the cost function. However, it is well known that due to the lack of prior information, there are the following two ambiguities in ICA [9]:

1. scaling ambiguity
2. permutation ambiguity.

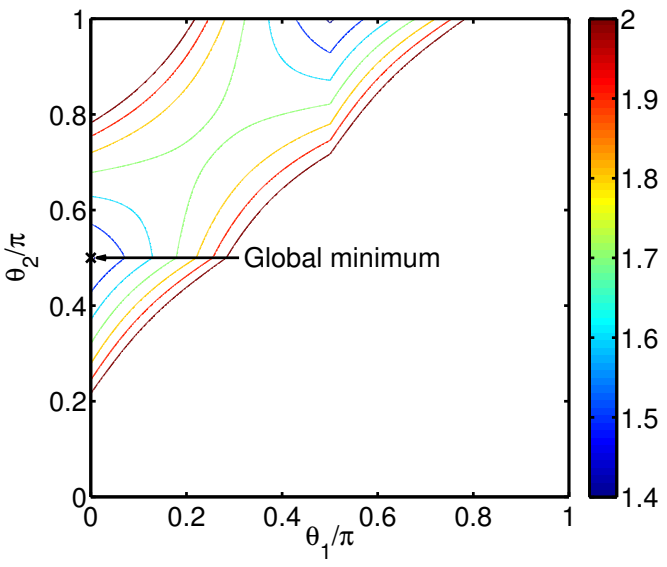


Fig. 2. Contours of $J(\theta_1,\theta_2)$ of Eq. (23) in case of Eq. (26).

That is, the de-mixing matrix $\boldsymbol{W}$ has the form

$$\boldsymbol{W} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{A}^{-1} \quad (27)$$

where $\boldsymbol{P}$ is a permutation matrix and $\boldsymbol{D}$ is a diagonal matrix. From Eq. (22), $\boldsymbol{D}$ is fixed and constrains the unit norm of the rows of $\boldsymbol{W}$. On the other hand, in this paper we set the constraints $\boldsymbol{P} = \boldsymbol{I}$ ($\boldsymbol{I}$ is the identity matrix) without loss of generality. Summarizing, in this section we obtain the global minimum of $J(\theta_1,\theta_2)$ of Eq. (23) by the following steps:

Step 1: Let $\boldsymbol{W} = \boldsymbol{A}^{-1}$.
Step 2: Divide the rows of $\boldsymbol{W}$ by the norm.
Step 3: Calculate θ_i from Eq. (22).

In case of Eq. (26), $J(\theta_1,\theta_2)$ of Eq. (23) is minimized by

$$(\theta_1,\theta_2) = \left(0, \frac{\pi}{2}\right). \quad (28)$$

Eq. (28) is the global minimum of $J(\theta_1,\theta_2)$ in Fig. 2. From Fig. 2, $J(\theta_1,\theta_2)$ has the global minimum and no other local minima in case of Eq. (26).

B. Case 2: $\boldsymbol{A}$ is a Symmetric Matrix

The mixing matrix is

$$\boldsymbol{A} = \begin{bmatrix} 1 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & 1 \end{bmatrix}. \quad (29)$$

Substituting Eq. (29) into $J(\theta_1,\theta_2)$ of Eq. (23), we plot $J(\theta_1,\theta_2)$ in Fig. 3. In case of Eq. (29), $J(\theta_1,\theta_2)$ of Eq. (23) is minimized by

$$(\theta_1,\theta_2) = \left(\frac{2}{3}\pi, \frac{5}{6}\pi\right). \quad (30)$$

Eq. (30) is the global minimum of $J(\theta_1,\theta_2)$ in Fig. 3. From Fig. 3, $J(\theta_1,\theta_2)$ has the global minimum and no other local minima in case of Eq. (29).

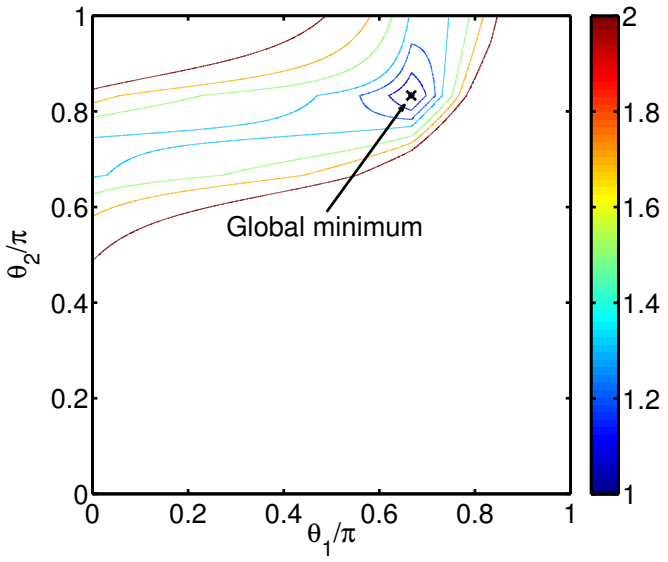


Fig. 3. Contours of $J(\theta_1, \theta_2)$ of Eq. (23) in case of Eq. (29).

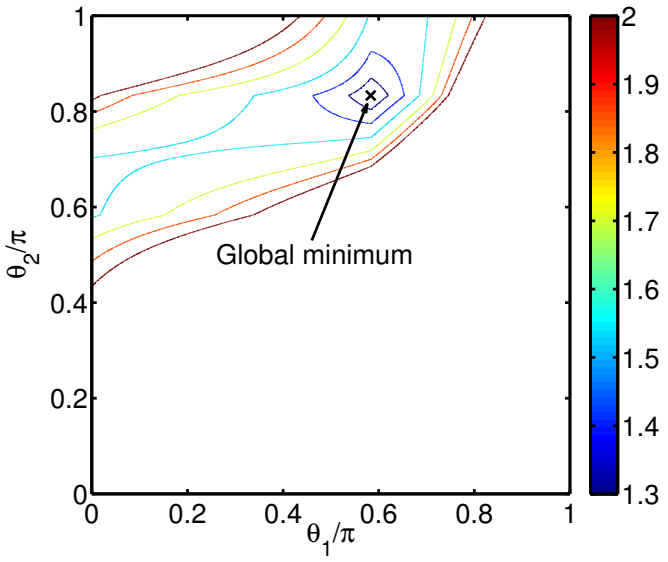


Fig. 4. Contours of $J(\theta_1, \theta_2)$ of Eq. (23) in case of Eq. (31).

C. Case 3: $\boldsymbol{A}$ is an Asymmetric Matrix

The mixing matrix is

$$\boldsymbol{A} = \begin{bmatrix} 1 & \frac{1}{\sqrt{3}} \\ 2-\sqrt{3} & 1 \end{bmatrix}. \tag{31}$$

Substituting Eq. (31) into $J(\theta_1, \theta_2)$ of Eq. (23), we plot $J(\theta_1, \theta_2)$ in Fig. 4. In case of Eq. (31), $J(\theta_1, \theta_2)$ of Eq. (23) is minimized by

$$(\theta_1, \theta_2) = \left(\frac{7}{12}\pi, \frac{5}{6}\pi\right). \tag{32}$$

Eq. (32) is the global minimum $J(\theta_1, \theta_2)$ in Fig. 4. From Fig. 4, $J(\theta_1, \theta_2)$ has the global minimum and no other local minima in case of Eq. (31).

V. Conclusion

In this paper, we derived the cost function for uniformly distributed source signals using the Kullback-Leibler divergence for ICA. To derive the cost function, we assumed that two source signals are mutually independnet and follow the uniform distribution. Under this assumption, we derived the PDF of a output signal using the linear transformation of the PDF of the source signals. From the PDF of the output signal, we derived the cost function for the uniformly distributed source signals. Our future task is to examine the existence of the local minima except the global minimum of the cost function theoretically. If the task is achieve, the result probably yields developments of the algorithms for ICA.

References

[1] P. Comon, "Independent component analysis, a new concept?" *Signal Processing*, vol. 36, no. 3, pp. 287–314, Apr. 1994.

[2] A. Hyvarinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural Computation*, vol. 9, no. 7, pp. 1483–1492, Oct. 1997.

[3] S. Choi, A. Cichocki, and S. Amari, "Flexible independent component analysis," *The J. of VLSI Signal Processing*, vol. 26, no. 1-2, pp. 25–38, Aug. 2000.

[4] A. Cichocki and S. Amari, *Adaptive blind signal and image processing*. Wiley, 2002.

[5] S. Amari, "Natural gradient works efficiently in learning," *Neural Computation*, vol. 10, no. 2, pp. 251–276, Feb. 1998.

[6] J.-F. Cardoso and B. Laheld, "Equivariant adaptive source separation," *IEEE Trans. Signal Processing*, vol. 44, no. 12, pp. 3017–3030, Dec. 1996.

[7] R. Boscolo, H. Pan, and V. P. Roychowdhury, "Independent component analysis based on nonparametric density estimation," *IEEE Trans. Neural Networks*, vol. 15, no. 1, pp. 55–65, Jan. 2004.

[8] H. Stark and J. W. Woods, *Probability and random processes with applications to signal processing*. Prentice Hall, 2002.

[9] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent component analysis*. Wiley, 2001.

A Novel Training based Channel Estimation Method in flat-block MIMO System

Junchao Yan, Qun Wan
Dept. of Electrical Engineering
University of Electrical Science and Technology of China
Chengdu, China
yjc801@uestc.edu.cn, wanqun@uestc.edu.cn

Abstract—**In this paper, we propose a novel method for training based channel estimation in multiple-input multiple output system. We study the computational performance of least squares (LS) and linear minimum mean-square-error (MMSE) by using optimal training matrices, and propose a Householder QR based MMSE estimator. It is shown that the proposed estimator has not only a good performance but also fewer computational complexes than conventional MMSE. Computer simulations confirm the effectiveness of the introduced method comparing with the existing methods.**

Keywords- multiple-input multiple-output; Householder transformation; MMSE; Matrix factorization

I. INTRODUCTION

With the growing demand for high transmission rates in wireless communication systems, multiple-input multiple-output (MIMO) communications have gained much attention due to their potential of greatly improving the performance over single-antenna systems. It has been shown that in the MIMO wireless systems the capacity and spectral efficiency can be greatly increased [1]–[2].

However, an accurate channel state information (CSI) is a key role in MIMO communications. So the method based on pilot signal (which is also known as training sequence) is widely used to estimate the channel. One of the most popular training based methods is the least square (LS) method, which does not need any a priori of the channel. When partial or full information of the channel is known, a better estimation can be achieved by minimum mean square error (MMSE) method. However, as MMSE method additionally exploits prior knowledge of the channel coefficients, it will inevitably cost higher complexity [3].

In [4], a study of the performance of training-based flat-block fading MIMO channel estimation is presented. And the optimal training matrices are studied for different training based channel estimation. But when it comes to acquiring real-time estimation, the matrix inversion resulted in high computational complexity will definitely decrease the performance of estimation and hence undesirable for hardware implementation. Matrix factorization is used to conduct large matrix calculations in alternate ways, and applied for system complexity reduction. In this paper, we propose a novel method for training based channel estimator based on Householder QR, which has fewer complexes than traditional MMSE.

The rest of paper is organized as follows. In Section II, the system model is introduced. Section III discusses the computational complexity of traditional LS and MMSE estimator by using the optimal training sequence. In Section IV, a less complex training based MMSE channel estimation approach is proposed. In Section V, experimentations are shown based on the proposed approach in Section IV. Finally, in Section VI, some concluding remarks are given.

We use the following notation. Matrices are represented by boldface upper case letters, and vectors are boldface lower case letters. For example, we use $\mathbf{A}$ to denote a matrix, $\mathbf{a}$ to denote a vector, and a to denote a scalar. To indicate the entries of matrices, we use subscripts. For example, $\mathbf{A}_{k,l}$ denotes the component at the k th row and l th column of the matrix $\mathbf{A}$.

$\mathrm{Tr}(\cdot)$ denotes the trace of a matrix, and $(\cdot)^H$ and $(\cdot)^T$ denotes the Hermitian transpose and the transpose of a matrix, respectively. The symbol $\|\cdot\|_F$ denotes the Frobenius norm of a matrix. $\mathbf{I}$ stands for the identity matrix.

II. SYSTEM MODEL

Consider a flat block-fading MIMO system with t transmit and r receive antennas, then the system model can be represented by

$$\mathbf{y}_i = \mathbf{H}\mathbf{s}_i + \mathbf{n}_i \tag{1}$$

where $\mathbf{H}$, $\mathbf{s}_i$, and $\mathbf{n}_i$ are the complex random channel matrix, transmitted signal and white Gaussian noise, respectively.

To estimate $\mathbf{H}$, we have the training sequence $\mathbf{s}_1,\ldots,\mathbf{s}_N$ transmitted, $N \geq N_t$. So the corresponding matrix $\mathbf{Y} = [\mathbf{y}_1,\ldots,\mathbf{y}_N]$ of received signals can be written as

$$\mathbf{Y} = \mathbf{H}\mathbf{S} + \mathbf{N} \tag{2}$$

This work was supported in part by the National Natural Science Foundation of China under grant 60772146, the Key Project of Chinese Ministry of Education under grant 109139 and in part by Open Research Foundation of Chongqing Key Laboratory of Signal and Information Processing (CQKLS&IP), Chongqing University of Posts and Telecommunications (CQUPT).

where

$$\mathbf{S} = \left[\mathbf{s}_1, \ldots, \mathbf{s}_N\right]$$

$$\mathbf{N} = \left[\mathbf{n}_1, \ldots, \mathbf{n}_N\right]$$

is the $N_r \times N$ training matrix, and the $N_t \times N$ i.i.d Gaussian random noise matrix with zero mean, respectively.

III. LS AND MMSE ESTIMATOR

In this section, we will first describe the LS and LMMSE estimator and then discuss their performance with the optimal training matrix.

A. LS Estimator

Considering that $\mathbf{H}$ and $\mathbf{S}$ are known, the estimation of channel matrix can be obtained by

$$\hat{\mathbf{H}}_{LS} = \mathbf{Y}\mathbf{S}^{\dagger} \tag{3}$$

where

$$\mathbf{S}^{\dagger} = \mathbf{S}^H(\mathbf{S}\mathbf{S}^H)^{-1} \tag{4}$$

is the pseudo-inverse of $\mathbf{S}$.

Under the power constraint of transmitted matrix,

$$\left\|\mathbf{S}\right\|_F^2 = P \tag{5}$$

the cost function can be written as

$$\min_{\mathbf{P}} E\left\{\left\|\mathbf{H} - \hat{\mathbf{H}}_{LS}\right\|_F^2\right\} \quad subject\ to\ \left\|\mathbf{S}\right\|_F^2 = P \tag{6}$$

where the P is a constant.

Solving the optimization problem (6), it can fairly be shown that the training matrix is optimal if it satisfices with the equation

$$\mathbf{S}\mathbf{S}^H = \frac{P}{t}\mathbf{I} \tag{7}$$

In [4], a normalized submatrix of the discrete Fourier transform (DFT) matrix is used as the training matrix.

$$\mathbf{S} = \sqrt{\frac{P}{NN_t}}\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & e^{j2\pi} & \cdots & e^{j2\pi(N-1)/N} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{j2\pi(N_t-1)/N} & \cdots & e^{j2\pi(N_t-1)(N-1)/N} \end{bmatrix} \tag{8}$$

Using (8) along with (3), the LS channel estimation can be obtained by

$$\hat{\mathbf{H}}_{LS} = \frac{N_t}{P}\mathbf{Y}\mathbf{S}^H \tag{9}$$

Note that the computation of (9) does not require matrix inversion by using the optimal training matrices, so it costs low complexity. However, LS estimation does not exploit any a priori knowledge of the channel coefficients. Thus, it is meaningfully to apply the MMSE estimator for a better performance of channel estimation by using the first and second order system statistics.

B. LMMSE Estimator

In general, a MMSE estimator can be obtained by the optimization function

$$\min E\left\{\left\|\mathbf{H} - \hat{\mathbf{H}}\right\|_F^2\right\} \tag{10}$$

Note that that the general MMSE estimator derived from (10) is in fact linear, but nonetheless it has repeatedly been referred to as the linear MMSE (LMMSE) estimator [6]–[7] which is correct but could lead to the incorrect conclusion that there may exist better non-linear estimators [5].

We assume that $\mathbf{A}$ is a linear operation, which can be obtained by minimize the mean square errors (MSE).

$$\mathbf{H}_{MMSE} = \mathbf{Y}\mathbf{A} \tag{11}$$

The estimation error can be written as

$$\begin{aligned} \varepsilon &= E\left\{\left\|\mathbf{H} - \mathbf{Y}\mathbf{A}\right\|_F^2\right\} \\ &= E\left\{\left(\mathbf{H} - \mathbf{Y}\mathbf{A}\right)^H\left(\mathbf{H} - \mathbf{Y}\mathbf{A}\right)\right\} \\ &= tr\left\{\mathbf{R}_{\mathbf{H}}\right\} - tr\left\{\mathbf{R}_{\mathbf{H}}\mathbf{S}\mathbf{A}\right\} - tr\left\{\mathbf{A}^H\mathbf{P}^H\mathbf{R}_{\mathbf{H}}\right\} \\ &\quad + tr\left\{\mathbf{A}^H(\mathbf{S}^H\mathbf{R}_{\mathbf{H}}\mathbf{S} + \sigma_n^2 N_r\mathbf{I})\mathbf{A}\right\} \end{aligned}$$

Let $\frac{\partial \varepsilon}{\partial \mathbf{A}} = 0$,

$$\mathbf{A} = (\mathbf{S}^H\mathbf{R}_{\mathbf{H}}\mathbf{S} + \sigma_n^2 N_r\mathbf{I})^{-1}\mathbf{S}^H\mathbf{R}_{\mathbf{H}} \tag{12}$$

$$\hat{\mathbf{H}}_{MMSE} = \mathbf{Y}(\mathbf{S}^H\mathbf{R}_{\mathbf{H}}\mathbf{S} + \sigma_n^2 N_r\mathbf{I})^{-1}\mathbf{S}^H\mathbf{R}_{\mathbf{H}} \tag{13}$$

where $\mathbf{R}_{\mathbf{H}} = E\left\{\mathbf{H}^H\mathbf{H}\right\}$.

And the $MSE \triangleq E\left\{\left\|\mathbf{H} - \hat{\mathbf{H}}_{MMSE}\right\|_F^2\right\}$.

By assuming that the channel is uncorrelated, (which means $\mathbf{R}_{\mathbf{H}} \propto \mathbf{I}$), the LMMSE estimator has the same condition on optimal training matrices as the LS estimator [4]. Moreover, replacing with $\mathbf{R}_{\mathbf{H}} = \alpha\mathbf{I}$, (13) can be expressed as

$$\hat{\mathbf{H}}_{MMSE} = \frac{N_t}{\sigma_n^2 N_r N_t + \alpha P}\mathbf{Y}\mathbf{S}^H\mathbf{R}_{\mathbf{H}} \tag{14}$$

where α is a parameter to minimize MSE.

In this case, the LMMSE estimator can also avoid the operation of matrix inversion, but achieve a better performance than LS estimator. However, the assumption may be unrealistic in practical applications. Therefore, we propose a method that does not involve matrix inversion, but has the same performance as LMMSE estimator.

IV. HOUSEHOLDER QR-BASED LMMSE ESTIMATOR

In this section, we propose a novel method for LMMSE estimator based on Householder QR (H-LMMSE).

The matrix triangularization is a widely used technique to avoid matrix inversion and is preferable due to its implementation in highly parallel array architecture [3] [9].

In Householder QR approach, a series of reflection matrix is applied to the matrix to annihilate the lower triangular elements. The reflection matrix is orthonormal matrix that can be written as

$$\mathbf{P} = \mathbf{I} - \mathbf{v}\mathbf{v}^H \tag{15}$$

where $\mathbf{v}$ is the Householder vector.

Support we have an arbitrary matrix $\mathbf{B}$,

$$\mathbf{P}_n, \cdots, \mathbf{P}_1\mathbf{B} = \mathbf{R} \tag{16}$$

where $\mathbf{R}$ is the upper-triangular matrix. The sequence of reflection matrices form the Householder matrix, where we have $\mathbf{P}_n, \cdots, \mathbf{P}_1 = \mathbf{Q}^H$, $\mathbf{Q}\mathbf{Q}^H = \mathbf{I}$.

$$\mathbf{B} = \mathbf{Q}\mathbf{R} \tag{17}$$

So the inversion of $\mathbf{B}$ can be obtained by

$$\mathbf{B}^{-1} = \mathbf{R}^{-1}\mathbf{Q}^H \tag{18}$$

The detail of estimation steps is given as follows.

Householder QR-based LMMSE Channel Estimator

Input: *Training sequence* $\mathbf{S}$, *observation signal matrix* $\mathbf{Y}$. *Choose the initial channel estimator* $\mathbf{H}_0$ *by LS.*

Output: *Householder QR based LMMSE estimator* $\hat{\mathbf{H}}$.

1) Find $\mathbf{Q}$, $\mathbf{R}$ by Householder QR matrix decomposition

Initialization $\mathbf{R} = \mathbf{S}^H\mathbf{R}_{\mathbf{H}_0}\mathbf{S} + \sigma_n^2 N_r\mathbf{I}$

$\mathbf{Q} = eye(size(\mathbf{R},1))$

for n=1,…,M-1

- Let $\mathbf{x} = \mathbf{R}(n:M, n)$ $\mathbf{y} = [1, zeros(M-n, 1)]$
- Update $\mathbf{v}$ and β

  by $\mathbf{v} = \dfrac{\mathbf{x} - \beta\mathbf{y}}{\|\mathbf{x} - \beta\mathbf{y}\|}$, $\beta = -\dfrac{x_1}{|x_1|}\cdot\|\mathbf{x}\|$,

  respectively.
- Compute $\mathbf{P}_n = \mathbf{I} - \mathbf{v}\mathbf{v}^T$, $\mathbf{P}_n' = \begin{bmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{P}_n\end{bmatrix}$
- Update $\mathbf{Q} = \mathbf{Q}\cdot\mathbf{P}'$, $\mathbf{R} = \mathbf{R}\cdot\mathbf{P}'$

end

2) $\hat{\mathbf{H}}_{MMSE} = \mathbf{Y}\mathbf{R}^{-1}\mathbf{Q}^H\mathbf{S}^H\mathbf{R}_{\mathbf{H}}$

V. SIMULATION

The parameters used in the simulation are listed in Tab. 1.

A. MSE Comparision with SNR

As is shown in Fig.1, we give the MSE performance comparisons of the three channel estimation methods under SNR ranged 0~20dB. The proposed estimation method has a better MSE performance than LS, and the same as LMMSE. The LMMSE method requires $\mathbf{R}_H$ as prior information. However, for practical systems, it is usually unknown. We usually use $\hat{\mathbf{R}}_H = \mathbf{S}\mathbf{S}^H$ instead of $\mathbf{R}_H$.

B. Computational Complexity

To study the computational complexity of the introduced method, we have evaluated the CPU time in second to complete the channel estimation for SNR=10dB. It is worth mentioning that although the CPU time is not an exact measure of complexity, it can give us a rough estimation of computational complexity. Our simulations are performance in MATLAB 2011a environment using a 2.40GHz Intel Core-2 processor with 4GB of memory and under Microsoft Windows 7 operating system.

The comparison of computational complexity between LS, LMMSE and H-LMMSE estimator is shown in Fig. 2. In common, the complexity of Householder QR is $O(m^2)$ where the complexity of the matrix inversion is $O(m^3)$. In [10], a fast Householder transform algorithm was proposed, which the computational complexity is only $O(m)$. As for the LMMSE, the complexity is increasing as the number of training sequence increasing. Thus, training based channel estimation with H-LMMSE is lower complexity method and hence easy implement at receiver on practical communications system.

TABLE I. SIMULATION CONDITION

Estimation Methods	LS	
	LMMSE	
	Householder QR-based LMMSE	
Antenna		4x4
Channel Fading		Flat block fading
Training Sequency S		DFT
Length of S		10, 20, 30, 40, 50, 60, 70, 80, 90, 100
Iterations		100

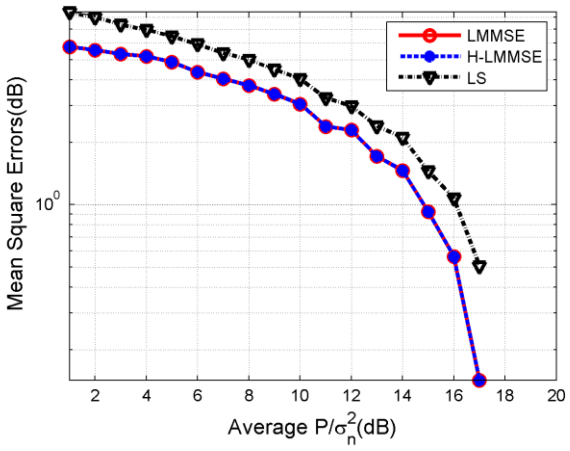


Figure 1. Channel estimation MSEs of LS, LMMSE and H-LMMSE estimator versus P / σ_n^2

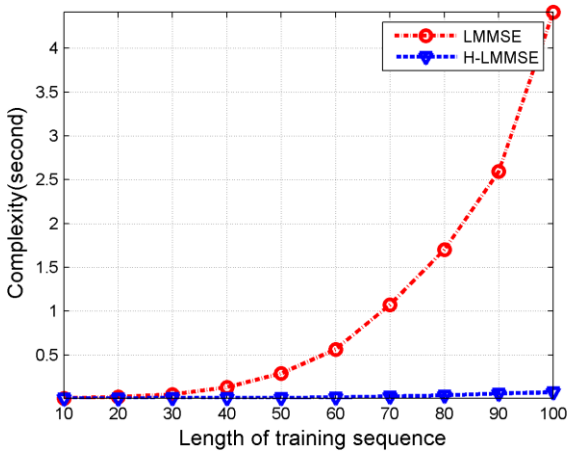


Figure 2. Complexity of LMMSE and H-LMMSE estimator versus Length of training sequence

VI. CONCLUSION

In this paper, we have proposed a novel training based channel estimation method based on Householder QR. This method has the advantages of the LMMSE estimator, but fewer complexes. It has been shown that, when compared with the existing estimator, our introduced method is both accurate and computationally efficient.

REFERENCES

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," Wireless Personal Commun., vol. 6, pp. 311–335, 1998.

[2] E. Telatar, "Capacity of multi-antenna Gaussian channels," Eur. Trans. Telecommun., vol. 10, pp. 585–595, 1999.

[3] N. M. † M. W. Numan, M. T. Islam, "Performance and Complexity Improvment of Training based Channel Estimation in MIMO Systems," *Electromagnetics*, vol. 10, pp. 1-13, 2009.

[4] M. Biguesh and a B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," IEEE Transactions on Signal Processing, vol. 54, no. 3, pp. 884-893, Mar.2006.

[5] E. Bjornson and B. Ottersten, "A Framework for Training-Based Estimation in Arbitrarily Correlated Rician MIMO Channels With Rician Disturbance," IEEE Transactions on Signal Processing, vol. 58, no. 3, pp. 1807-1820, Mar. 2010.

[6] Y. Liu, T.Wong, and W. Hager, "Training signal design for estimation of correlated MIMO channels with colored interference," IEEE Trans. Signal Process., vol. 55, no. 4, pp. 1486–1497, 2007.

[7] X. Ma, L. Yang, and G. Giannakis, "Optimal training for MIMO frequency-selective fading channels," IEEE Trans.Wireless Commun., vol. 4, no. 2, pp. 453–466, 2005.

[8] J. Kotecha and A. Sayeed, "Transmit signal design for optimal estimation of correlated MIMO channels," IEEE Trans. Signal Process., vol. 52, no. 2, pp. 546–557, 2004.

[9] K. Islam, N. Al-Dhahir, and R. McKown, "A new fast householder-based fractionally-spaced FIR MMSE-DFE computation algorithm and its real-time implementation," Communications, vol. 14, no. 11, pp. 1065-1067, 2010.

[10] Zheng-She Liu, "QR methods of O(N) complexity in adaptive parameter estimation," Signal Processing, IEEE Transactions on , vol.43, no.3, pp.720-729, Mar 1995

MIMO Based AOA Positioning with Phase Difference in Multipath Environments

Ik-Hyun Park
Department of Information and Communication Engineering
Yeungnam University
Gyeongsan, Republic of Korea
ihpark@ynu.ac.kr

Kook-Yeol Yoo
Department of Information and Communication Engineering
Yeungnam University
Gyeongsan, Republic of Korea
kyoo@yu.ac.kr

Abstract— **In this paper, we have studied the performance of the AOA (Angle of Arrival) in MIMO (Multi Input Multi Output) systems for LBS (Location Based Services) of 4G communication environment and we also analyzed the performance of the AOA in SISO (Single Input Single Output) in multipath environments and their differences. The adequacy of AOA positioning in 4G communication environments was determined. Currently used positioning methods in 3G communication environment has been developed based on SISO. However, the accuracy of SISO-based TOA (Time of Arrival), TDOA (Time Difference of Arrival), AOA positioning techniques degraded in multipath environments. The communication system of the future is 4G communication systems such as LTE-Advanced and enhanced positioning techniques are required. Using antenna characteristics and the phase difference between antennas of LTE-Advanced standard's key technique MIMO system AOA positioning, and SISO based AOA positioning performance were analyzed. We found that AOA technique potential for use based on MIMO systems by computer simulations.**

Keywords-component; LTE-Advanced, SISO, MIMO, AOA, phase difference, etc.

i. Introduction

Mobile internet devices such as smartphones and tablet PC are widely spread all over the world. Accordingly, LBSs which use location information of user are utilized in various fields, and the importance is growing bigger. Positioning in 3G communication environments are based on various methods and algorithms. Currently positioning systems for outdoor LBSs use mobile network and Global Positioning System (GPS).

The accuracy of GPS based positioning systems is high but they have some disadvantages such as an additional device is necessary and low accuracy in urban areas and not availability in indoor environments. Network based method solves the disadvantage of GPS based methods, but will have positioning error caused by multipath and NLOS. Network based positioning method mainly uses Cell ID positioning method that uses a cell radius of a base stations and an AOA positioning method that uses measurement of the azimuth in direction of a transmitted signal from two or more base stations and TOA positioning method that uses signal's time of flight transmitted from base stations and TDOA positioning method that uses time difference of received signal from base stations.

Cell ID positioning method grow away from the cell radius, positioning accuracy is greatly reduced. TOA and TDOA positioning method's positioning accuracy is decreased in multipath environments because they are sensitive to time changes caused by propagation delay. AOA positioning method error occurs due to the phase change cause by NLOS. To overcome these disadvantages in current 3G communication systems pattern matching algorithm is used for positioning. Pattern matching algorithm stores propagation characteristics of RF signal at measured location in database. And the location of mobile is found by comparing the propagation characteristics of RF signal of mobile user with database. However, pattern matching algorithm cannot build database when GPS signal is shadowed or in indoors. It is highly database dependent.

These network based positioning was current used in 3G communication systems. However, in near future those systems will be changed into 4G communication systems such as WiBro, LTE-Advanced. Accordingly, communication techniques and communication environments that are base of LBS will be changed and the positioning performances of each method will also be changed. The key techniques of 4G communication systems include MIMO, OFDM (Orthogonal Frequency Division Multiplexing), smart antenna, etc. In terms of positioning, MIMO technique based mobile and base stations with multiple antennas through spatial diversity in a multipath environments can result in a robust performance. The influence of a multi-path error onto transmission time is decreased due to short chip duration in OFDM compared to CDMA. Thus, it is important for positioning systems to adapt positioning techniques to changing mobile communication environments.

In this paper, we were confirmed a positioning error that used an AOA technique closely related an antenna characteristic for positioning in MIMO system. And we were confirmed the AOA positioning through azimuth estimate of the received signal using the phase difference between antennas. Through this, SISO system based AOA positioning performance was comparative analyzed.

ii. SISO system based AOA positioning

The AOA is a network based positioning technique; it measures the signal which was transmitted from each base station and its azimuth directions. AOA technique estimates

position of the user after obtaining azimuth angle which is known in different ways. Orientation information is defined clockwise direction from the north. One way to achieve a common base station for AOA measurements is to use antenna array in each sensor. Another way is measuring the transmitted signal's angle from the base station.

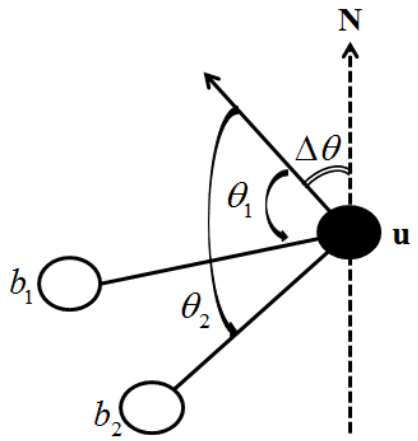


Fig. 1 Positioning with orientation information

Angle θ_1 and θ_2 in Figure 1 are relative signals which are transmitted by base station b_1, b_2 which is measured by each mobile, u. Assume mobile orientation information to be $\Delta\theta$, AOA at base stations b_1, b_2 become $(\theta_i + \Delta\theta)(\mathrm{mod}\, 2\pi), i = \{1,2\}$. Location of the mobile is determined by signals form base station which is perfect match for each AOA measurement. Mobile, u, is located at the intersection of lines when all base stations do not lie on a straight line.

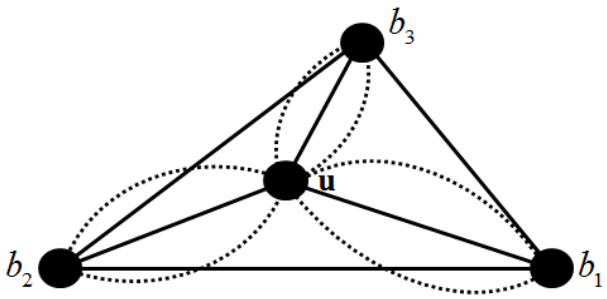


Fig. 2 Positioning without orientation information

Angles $\angle b_1 u b_2, \angle b_1 u b_3, \angle b_2 u b_3$ are calculated using the full AOA in Figure 2. All angles determine the range in front as the same line Therefore, arc is determined by the two base stations, theirs line and fixed angle. And mobile position's range is determined by its arc. Mobile u is located in intersection of b_1, u, b_2 which are determined by $\angle b_1 u b_2$'s angle and $b_1 b_2$ line.

These AOA techniques need LOS between the base station and mobile. In a multipath environment, it is difficult to measure AOA because of NLOS, which introduces range error. These errors are bigger when the distance between mobile and base station is longer. So the positioning performance degrades.

SISO system is single antenna system that is currently used communication system. It uses single transmit and receive antenna. Positioning accuracy in SISO system reduced by a delay under NLOS in multipath environments. When using AOA technique, due to this propagation delay the angle changes excessivly. Due to these problems in a network-based positioning with SISO TOA, TDOA and pattern-matching algorithms are widely used.

iii. MIMO system based AOA positioning

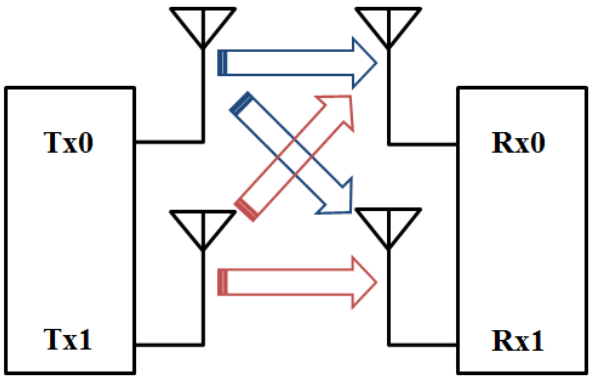


Fig. 3 Antenna configurations of MIMO system

MIMO system that was developed conventional to SISO, it uses multiple antennas. Using multiple antennas based on send and receives multiple signals. The MIMO system is a key technology for 4G mobile communication. It can increase system capacity without adding bandwidth by transmitting simultaneous data in different antennas. In addition, it can reduce the distortion in multipath environments through spatial multiplexing when the same data transmits to each antenna from multiple transmitters. Based on these spatial multiplexing in MIMO systems, we can lead to more robust performance for positioning. Network based positioning will improve accuracy when increased number of base stations. Thus, MIMO system using multiple antennas lead to the same effect that use a large number of base stations during positioning without additional base stations. Using multiple antennas, the positioning performance will be increased due to complementation of each signal's error in a multipath environment.

For AOA positioning, we have to measure the received signal's angle from the base station. To accomplish this, we have to distinguish the signal sent by multi-path of the

received signal and find the first arriving path. These paths become LOS signals and we can measure angle using these LOS path.

In this paper, SISO and MIMO based AOA positioning error is analyzed for determining the performance of MIMO based positioning in real communication environment. For this reason, we set up the channel had the Gaussian distribution, mean is LOS signal value and variance is 1. Transmitted signal is received through this channel. At this time, an azimuth of the received signal is changed randomly on the basis LOS component. This is to have set up the Rician fading channel model to include a LOS component.

IV. MIMO based AOA positioning using the phase difference

Another AOA positioning is possible if we use a characteristic of the MIMO system that has multiple antennas. The received signal at each antenna is will have a transmission delay due to the distance between antennas. An amplitude and phase of received signal at each antenna is differed by this delay. In this case, we use the phase difference of each antenna we can estimate the received angle of signal. AOA positioning is possible using the received angle of signal.

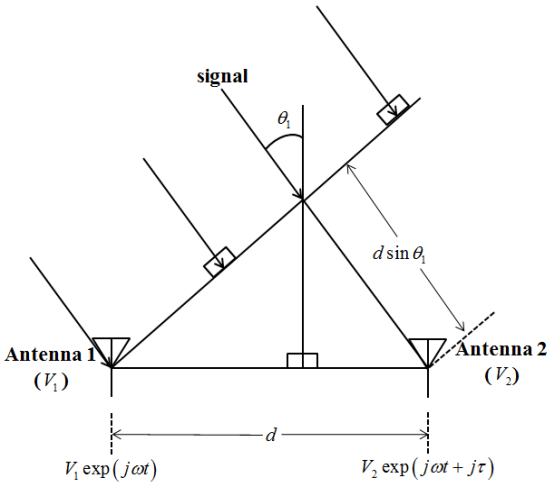


Fig. 4 Azimuth estimate using the phase difference between antennas

Figure 4 shows how to estimate the received angle that used a phase difference. D shows the distance between each antenna. θ_1 shows the received angle of signal. We can calculate delay distance and delay time between each antenna using this. Delay distance can show to $d\sin\theta_1$, and we use a formula of a trigonometric function. We can calculate delay time as dividing delay distance by speed of light. A phase difference between antenna v1 and v2 can show product of frequency of transmitted signal and delay time. This can be expressed as follows

$$\tau = f\left(d\sin\theta_1 / C\right). \quad (1)$$

where τ is phase difference of each antenna and f is frequency of transmitted signal and C is the speed of light.

The received angle of signal is estimated using the phase difference of each antenna as shown in (1). The received angle which is similar to LOS is got using these results, and the AOA positioning becomes possible. Because the AOA positioning is sensitive to change of the received angle due to multipath, it has a disadvantage to increase positioning error even if the angle has a little change. But the phase change of signal is smaller than the received angle change in multipath environment. Therefore the positioning error decreases when the AOA positioning to estimate the received angle using the phase difference is performed, and the phase difference can be calibrated using phase estimation. In conclusion, accuracy of the AOA positioning can increase in the multi-path environment.

V. Simulation result

To check a positioning error characteristic based on MIMO and SISO, we simulated the AOA positioning, and made CDF (Cumulative Distribution Function) of positioning error distance through the repetitive simulation. In this way, we analyzed error distribution of each method.

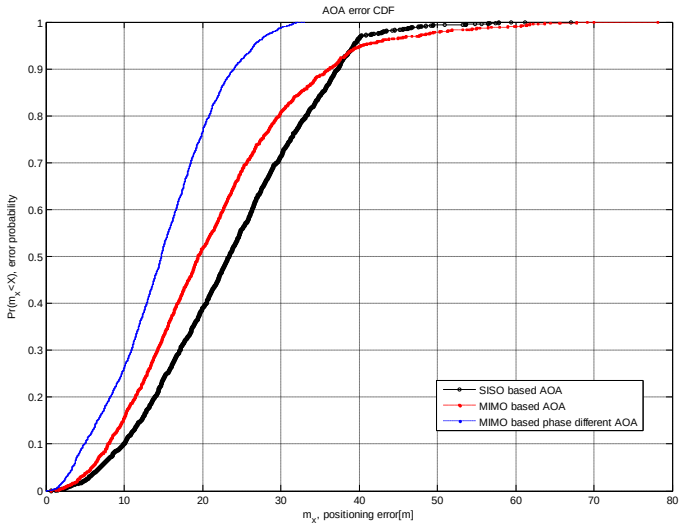


Fig. 5 Cumulative Distribution Function of AOA positioning error

Figure 5 shows the error CDF of each method. Black line shows error CDF of SISO system based AOA positioning. Red line shows error CDF of MIMO system based AOA positioning. And blue line shows error CDF of MIMO system based AOA positioning using phase difference. Each error is shown with the difference of distance between actual mobile position and estimate position.

Graphs depict the maximum distance error of SISO system based AOA positioning is approximately 70m. And it depict distributed mainly to 0~40m error. Also, the maximum distance error of MIMO system based AOA positioning is approximately 70m. But we can observe that less than 40m distance error appears much more compare with SISO system based AOA positioning. Otherwise, we

can observe that the maximum distance error of AOA positioning using phase is 30m and it is very small compared to the conventional methods.

An average error of each method is approximately 23m in case of SISO system based AOA positioning, that is approximately 21m in case of MIMO system based AOA positioning. And average error is approximately 14m in case of MIMO system based AOA positioning using phase difference. Through this, we can know that AOA positioning using multiple antennas is more accurate. Moreover, we can observe that the proposed method achieves more accuracy compared to AOA method only using received angle under the multipath environment.

VI. Conclusions

We analysis an error of MIMO and SISO systems based AOA positioning in multipath environment. Through this, we verified the performance of MIMO based AOA positioning in multipath environment. In MIMO based AOA positioning produces more correct location information even some signals have their azimuth angle changed greatly by multipath.

In the case of the AOA positioning using the phase difference, the error distance is smaller than other methods. It is because the phase change of multipath is smaller than the received angle change, and we can estimate the received angle of signal at closely the received angle of LOS signal using this method. If the phase error is calibrated through channel estimation, we can estimate the closer received angle at LOS signal, and the AOA positioning accuracy will increase.

Accordingly, MIMO based AOA positioning due to multipath is less sensitive to changes in azimuth and it has robust performance compare to SISO based AOA positioning. Moreover, MIMO based received angle estimate AOA positioning using the phase difference and it has robust performance compare to AOA using as an azimuth.

We will use 4G communication systems including MIMO. 4G based AOA positioning compare to SISO based AOA positioning of the 3G communication systems are expected to have an excellent positioning performance. In addition to MIMO, key techniques of 4G mobile communication systems are expected to have an improved positioning performance based on development of various positioning methods.

References

[1] Rong Peng, Mihail L. Sichitiu, "Angle of Arrival Localization for Wireless Sensor Networks, " in proc.SensorandAdHocCommunicationsandNetworks2006,pp.374-382 ,2006.

Dragos Niculescu, Badri Nath, "Ad Hoc Positioning System (APS) Using AOA," in proc.INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications, pp.1734-1743,2003.

Grzegorz Czopik, Adam Kawalec, "A HYBRID AOA-TDOA SIGNAL SOURCE LOCALIZATION ALGORITHM," in proc.Microwaves, Rader and Wireless Communications, 2008, pp. 1-4, 2008.

Ali Abdi, Janet A. Barger, Mostafa Kaveh, "A Parametric Model for the Distribution of the Angle of Arrival and the Associated Correlation Function and Power Spectrum at the Mobile Station," Vehicular Technology, IEEE Transactions on, pp. 425-434, 2002

Winfree Sean, "Angle of Arrival Estimation using Received Signal Strength with Directional Antennas," Engineering Honors Theses, 2010.

T.N. Lin, Po-Chian Lin, "Performance comparison of indoor positioning techniques based on location fingerprinting in wirelessnetworks", in Proc. WIRLES'05, 2005, pp. 1469-1574, vol. 2.

Quentin H. Spencer, Brian D. Jeffs, Michael A.Jensen, A. Lee Swindlehurst, "Modeling the Statistical Time and Angle of Arrival Characteristics of an Indoor Multipath Channel ", SelectedAreasinCommunications,IEEEJournalon,vol.18,pp.347-360,2 002.

Richard B. Ertel, Paulo Cardieri, "Overview of Spatial Channel Models for Antenna Array Communication Systems, " Personal Communications, IEEE, vol. 5, pp. 10-22, Feb 1998.

N.J. Thomas, D.G.M. Cruickshank, D.I. Laurenson, "PERFORMANCE OF A TDOA-AOA HYBRID MOBILE LOCATION SYSTEM", in Proc. 3G Mobile Communication Technologies,2001, pp. 216-220, 2001

Kenichi Kobayashi, Takao Someya, Tomoaki Ohtsuki, Sigit PW Jarot, Tsuyoshi Kashima, "MIMO System with Relative Phase Difference Time-Shift Modulation in Rician Fading Environments", IEICE Transaction on Communications, Vol E91.B, Issue 2, pp.459-465, 2010

Joint MMSE-FDE & Spectrum Combining for Antenna Diversity Reception on Single-carrier Transmission

Tatsunori OBARA† and Fumiyuki ADACHI‡

Dept. of Electrical and Communication Engineering, Graduate School of Engineering, Tohoku University
6-6-05, Aza-Aoba, Aramaki, Aoba-ku, Sendai, 980-8579, JAPAN
E-mail: †obara@mobile.ecei.tohoku.ac.jp ‡adachi@ecei.tohoku.ac.jp

***Abstract*—Frequency-domain equalization (FDE) based on minimum mean square error (MMSE) criterion is a powerful equalization technique for the broadband single-carrier (SC) transmission. However, the presence of timing offset produces the inter-symbol interference (ISI) and degrades the bit error rate (BER) performance. As the roll-off factor of the transmit filter increases, the performance degrades more. Recently, we proposed joint MMSE-FDE & spectrum combining which can achieve the frequency diversity gain while suppressing the negative impact of timing offset for the SC transmission. In this paper, we extend the joint MMSE-FDE & spectrum combining to include the antenna diversity reception.**

Keywords; Frequency-domain equalization, Nyquist filter, oversampling, timing offset, antenna diversity reception, single-carrier transmission

I. INTRODUCTION

The broadband wireless channel is composed of many propagation paths with different time delays and the strong frequency-selective fading channel is produced [1]-[3]. Therefore, the bit error rate (BER) performance of the broadband single-carrier (SC) transmission degrades due to the strong inter-symbol interference (ISI). The use of the frequency-domain equalization (FDE) based on the minimum mean square error (MMSE) criterion can improve significantly the BER performance [4]-[6]. This is only true in the case of no timing offset between a transmitter and a receiver.

In many spectrum-efficient wireless communication systems, a square-root Nyquist filter is used at the transmitters to limit the signal bandwidth and the same filter at the receivers. However, the presence of timing offset between a transmitter and a receiver produces the ISI and degrades the BER performance as the roll-off factor of Nyquist filter increases as shown in Fig. 1. The reason for this performance degradation is that, when the received signal is sampled by the symbol rate, the received signal spectrum is distorted since adjacent frequency-shifted spectra are given different phase rotations and overlapped if the roll-off factor of Nyquist filter is larger than 0.

To solve the above problem, we proposed joint MMSE-FDE & spectrum combining [7]. The overlapping of spectra phase-rotated due to the timing offset can be avoided by 2-times oversampling. Therefore, when MMSE-FDE is applied to the oversampled received signal, the spectrum distortion due to the channel frequency-selectivity and the phase rotation due to the timing offset can be simultaneously compensated. MMSE-FDE and the spectrum combining are jointly performed to restore the ISI-free spectrum over the desired frequency range. The proposed MMSE-FDE can achieve a better BER performance as the filter roll-off factor increases. In our previous paper, we assumed the single antenna reception. The combination of MMSE-FDE and antenna diversity reception significantly improves the BER performance of the SC transmission [6]. In this paper, we extend the joint MMSE-FDE & spectrum combining to include the antenna diversity reception.

The remainder of this paper is organized as follows. Section II presents the system model of the proposed MMSE-FDE. The computer simulation results are discussed in Sect. III. Section IV offers some conclusions.

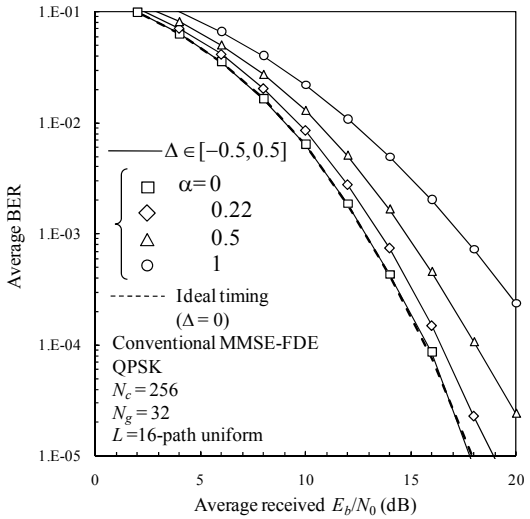


Fig. 1 BER performance of the conventional MMSE-FDE in the presence of timing offset.

II. JOINT MMSE-FDE & SPECTRUM COMBINING WITH ANTENNA DIVERSITY

In Fig. 2, the receiver structure of the SC transmission using the proposed joint MMSE-FDE & spectrum combining is illustrated. First, the received signal of each receive antenna is oversampled at a faster rate than the symbol rate to avoid the spectrum overlapping. When the square-root raised cosine filter is used as transmit filter, the spectrum overlapping can be

avoided by using double oversampling (the received signal sampled at the rate $2/T_s$), as shown in Fig. 3. Then, MMSE-FDE is applied over the frequency range of $-N_c \le k < N_c$ to simultaneously compensate for both the phase rotation due to the timing offset and the spectrum distortion due to the channel frequency-selectivity. Finally, the spectrum combining (or the frequency-domain down sampling) and antenna diversity combining are performed to recover the desired signal spectrum over the frequency range of $-N_c/2 \le k < N_c/2$.

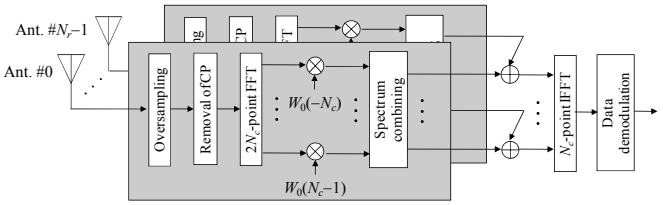


Fig. 2 Receiver structure of SC transmission using the proposed joint MMSE-FDE & spectrum combining.

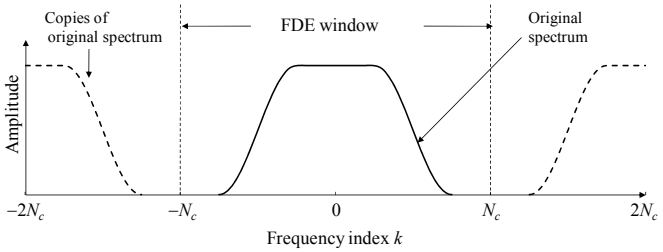


Fig. 3 Signal spectrum after double oversampling.

A. Signal representation

At the transmitter, the data-modulated symbol sequence is divided into a sequence of N_c-symbol blocks, where N_c is the size of fast Fourier transform (FFT). An N_g-symbol cyclic prefix (CP) is inserted into the guard interval (GI) of each symbol block. The GI-inserted symbol block is transmitted after passing through the square-root Nyquist transmit filter to limit the signal bandwidth.

The transmitted symbol block is received at the receiver via a frequency-selective fading channel. The n(=0~N_r−1)th antenna received signal oversampled at the rate $2/T_s$ can be expressed as

$$r_n(i) = \sqrt{\frac{2E_s}{T_s}} \sum_{l=0}^{L-1} \sum_{i'=-\infty}^{\infty} h_{n,l}\, s(i' \bmod N_c)\, \varphi\left(\frac{i}{2} + \Delta_n - \tau_l - i'\right), \\ + \nu(i) + \eta(i) \tag{1}$$

where E_s is the symbol energy, $h_{n,l}$ and τ_l are respectively the complex-valued channel gain with $\sum_{l=0}^{L-1} E[|h_{n,l}|^2] = 1$ and delay time of l-th path, $\{s(m);m=0\sim N_c-1\}$ is the transmitted symbol block, $\nu(i)$ and $\eta(i)$ are respectively the inter-block interference (IBI) and the filter output of the additive white Gaussian noise (AWGN) with zero mean and variance $2N_0/T_s$ with N_0 being the single-sided power spectrum density, $\varphi(t)$ is the transmit filter impulse response, and Δ_n is the timing offset. In this paper, we assume that the root raised cosine filter with the roll-off factor α is used as the transmit filter.

After the removal of $2N_g$-sample CP, $2N_c$-point FFT is applied to transform the oversampled signal block $\{r_n(i);i=0\sim 2N_c-1\}$ of the nth receive antenna into the frequency-domain signal $\{R_n(k);k=-N_c\sim N_c-1\}$. The kth frequency component $R_n(k)$ can be expressed as

$$R_n(k) = \frac{1}{\sqrt{2N_c}} \sum_{i=0}^{2N_c-1} r_n(i) \exp\left(-j2\pi k \frac{i}{2N_c}\right) \\ = \sqrt{\frac{2E_s}{T_s}} \tilde{H}_n(k,\Delta_n) S(k) + \mathrm{N}_n(k) + \Pi_n(k) \tag{3}$$

where $\tilde{H}_n(k,\Delta_n)$, $S(k)$, $\mathrm{N}_n(k)$, and $\Pi_n(k)$ are the overall (transmit/receive filter + channel) transfer function, the signal component, the IBI component, and the noise component, respectively. $\tilde{H}_n(k,\Delta_n)$ and $S(k)$ are respectively given as

$$S(k) = \frac{1}{\sqrt{N_c}} \sum_{i=0}^{N_c-1} s(i) \exp\left(-j2\pi k \frac{i}{N_c}\right) \tag{4}$$

$$\tilde{H}_n(k,\Delta_n) = \sqrt{2} \sum_{p=-\infty}^{\infty} H_n(k-2pN_c)\Phi(k-2pN_c) \\ \times \exp\left\{j2\pi(k-2pN_c)\frac{\Delta_n}{N_c}\right\}, \tag{5}$$

where $H_n(k)$ is the channel gain at the kth frequency given as

$$H_n(k) = \sum_{l=0}^{L-1} h_{n,l} \exp\left(-j2\pi k \frac{\tau_l}{N_c}\right). \tag{6}$$

$\Phi(k)$ is the transfer function of the transmit filter given as

$$\Phi(k) = \begin{cases} 1, & 0 \le \left|\frac{k}{N_c}\right| \le \frac{1-\alpha}{2} \\ \cos\frac{\pi}{2\alpha}\left(\left|\frac{k}{N_c}\right| - \frac{1-\alpha}{2}\right), & \frac{1-\alpha}{2} \le \left|\frac{k}{N_c}\right| \le \frac{1+\alpha}{2} \\ 0, & \text{elsewhere} \end{cases}. \tag{7}$$

$\tilde{H}_n(k,\Delta_n)$ can be estimated by using pilot-assisted channel estimation [8]-[10]. From Eq. (5), it can be understood that the copies of the original spectrum are phase-rotated due to the timing offset and frequency-shifted by an integer multiple of $2/T_s$.

B. Joint MMSE-FDE & spectrum combining

One-tap MMSE-FDE is performed over the frequency range of $-N_c \le k < N_c$ to simultaneously compensate for both

the phase rotation due to the timing offset and the spectrum distortion due to the channel frequency-selectivity as

$$\begin{aligned}\hat{R}_n(k) &= R_n(k)W_n(k) \\ &= \sqrt{\frac{2E_s}{T_s}}\hat{H}_n(k,\Delta)S(k) + \hat{\mathrm{N}}_n(k) + \hat{\Pi}_n(k)\end{aligned}, \quad (8)$$

where $W_n(k)$ is the MMSE-FDE weight.

After MMSE-FDE, the spectrum combining and antenna diversity combining is performed to restore the ISI-free condition over the desired frequency range $-N_c/2 \le k < N_c/2$ as shown in Fig. 4. The frequency-domain signal after the spectrum combining and antenna diversity combining is given by

$$\begin{aligned}\breve{R}(k) &= \sum_{n=0}^{N_r-1}\sum_{q=-1}^{1}\hat{R}_n(k-qN_c) \\ &= \sqrt{\frac{2E_s}{T_s}}\breve{H}(k)S(k) + \breve{\mathrm{N}}(k) + \breve{\Pi}(k)\end{aligned}, \quad (9)$$

where

$$\begin{cases}\breve{H}(k) = \sum_{n=0}^{N_r-1}\sum_{q=-1}^{1}\hat{H}_n(k-qN_c,\Delta_n) \\ \breve{\mathrm{N}}(k) = \sum_{n=0}^{N_r-1}\sum_{q=-1}^{1}\hat{\mathrm{N}}_n(k-qN_c) \\ \breve{\Pi}(k) = \sum_{n=0}^{N_r-1}\sum_{q=-1}^{1}\hat{\Pi}_n(k-qN_c)\end{cases}. \quad (10)$$

The frequency-domain signal $\{\breve{R}(k); k = -N_c/2 \sim N_c/2-1\}$ after MMSE-FDE and spectrum combining is transformed by N_c-point IFFT into the time-domain signal block for succeeding data demodulation. Below, we derive the MMSE-FDE weight.

We define the equalization error $e(k)$ after the spectrum combining at the kth frequency as

$$\begin{aligned}e(k) &= \breve{R}(k) - \sqrt{\frac{2E_s}{T_s}}S(k) \\ &= \sum_{n=0}^{N_r-1}\sum_{q=-1}^{1}\hat{R}_n(k-qN_c) - \sqrt{\frac{2E_s}{T_s}}S(k)\end{aligned}, \quad (11)$$

where $-N_c/2 \le k < N_c/2$. The MMSE weight $\{W_n(k); k=-N_c \sim N_c-1\}$ for joint FDE & spectrum combining which minimizes the MSE $E[|e(k)|^2]$ can be derived as

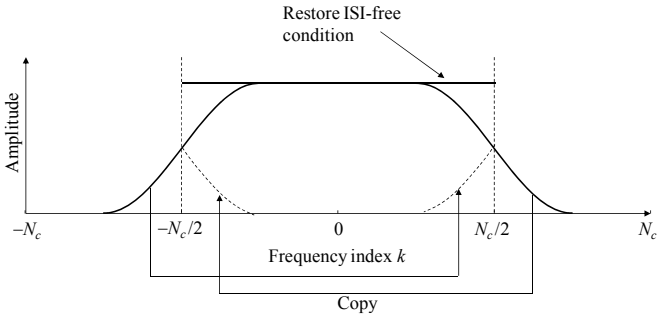


Fig. 4 Spectrum combining.

$$W_n(k) = \frac{\tilde{H}_n^*(k,\Delta_n)}{\sum_{n'=0}^{N_r-1}\sum_{q=-1}^{1}\frac{\Lambda_n^{-1}(\Delta_n)}{\Lambda_{n'}^{-1}(\Delta_{n'})}\left|\tilde{H}_{n'}(k-qN_c,\Delta_{n'})\right|^2 + \Lambda_n^{-1}(\Delta_n)}, \quad (12)$$

where $\Lambda_n(\Delta_n)$ denotes the signal-to-IBI plus noise power ratio (SINR) given by

$$\Lambda_n(\Delta_n) = \frac{(2E_s/T_s)}{E[|\mathrm{N}_n(k)|^2 + |\Pi_n(k)|^2]}. \quad (13)$$

Since $\tilde{H}_n(k,\Delta_n)$ includes the transfer function of the transmit filter, the MMSE-FDE weight also takes a role of the receive filter matched to the transmit filter (this is the reason why no receive filter is necessary in the receiver structure of Fig. 2).

III. Computer Simulation

A. Simulation condition

The computer simulation condition is summarized in Table I. We assume QPSK data-modulation, a signal block length of N_c=256 symbols, and a CP length of N_g=32 symbols. The propagation channel is assumed to be L=16-path frequency-selective block Rayleigh fading channel having uniform power delay profile. The receiver has N_r(=2, 4) antennas. The timing offset Δ_n normalized by the symbol duration T_s is assumed to be uniformly distributed over [–0.5, 0.5] for n=0~N_r–1. The ideal channel estimation is also assumed.

TABLE I. Simulation condition

Data modulation	QPSK	
Block length	N_c=256	
CP length	N_g=32	
Channel model	Frequency-selective block Rayleigh fading	
	Power delay profile	L=16-path uniform
Transmit filter	Root raised cosine filter	
	Roll-off factor	α=0~1
Receiver	No. of receive antennas	N_r=2, 4
	Timing offset	$\Delta_n \in$[–0.5, 0.5] (n=0~N_r–1)
	Channel estimation	Ideal

B. *Impact of timing offset on the conventional MMSE-FDE*

Figure 5 plots the average BER performance of the conventional MMSE-FDE as a function of the average received bit energy-to-noise power spectrum density ratio $E_b/N_0(=0.5(E_s/N_0)(1+N_g/N_c))$ per receive antenna. For comparison, the BER performance for N_r=1 and Δ_0=0 (no timing offset case) is also plotted.

As shown in Fig. 5, as N_r increases, the BER performance significantly improves compared to the case of N_r=1. However, when the timing offset is present, the BER performance degrades as α increases. This is because, as α increases, the overlapping interval of adjacent spectra which are phase-rotated due to the timing offset becomes wider, thereby enhancing the spectrum distortion. When α=0, the conventional MMSE-FDE in the presence of timing offset achieve almost the same performance as in the no timing offset case. This is because phase-rotated adjacent spectra do not overlap and therefore, no spectrum distortion is produced.

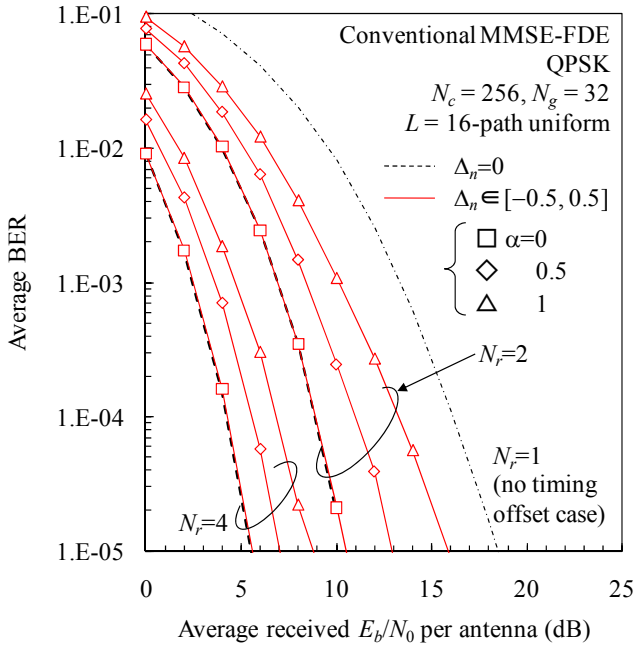


Fig. 4 Impact of timing offset on the conventional MMSE-FDE.

C. *BER performance of the proposed joint MMSE-FDE & spectrum combining*

Figure 6 plots the BER performance of the proposed joint MMSE-FDE & spectrum combining with α and N_r as parameters. For comparison, the BER performances for N_r=1 are also plotted.

The proposed joint MMSE-FDE & spectrum combining can achieve almost the same performance as in the no timing offset case irrespective of N_r and α. Furthermore, the performance of joint MMSE-FDE & spectrum combining improves as N_r and α increase. The reason for this is that, as the signal bandwidth becomes wider, the proposed joint MMSE-FDE & spectrum combining with antenna diversity can achieve increased frequency diversity gain due to joint equalization and frequency and antenna diversity combining.

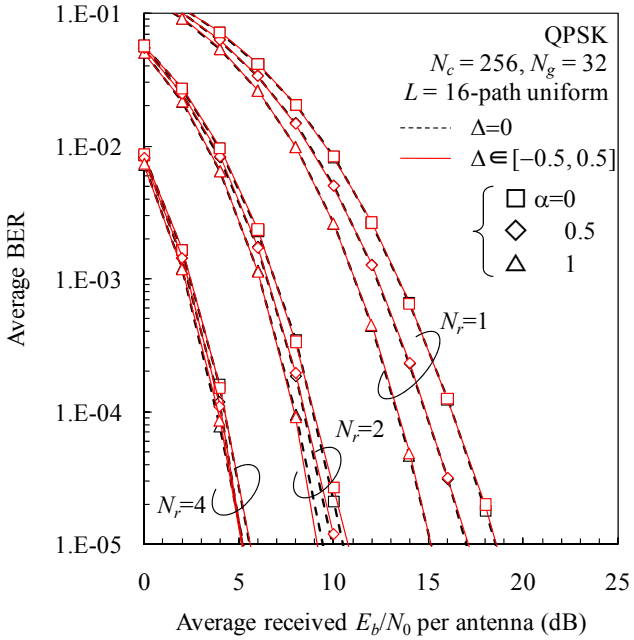


Fig. 4 BER performance of joint MMSE-FDE & spectrum combining.

IV. Conclusion

In this paper, we extended the joint MMSE-FDE & spectrum combining to include the antenna diversity reception. The joint MMSE-FDE & spectrum combining can provide almost the same performance as in the no timing offset case and much better performance due to larger frequency diversity gain as the filter roll-off factor and the number of receive antennas increase.

References

[1] W. C., Jakes Jr, Ed, *Microwave mobile communications*, Wiley, Newyork, 1974.

[2] J. G. Proakis, *Digital communication*, 4th ed., McGraw-Hill, 2001.

[3] Y. Akaiwa, Introduction to digital mobile communication, Wiley, Newyork, 1997.

[4] D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar and B. Eidson, "Frequency domain equalization for single-carrier broadband wireless systems," IEEE Commun. Mag., Vol. 40, No. 40, pp.58-66, Apr. 2002.

[5] F. Adachi, T. Sao, and T. Itagaki, "Performance of multicode DS-CDMA using frequency domain equalization in a frequency selective fading channel," IEE Electronics Letters, vol. 39, No.2, pp. 239-241, Jan. 2003.

[6] F. Adachi and K. Takeda, "Bit error rate analysis of DS-CDMA with joint frequency-domain equalization and antenna diversity combining," IEICE Trans. Commun.,Vol. E87-B, No. 10, pp. 2991-3002, Oct. 2004.

[7] T. Obara, K. Takeda and F. Adachi, "Joint MMSE-FDE & spectrum combining for a broadband single-carrier transmission in the presence of timing offset," IEICE Trans. Commun., Vol. E94-B, No. 5, May 2011.

[8] H. Ando, M. Sawahashi and F. Adachi, "Channel estimation filter using time-multiplexed pilot channel for coherent Rake combining in DS-CDMA mobile radio," IEICE Trans. Commun., Vol. E81-B, No. 7, pp. 1517-1526, July 1998.

[9] S. Takaoka and F. Adachi, "Pilot-aided adaptive prediction channel estimation in a frequency-nonselective fading channel," IEICE Trans., Commun., Vol. E85-B, No. 8, pp. 1552-1560, Aug. 2002.

[10] K. Takeda and F. Adachi, "Frequency-domain MMSE channel estimation for frequency-domain equalization of DS-CDMA signals," IEICE Trans., Commun., Vol. E90-B, No.7, pp. 1746-1753, July 2007.

Design and Performance of a Lumped Resistively Loaded Spiral Antenna

Meng Zou , Jin Pan , Meng Quan
Department of Microwave Engineering
University of Electronic Science and Technology of China
Chengdu, China
E-mail: zm.1598@163.com

Le Zuo
Chengdu SEEKON Microwave Communications CO., Ltd.
Chengdu, China

***Abstract*—A novel spiral antenna with miniature configuration is proposed by introducing lumped resistive loading along the spiral arms. A transmission-line model incorporating the loss resistance is introduced to help study the method. Numerical simulation examples have been demonstrated that antenna size could be reduced significantly while maintaining the axial ratio characteristics. A partly loading structure was investigated to raise the gain at lower frequencies. The designed antenna was implemented and measured, which achieved a size reduction of 58%.**

Keywords-spiral antenna; miniaturization; resistive loading; axial ratio

I. INTRODUCTION

Circularly polarized antennas have significant applications in wireless communication, especially in satellite communications. Spiral antennas have attracted significant attention since their introduction in the late 1950s because of their circular polarization and wide bandwidth [1]-[2]. For the typical wire spiral antenna, the above mentioned advantages come at the price of antenna size. Three conditions should be satisfied to radiate circular polarization wave from spiral antennas at one frequency, 1) there must be one turn in the spiral whose length is equal to the current wavelength; 2) the current is lossless on this turn; 3) current on the arms should be travelling wave. Due to the three conditions the aperture of traditional circular spiral is limited by $D=\lambda_{e\max}/\pi$ [2] and the physical size at the low-frequency becomes too large for practical applications. Therefore, the concept of spiral antenna miniaturization was investigated quite intensively in recent years. In previous researches, most work on miniaturizing spiral antennas attempted to meet condition 1and condition 3 assuming that condition 2 was always satisfied. There are two ways to meet condition 1, one is to increase the physical length of the turn, and the other is to reduce the guided wave's wavelength within the antenna structure. A proper termination is useful to suppress the travelling wave reflection from the end of the spiral [3]. A square spiral or planar zigzag/ meandrous spiral rather than a traditional circular spiral can be used to increase physical length [4], while dielectric loading [5] and distributed/lumped reactive loading (adding distributed/lumped reactive components to the conducting arms) [6]-[8] were investigated to reduce the wavelength. All of these methods have performed well in reducing antenna size. However they were designed under the assumption that there was no transmission loss on the arms. In lossless transmission-line model the wavelength only concerned with circuit elements (C, L) at one frequency, where C and L refer to the unit-length capacitance and inductance. Meanwhile it is also proportional to R in lossy transmission-line model. This means that one can reduce the wavelength by introducing lumped resistive loading on the arms, however, the introducing of resistive loading destroys condition 2. So whether the antenna miniaturization can be achieved while maintaining the circular polarization properties is a question which is worth further researching.

The preliminary assessment of the above mentioned method is presented in this paper. Then a 96mm diameter spiral operating at 420MHz-500MHz with AR (axial ratio) < 4 and VSWR < 2:1 is optimized based on simulation and measurement.

The remaining structure of the paper is as follows. Section II presents the theoretical basis of the method and the verification by simulation (using HFSS). An example takes advantage of this method for antenna miniaturization is introduced in Section III followed by the conclusion drawn in Section Ⅳ.

II. ANTENNA MINIATURIZATION USING LUMPED RESISTIVE LOADING APPROACH

Most spirals were investigated under the assumption that there was no loss on the arm, therefore a lossless transmission line model could be used to help study the antenna. In the context of the lossless transmission line model we have

$$\lambda = \frac{2\pi}{\omega\sqrt{LC}}$$

The correspondences between the material parameter (ε, μ) and the circuit elements (L, C) are $\varepsilon \leftrightarrow L$ and $\mu \leftrightarrow C$ [6].This is the reason why previous antenna miniaturization was focused on increasing the equivalent permittivity and permeability to reduce the wavelength. If

taking the transmitting loss into consideration a transmission line model with a loss resistance R is employed. Now the wave is decided by

$$\beta=\omega\sqrt{\frac{LC}{2}*\left[\sqrt{1+\left(\frac{R}{\omega L}\right)^2}+1\right]}$$

$$\partial=\omega\sqrt{\frac{LC}{2}*\left[\sqrt{1+\left(\frac{R}{\omega L}\right)^2}-1\right]}$$

$$\lambda=\frac{2\pi}{\beta}$$

This indicates that the wavelength is inversely proportional to the loss resistance in transmission line model. The attenuation constant is increased when reducing the spiral size by raising the loss resistance, this will do harm to the gain of antenna but its effect on AR is uncertain. Lumped resistive loading spirals were developed to investigate it. For these configurations the following geometrical parameters apply (presented in Fig. 1): Archimedean growth rate $15.4mm/rad$, maximum growth angle $\phi_{\max}=6\pi$, and arm width $w=2mm$ and aperture dimensions of $96\times96mm$. The circular spiral was printed on a 2mm thick FR4 substrate ($\varepsilon_r=4.2,\tan\delta_\varepsilon=0.02$). The arms were populated with resistive loads placed approximately $10mm<\lambda_{\min}/10$ apart resulting in 33 per arm.

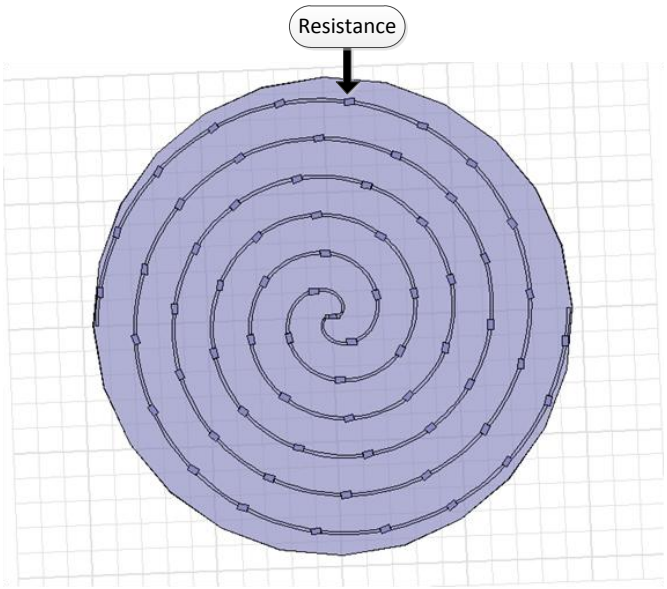


Figure 1. Structure of lumped resistive loaded spiral antenna

The antenna performance of AR and Gain between 100MHz-2GHz of the antennas with various R are shown in Fig. 2.

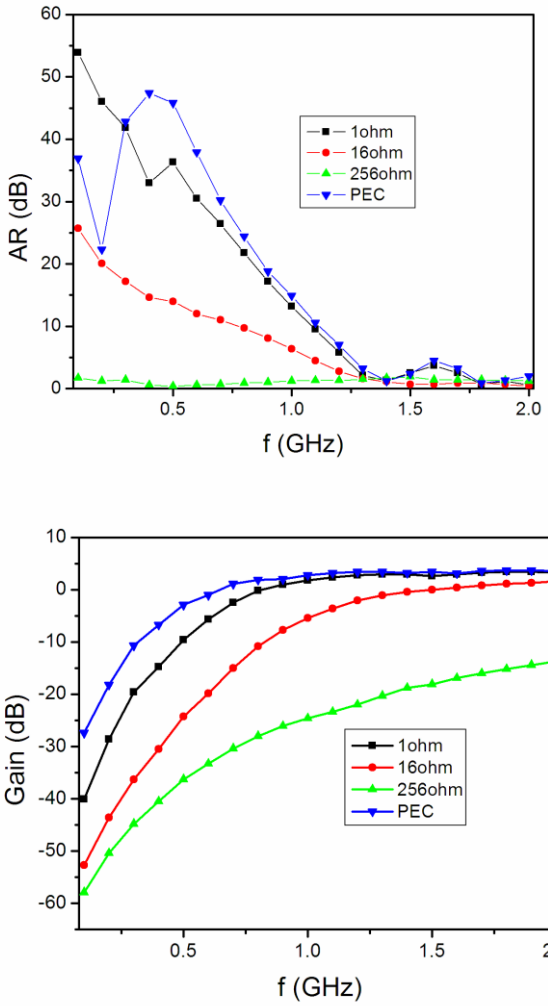


Figure 2. Simulation results of an spiral antenna with different resistive loadings.

From Fig. 2 it is observed that the AR of spiral can be reduced when doing antenna miniaturization by loading lumped resistance and the impedance increases with higher miniaturization. However, the gain of these configurations decreasing seriously when $f<1\text{GHz}$, this maybe results from three reasons：1) the spiral is electrically small and can be thought of as a small dipole antenna when loading lightly, 2) loss caused by the resistances becomes large when loading heavily, 3) the return loss worsened when loading heavily. For raising the gain at low frequencies we can use lumped resistive loading only at the outer region of the spiral. By this method the performance at middle frequencies will be deteriorated, however, it is verified practical when the operating frequency band is narrow.

III. Example of Antenna Miniaturization Using the Partly Loading Concept

A 420-500MHz spiral with 96mm diameter was devised taking use of the method motioned above. The geometrical parameters of the substrate were same with the ones in Section

Ⅱ except that the resistances were distributed along the last quarter arm and the number of them was reduced to 15 per arm (shown in Fig.3).

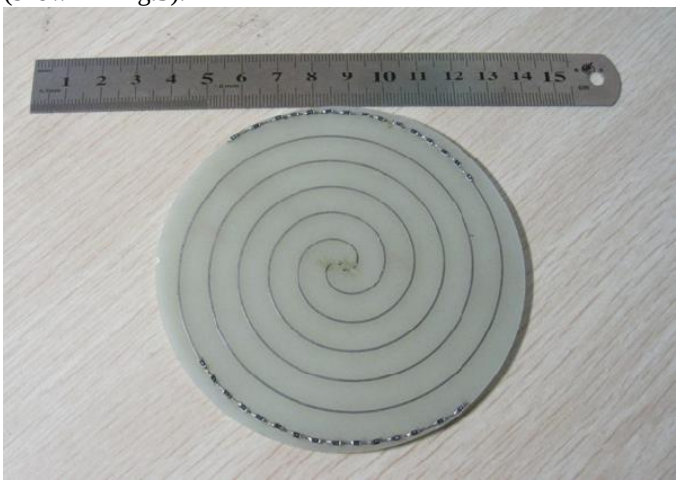


Figure 3. Partly resistive loaded spiral antenna

Two 35mm thick absorber with unknown parameters, which made it more difficult to design, and a balun converting 50Ω input to 188Ω output were employed(shown in Fig.4). The entire antenna is enclosed by a plastic cavity. Fig. 5 is the photo of the fabricated spiral antenna.

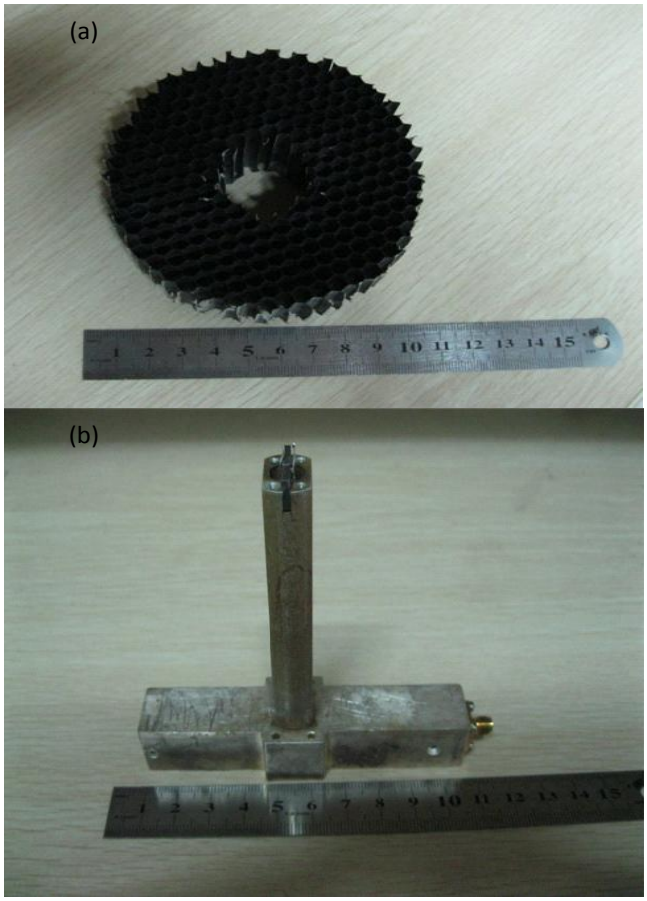
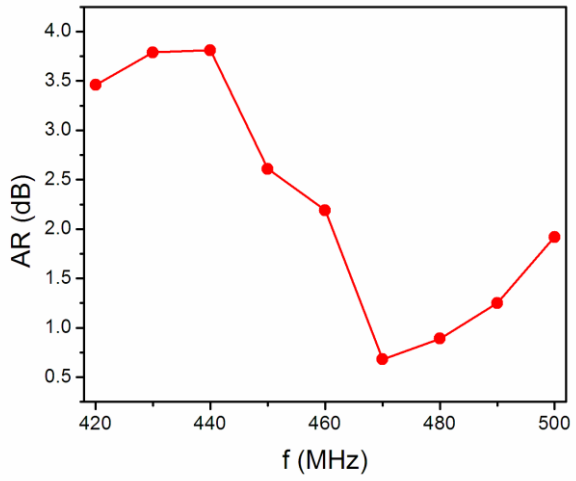


Figure 4. (a) Absorber with unknown parameters (b) balun employed in the antenna

Although the high dielectric constant of the absorber benefited the size reducing, it increased the difficulty of designing due to its unknown parameters. The iteration of simulation and measurement was used to estimate the equivalent parameters concerned with AR, which is a high light in the investigation, beyond the scope of this letter and will be discussed in another paper. The antenna was significantly affected by the existing balun, because it had a lower frequency limit at 420MHz and an insufferable height of 80mm which prevented from reducing antenna size. A set of resistances were obtained to meet the specification by optimizing, the resistances along the arm from inside to outside were 8.2, 10, 15, 18, 24, 30, 33, 20, 24, 27, 30, 33, 36, 39, 43(Ω).

Figure 5. Photo of the fabricated spiral antenna

The bore-sight axial ratio, bore-sight gain and voltage standing-wave ratio of this antenna are presented in Fig. 6, with the measurements done at the anechoic chamber in UESTC.



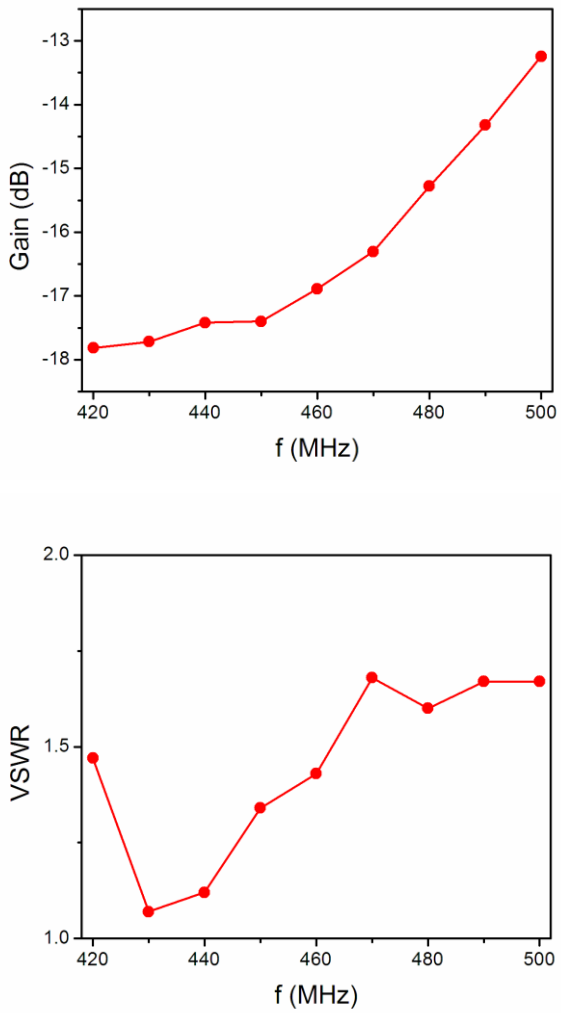


Figure 6. Measured results of a partly resistive loaded spiral antenna.

From Fig.6 antenna's axial ratio of AR < 4 was achieved throughout the band of 420-500MHz which needed a 227mm diameter for traditional circular spiral. The gain of the antenna was significantly raised by using partly loading and the employing of a balun.

IV. CONCLUSION

The use of lumped resistive loading for antenna miniaturization while maintaining axial ratio characteristics is presented in this paper. A novel lossy transmission line model with loss resistance *R* was introduced to help understand this method. Simulation and experiment provided a demonstration of this concept. Simulation of full lumped resistive loading show it can indeed keep the AR<3dB at or even lower than 100MHz when resistances were properly chosen. This is at the price of sacrificing the gain of antenna. A spiral antenna with 96 mm diameter was designed to operate at 420-500 MHz and a more than 58% size reduction was achieved. The measurement results show that excellent circular polarization property is achieved and the gain was significantly raised by partly loading instead of fully loading.

We conclude the paper by noting that better performance can be achieved at low frequency by optimizing the topology of resistive loading. And the improvement of the balun is also needed to reduce the size of antenna at lower frequencies.

REFERENCES

[1] J. D. Dyson, "The equiangular spiral antenna," IRE Trans. Antennas Propagat., vol. AP-7, pp. 181-187,Apr. 1959.

[2] J. Kaiser, "The Archimedean two-wire spiral antenna," IRE Trans. Antennas Propagat., vol. 8, pp. 312–323, May 1960.

[3] M. W. Nurnberger and J. L. Volakis, "New termination and shallow reflecting cavity for ultra wide-band slot spirals," in Proc. IEEE APS Symp., vol. 3, 2001, pp. 1528-531.

[4] D. S. Filipovic and J. L. Volakis, "Broadband meanderline slot spiral antenna," Inst. Elect. Eng. Proc. Microw., Antennas Propag., vol. 149, no. 2, pp. 98-105, Apr. 2002.

[5] B. A. Kramer, M. Lee, C.-C. Chen, and J. L. Volakis, "Design and performance of an ultrawide-band ceramic-loaded slot spiral," IEEE Trans. Antennas Propag., vol. 53, no. 7, pp. 2193–2199, 2005.

[6] M. Lee, B. A. Kramer, C. Chen and J. L. Volakis, "Distributed lumped loads nd lossy transmission line model for wideband spiral antenna miniaturization and characterization," IEEE Transactions on Antennas and Propagation, Vol.55, No.10, pp.2671-2678, Oct 2007.

[7] I. Tzanidis, C.-C. Chen, and J. L. Volakis, "Smaller uwb conformal antennas for VHF/UHF applications with ferrodielectric loadings," in Proc.IEEE, Antennas and Propagation Society Int. Symp., Jul. 2008, pp. 1-4.

[8] I. Tzanidis, Chi-Chih Chen, and John L. Volakis, "Low Profile Spiral on a Thin Ferrite Ground Plane for 220–500 MHz Operation," IEEE Transactions on Antennas and Propagation, Vol.58, No.11, pp.3715-3720, Nov 2010.

Speech and Hand Gesture Fusion Interface for the Control of Home Automation System

Joo-Gon KIM
Department of Information and Communication Engineering
Yeungnam University
Kyongsan-City, Korea
speech@ynu.ac.kr

Soo-Young SUK
Department of Research Development
Gyeongbuk Institute of IT Convergence Industry Technology
Kyongsan-City, Korea
sysuk@gitc.or.kr

Hyun-Koo KIM , Kook-Yeol YOO, Hyun-Yeol CHUNG

Department of Information and Communication Engineering
Yeungnam University
Kyongsan-City, Korea
kim-hk@ynu.ac.kr, {kyoo, hychung}@yu.ac.kr

Abstract— **In this paper, we present a speech and hand gesture fusion recognition interface for the control of home automation systems. The investigation of this recognition has been the focus of research about multimodal interfaces for speech recognition with real-time hand gesture recognition using 3D depth camera. Voice command recognition is commercially successful systems for the home automation control. However, continuous voice command system produces poor performance due to the misrecognition of the non-command voice, noise and conversational voice. In this system, as auxiliary information hand gestures are used for voice activity detection (VAD) in the speech recognition. With the improved VAD performance, the proposed system has robust performance under noisy environment. Our gesture recognition process was performed Support Vector Machine (SVM) classifier method and tracking method using Kalman-filter. Further, our speech recognition system uses multi-path HMnet model with the Julius system.**

Keywords- Home automation system; multi-path HMnet model; speech and hand gesture fusion recognition; voice activity detection

I. Introduction

Hand gestures and speech constitute one of the most important modalities. The investigation of this phenomenon has been the focus of research about multimodal interfaces for speech and hand gesture recognition in Human Machine Interaction (HMI). Recently, speech recognition system is widely used in navigation and smart phone because of that easy handling is possible. Speech recognition system needs simple and reliable operation that when people watch TV, in the normal condition state, voice input has to be stopped and in the voice input state, has to reduce environment noise and recognize speech. In the past, speech recognition system use button control to start action of recognition. In the control of home appliances, people should be able to input voice to maximize the convenience of speech recognition anywhere and anytime without additional method such as button device. However speech recognition system is mistaken on the fly mode, because of background noise, conversations between users, noise of TV and radio, and etc. Therefore, various non-voice rejection methods have been proposed for real-environment speech recognition

For reject function of abnormal voice input, previous methods have been researched as follows:

- Using speech recognition result with confidence measure in the post-processing [1]
- Using predefined control commands recognition in the post-processing
- Using probability difference in the post-processing
- Distinguish voice and non-voice using pitch information in the preprocessing [2]
- Limit voice input area using microphone array [3]

Multimodal system use hand-gesture recognition to independent control method rather than additional information method for speech recognition. So, proposed system is able to implement robust voice activate detection (VAD) and to describe degree of size information.

Speech recognition system is able to quickly control using large vocabulary word recognition but it is difficult to detect degree and amount of volume and time bar control. Hand-gesture recognition system is able to implement HMI system satisfying immediate and continuity. Using speech and hand-gesture, we can implement HMI to prevent error in the real world environment and to satisfy immediacy and continuity. There has been various research on hand gesture recognition. In general, hand gesture recognition system is divided into mono-vision system and stereo-vision system. Mono-vision system detects hands using mono-camera but, it depends on image brightness and user's location.

Stereo-vision system detects hands using depth map of stereo images but, it depends on image brightness. In particular, by using visual markings on the hands, previous researchers have recognized sign language and pointing gestures [4-6]. However, these methods require the placement of markers on the hands. The marking-free systems of [7, 8] can recognize specific finger or pointing events, but not general gestures. Employing special hardware or offline learning, several researchers have developed successful systems to recognize general hand gestures [9-13]. Blake and Isard [14] have developed a fast contour-based tracker in which they applied to hands, but the discrimination of different hand poses is limited. Extensive research has been conducted on hand gesture recognition for digital images [15, 16]. However, it's still ongoing research as most papers do not provide a complete solution to the previously mentioned problems. Range cameras can simultaneously capture a full 3D point cloud with an array sensor at video rates. They offer great potential for real-time measurement of static and dynamic scenes [17]. Investigation on 3D range cameras for automatic gesture recognition is in its infancy. However, some research has been conducted in this area [18, 19].

The final aim of this research is to design and build a human-machine interface using the 3D information provided by a range camera for automatic and real time identification of hand gesture. In this paper, we focus on two applications. The first one recognize the number of raised fingers that appear in a hand gesture and the second is intended for moving an object in a virtual environment using only a hand gesture on the acquired images. For a real-time application, the expectation is to obtain the best possible images of the hand gesture within the lowest possible time. Some experiments have been conducted with the purpose of defining the best configuration for imaging the hand. This configuration includes, the distance between the hand and the camera, the influence of the integration time of the camera, the amplitude threshold, and the lighting conditions of the environment, the surrounding objects and the skin color.

This paper is organized as follows. Section 2 introduces various non-command rejection methods. Next, Section 3 describes hand-gesture recognition method using 3D depth camera. In Section 4, we propose speech hand-gesture fusion multimodal system with multi-channel speech-tube recognition. Finally, we draw our conclusions in Section 5.

II. Non-Command Rejection Method

In the control of home appliances, people should be able to input voice to maximize the convenience of speech recognition anywhere and anytime without additional method such as button device. On the fly, however, speech recognition system mistakes because of background noise, conversations between users, noise of TV and radio. Therefore, it has been proposed a various ways to minimize the malfunction such as figure 1.

In general, preprocessing step of speech recognition for Voice Activation Detection (VAD) is to press the button to reduce the malfunction of abnormal speech input (Ravichandran et al., 2007). But, in the case of method 1, if a pitch frequency occurs for more than a certain amount of time, VAD was started and in the case of method 2, system recognizes user's position as nearness using camera, ultrasonic, and laser sensor, it was stared. In addition, it is proposed to recognize and estimate user-position using multi-microphone sensor and beam-forming method for VAD (method 4), method 5 is proposed to YIN pitch detection, which is entered by voice, a voice region by analyzing the reliability voice and nasal to determine availability.

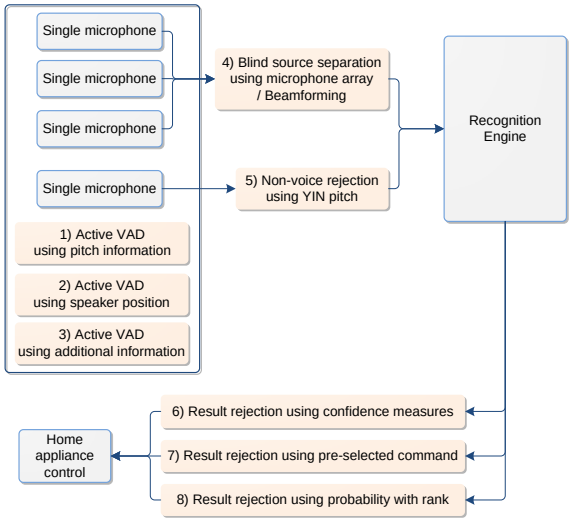


Figure 1. Various non-command rejection method

Typically, rejection function for the post-processing step has been used three methods such as speech recognition result (method 6), control commands recognition (method 7), and probability difference (method 8). But, when post-processing of method 6, 7, and 8 are used, if the absolute threshold is increased to prevent malfunction, rejection rate is also increased. And when pre-processing of method 2, 4, and 5 are used, it is difficult to distinguish control commands speech from general one. Therefore, this paper proposes robust recognition method of VAD starting condition using hand gesture for real world environment.

III. Hand-Gesture Recognithion

A. Hand-gesture Recognithon

After detection of the hand information from the captured scene, multiple step based recognition has been designed (Figure 5). First, hands are tracked by using Kalman filter. And then, the following step is the extraction of the hand gesture motion feature. To avoid a time-consuming segmentation on every acquired frame, tracking the hand gesture appears to be an appropriate alternative. The Kalman filter is a suitable tool designed for this purpose (Stenger et al., 2001). It is used to predict the position of the hand in the coming frame and after having measured the actual position of the hand in that frame, this prediction is corrected and the adjusted value is used for the prediction in the following frame. The Kalman filter is thus a recursive estimator for linear processes.

To track the hand gesture, we consider the centroid of the hand region. Then, we extract features in the following steps. A sequence of hand motion vectors are given by the hand-tracking module. Then we normalize the sequence by the normalization coefficient and quantize the selected motion vectors to make a feature to be used in gesture recognition. For input video, we can extract a feature as a string of symbols as describes above. Then we should classify the feature into one of the gesture classes. To accomplish this, we use a support vector machine (SVM) as a classifier, which is one of the best classifiers in general (Cortes et al., 1995). Our algorithm recognizes four types 3D hand gestures of push, pull, turn-right, and turn-left.

B. Position Tracking

Position tracking can provide improved results in noisy situations and generate other useful metrics important for the primary system objective. Kalman filtering provides a way to incorporate a linearized version of the system dynamics to generate optimal estimates under the assumption of Gaussian noise. Kalman filtering also provides estimates of state variables that are not directly observable, but may be useful for the system. It is important to have metrics such as rates of change of position robustly estimated. The Kalman filter [20] state variables are updated using the hand-position and angle estimates along with measurements of hand-moving-angle and moving velocity.

$$y_k = Mx_k \tag{1}$$

$$x_{k+1|k} = Ax_{k|k} + Bu_k \tag{2}$$

Where

$$x = \left[\phi, \dot{\phi} = \tan\theta, \ddot{\phi}, W\right]^T \tag{3}$$

$$A = \begin{bmatrix} 1 & v\Delta t & \frac{(v\Delta t)^2}{2} & 0 \\ 0 & 1 & v\Delta t & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4}$$

$$Bu_k = \left[0, \Phi\Delta t, 0, 0\right]^T \tag{5}$$

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{6}$$

These measurements are then used to update the discrete-time Kalman filter for the hand state. The system and measurement equations as well as the Kalman update equations at time are detailed in (1)–(6). The initial values for the estimation error covariance and state-noise covariance were determined by empirical testing. The measurement vector consists of the hand position, hand-moving angle, and hand size.

C. Motion Feature Extraction

After position tracking procedure, we extract motion feature by using vectors normalization and quantization method. By using vector quantization (VQ) we can transform a sequence of motion vectors into that of symbols (codewords). Vector quantization has been used in many applications such as image and voice compression and recognition. A vector quantizer consists of a codebook and a quantization function as follows:

$$M = \{Y_j \mid 1 \le j \le N_K\} \tag{7}$$

$$q : \{X_i\}_{i=1}^{N_X} \rightarrow M \tag{8}$$

where M is a set of codewords, called a codebook, q is a quantization function that associates each input data X_i's to a code. In the experiments, the quantization function q maps the motion vectors to the nearest codeword. To build the codewords, we perform the K-means algorithm to find K centroids using all four dimensional tracked training motion vectors of the concatenated left and right hand motion vectors. We implemented the K-means algorithm with $K = 20$ for vector quantization and assigned an integer number to each cluster. The size for code normalization was 10.

D. Gesture Recognition by SVM

To perform SVM , we should make sure that all features are in the same space so we need to make the size of features a constant length. In SVM, we used to the radial basis kernel. For multi-class classification we adopted the one-per-class (OPC) strategy to find class boundaries of all classes.

IV. Speech and Hand Gesture Fusion Multi-modal System

In general, speech and hand gesture recognition for multi-modal control system of information electronic appliance are used. Table 1 shows that voice recognition is able to quickly control using word and vocabulary recognition but it is difficult to detect degree and amount of volume and time bar control. Therefore, in this paper, it is proposed speech and gesture fusion multi-modal system for robust operating performance of information electronic appliance on the real world environment.

A. Speech Hand-gesture Fusion

A role of gesture recognition of proposed fusion recognition using 3D camera and microphone array is designed as follows:

Push/Pull gesture : start/stop VAD

Turning gesture: volume control and moving contents

First, when people open and hold out his hand (push action), VAD is started. On the contrary to this, when people hold in his hand (pull action), VAD is stopped. When voice is inputted, if this system recognizes action of VAD stop, VAD is stopped after continuous voice input. 3D camera uses to recognize the push and pull action and to reduce error. Therefore, in this system, gesture mode is operated when

speech recognition part detect voice commands of scroll action such as ‘volume’, ‘music list’, and etc.

B. Multi-channel Speech-tube Recognition System

Figure 2 shows block diagram of the proposed multi-channel recognition speech-tube recognition system. This system is composed of speech input control software module, independent recognition engine software module using socket, and analysis module of recognition results and each module is connected to tube.

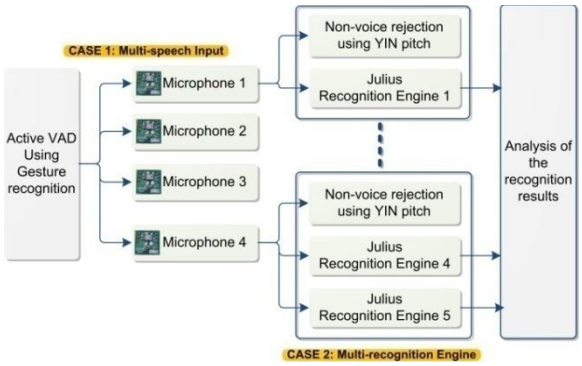


Figure 2. Block diagram of the multi-channel speech-tube recognition system

Features of this proposed system are received voice input to 16 channels and speech recognition is possible for each channel (case one). Moreover, in the part of case two, multiple recognition engines are used in individual voice. In other words, this system can receive voice input from several directions and can recognize multi-function and multi-vocabulary from one voice using multi-speech model. Figure 3 shows multi-channel recognition system, proposed system, for home automation. This system use independent four channel microphone arrays and control start TV, audio, VoIP telephone, and lamp.

Figure 3. A developed application for home automation

V. Conclusions and Future Works

This paper primarily discusses several key issues related to multimodal interfaces for speech and hand gesture recognition and presents developed fusion system with multi-channel speech-tube recognition. Hand-gesture of proposed system recognizes hand action to control start and stop conditions of VAD and to recognize degree of scroll such as volume control. Speech recognition uses Julius engine of multi-elements and speech-tube method using multi-channel microphone. In case of hand-gesture recognition, 3D depth perception improves performance, Kalman-filter is used, the direction vector recognition using SVM classifier were performed. We are planning to apply our method to wide area recognition and various real-time tests in real-environment.

References

[1] Lee A. and Kawahara T., “Recent Development of Open-Source Speech Recognition Engine Julius,” in Proc. APSIPA ASC. pp.131-137, 2009.

Suk S.Y., Kojima H., Chung H.Y., “Voice/Non-voice classification using reliable fundamental frequency estimator for voice activated wheelchair control,” LNCS vol. 4523, pp549-559, 2006.

Sasou A., “Head-orientation-estimation-integrated speech recognition for the smart-chair,” Proc. of 2nd International Symposium on Universal Communication (ISUC2008), pp.482-489, 2008.

Tartter V. C. and Knowlton K. C., “Perception of sign language from an array of 27 moving spots”. Nature, (239): pp. 676-678, Feb. 19, 1981.

Cipolla R., Okamoto Y., and Kuno Y., “Robust structure from motion using motion parallax”, In Proc. 4th Intl. Conf. Computer Vision, pp. 374 - 382, Berlin, Germany, 1993.

Dorner B., “Hand shape identification and tracking for sign language interpretation”, In Looking at people workshop, Chambery, France, 1993.

Fukumoto M., Mase K., and Suenaga Y., “Real-time detection of pointing actions for a glove-free interface”, In Workshop on Machine Vision Applications, Tokyo, 1992.

Rubine D. and McAvinney P., “Programmable finger-tracking instrument controllers”, Computer Music Journal, 14(1): pp. 26-41, 1990.

Segen J., “Gest: a learning computer vision system that recognizes gestures”, In Machine Learning IV. Morgan Kauffman, 1992. edited by Michalski et. al.

Ishibuchi K., Takemura H., and Kishino F., “Real time hand shape recognition using pipeline image processor”, In IEEE Intl. Workshop on robot and human communication, pp. 111-116. IEEE, 1992.

Darrell T. J. and Pentland A. P., “Space-time gestures”, In Proc. IEEE CVPR, pp. 335-340, 1993.

Davis J. and Shah M., “Gesture recognition”, Technical Report CS-TR-93-11, University of Central Florida, Orlando, FL 32816, 1993.

Rehg J. M. and Kanade T., “Digiteyes: vision-based human hand tracking”, Technical Report CMU-CS-93-220, Carnegie Mellon School of Computer Science, Pittsburgh, PA 15213, 1993.

Blake A. and Isard M., “3D position, attitude and shape input using video tracking of hands and lips”, In Proceedings of SIGGRAPH 94, pp. 185-192, 1994.

Tsuruta N., Yoshiki Y. and Tobely T.E., “A randomized hypercolumn model and gesture recognition”, 6th International Work-Conference on Artificial and Natural Neural Networks, Springer-Verlag, pp. 235-42, June 2001.

Mahesh R.J.K., Mahishi S., Dheeraj R., Sudheender S., Pujari S. and N.V., “Finger detection for sign language recognition”, International Association of Engineers pp. 489-93, March 2009.

Malassiotis S., and Strintzis M.G., “Real-time hand posture recognition using range data”, Image and Vision Computing, 26(7), 1027-37, 2008.

Liu X., and Fujimura K., “Hand gesture recognition using depth data”, Sixth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 529-534, May 2004.

Breuer P., Eckes C., and Muller S., “Hand gesture recognition with a novel IR time-of-flight range camera - a pilot study”, Proceedings, Springer-Verlag pp. 247-260, March 2007.

Kalman, R.E. “A new approach to linear filtering and prediction problems”. Journal of Basic Engineering 82 (1): 35–45, 1960.

Resonance Characteristics of Cylindrical Cavities utilizing Curved Resonant Gratings as a Mirror

Shohei IIJIMA, Yasuo OHTERA, and Hirohito YAMADA
Graduate School of Engineering
Tohoku University
Sendai, Japan
{iijima, ohtera, yamada}@ecei.tohoku.ac.jp

Abstract—**We demonstrated basic characteristics of cylindrical resonant cavities with micro-structured cavity walls. The cavity was surrounded by a circular dielectric waveguide with periodic rectangular corrugation. This periodic corrugation produces a function called Guided-Mode Resonance (GMR). GMR exhibits very high reflectivity at a resonance wavelength. Fundamental characteristics of the resonant grating mirror and the cavity were investigated through numerical simulation. Results obtained indicate that such a structure can be useful for the construction of large area single-mode resonators.**

Keywords- guided-mode resonance; grating; cylindrical resonator; quality factor; single-mode; FDTD

I. INTRODUCTION

High quality factor (Q) optical cavity is a fundamental building block for a number of optical devices from communication to sensing fields, such as wavelength filters, biosensors, and gas sensors. To construct such cavities, low-loss and highly reflective mirror structures with large degree of freedom in performance designing are needed. For this propose, structures that make use of Bragg reflection phenomena in dielectric periodic structures, such as multilayer or photonic crystals, have been widely utilized.

On the other hand, another class of reflection mechanisms, called Guided-Mode Resonance (GMR), have been gathering interests due to its highly wavelength sensitive characteristics and high reflectivity regardless of their simple configuration. GMR occurs as an interaction between an incident plane wave and a leaky mode of a periodic slab waveguide (Resonant Gratings: RGs). The RG traps the externally incident plane wave into its guiding layer. Enhancement of reflectivity occurs at a specific wavelength as a result of constructive interference of radiated waves escaping out from the guiding layer at each grating period [1, 2]. In spite of its practically-useful properties, almost all the former studies about RGs were based on a structure on a flat substrate. Therefore their application has been force limited to the fields where plane waves were assumed [3 - 5].

We have recently demonstrated that a thin film grating on a curved surface also exhibited GMR for concentric waves. According to our analysis it was confirmed that the reflectivity for cylindrical waves maintained > 99% if the curvature radius of the whole waveguide was, roughly, eight times larger than the grating period [6 - 8]. On the basis of this result, in this study we propose a cylindrical cavity that is surrounded by curved RGs, and clarify its resonance characteristics such as Q-factors through numerical simulation.

II. RESULT OF SIMULATION: CURVED RESONANT GRATING

(1) *Method of calculation*

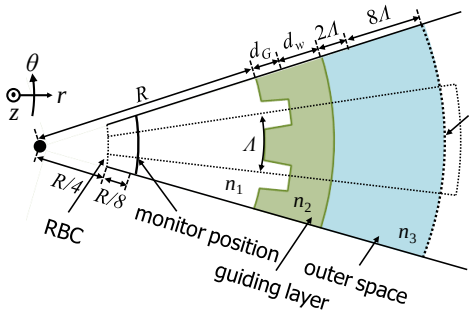


Fig. 1 Schematic view of the curved resonant grating structure. RBC denotes radiation boundary condition.

Before analyzing the cavity we calculated the mirror characteristics of a curved resonant grating (CRG) first. The structure is shown in Fig. 1. The parameters indicated are based on [3]. Radius of the curvature (R) was defined as the distance between the center of the curvature and the surface of the grating. Pitch of the grating (Λ) was defined as the length of an arc at the inner surface. Duty ratio, full thickness of the guiding layer ($d_g + d_w$) and the depth of the grating (d_g) was set as 1:1, 0.72Λ, and 0.33Λ, respectively. Refractive indexes of the cavity, guiding layer, and outer space were 1.0, 2.15, and 1.45, respectively. The structure is assumed to be uniform along the z axis.

We calculated the characteristics of this CRG by two-dimensional cylindrical coordinate version of the Finite-Difference Time-Domain (FDTD) method. The computational domain was denoted by dotted lines in Fig. 1. The angular direction (θ) was truncated by one angular period. Both ends

of the θ direction were connected by periodic boundary walls. The inner and outer bounds of the radial direction were terminated by radiation boundary condition (RBC) [9]. A pulse waveform with a finite duration having a center wavelength of λ = 1.25Λ was launched from the inner space. The reflected wave was monitored near the position of excitation. Throughout this paper, we focused on TM polarization (E_z, H_θ, H_r) only.

(2) Results of simulation

The results of calculation of resonance wavelength and peak reflectivity for RGs with various curvature radii are shown in Fig. 2. Horizontal axis corresponds to the ratio of pitch to radius (curvature). The left-edge of the figure correspond to flat RG (radius = infinity). The figure indicates that the resonance wavelength of CRG increases with curvature. This can be explained by that the modal field of a bent waveguide usually shifts toward outside of curvature, which stretches the effective pitch length Λ. Reflectivity almost maintains 100% if the curvature Λ/R is smaller than 0.157. In this region, 50% bandwidth of the resonance peak (FWHM) was found to be the order of 0.05Λ. Overall reflectivity, bandwidth and resonance wavelength were all kept as the same level as those of the flat RG structure. This indicates that this resonance peak comes from GMR. The reflectivity begins to decrease at about $\Lambda/R \cong 0.157$, and drops to 50% at Λ/R = 0.278. The resonance peak disappears finally at Λ/R = 0.392. The major origin for this degradation will be a large radiation loss due to the waveguide bending.

III. RESULT OF SIMULATION: CAVITY

(1) Method of calculation

Next, we investigated resonance characteristics of cylindrical cavities surrounded by CRG.

Structure for the analysis is shown in Fig. 3. Reflective indexes and thicknesses are the same as shown in Fig. 1. Inner radius of the cavity was set so that the CRG represent almost 100% reflectivity. We removed the inner RBC and extended the FDTD mesh to the center of the circular domain. A pulse waveform with a finite duration having a center wavelength of λ = 1.25Λ was launched from the center of inner space. The transmission power was monitored at 2Λ away from the outer surface of the waveguide layer.

Calculated resonance wavelengths of the cavity are shown in Fig. 4. GMR line of the CRG is shown by an almost flat dotted line. On this line, the CRG almost functions as a perfect mirror, while off the line the reflectivity drops immediately. Almost all the resonance wavelengths except near the GMR line are found to coincide with the lines drawn from upper left to bottom right. In this region, cavity wall almost behaves as a high index dielectric mirror with no periodic corrugation. This wall produces Fresnel reflection, thus in this paper, we call the resonance modes which are far away from GMR line "Fresnel modes". Fresnel modes are determined by the following equation:

$$J_0(nk_0R) = 0 \qquad (1)$$

where J_0, n, k_0 are the 0th order of Bessel function, reflective index, wave vector, respectively. On the other hand, in Fig. 4, the resonance wavelengths near the GMR line (outlined dots) are found to be anti-crossing. Hereafter we call the resonance modes around this region "GMR modes". At the resonant grating, the apparent reflection plane is drastically shifted by a half wavelength near the resonance wavelength. The anti-crossing in Fig. 4 can be explained by this characteristic. The dashed vertical lines in Fig. 4 show the structure where the cavity wavelength and the resonance wavelength by GMR become very close. These structures are expected to exhibit very high Q-factor where CRG represent almost 100% reflectivity.

To verify this assumption, we calculated Q-factors of the cavity using equation (2) at the cavity wavelengths near GMR line in Fig. 4.

We evaluated the Q-factor using the following relation:

$$U(t_0 + \Delta t) = U(t_0)\exp\left(-\frac{2\pi f}{Q}\Delta t\right) \qquad (2),$$

where f, $U(t_0)$ and $U(t_0+\Delta t)$ are the resonance frequency, stored energy at t_0 and $t_0 + \Delta t$, respectively.

Q-factor can thus be directly calculated by the decay rate of the stored energy at the resonant frequency. Because the stored energy is proportional to the power escaping out of the cavity,

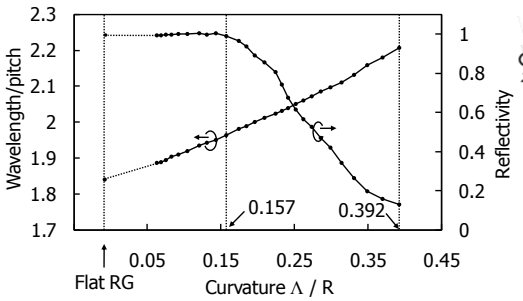


Fig. 2 Curvature vs GMR wavelength and reflectivity.

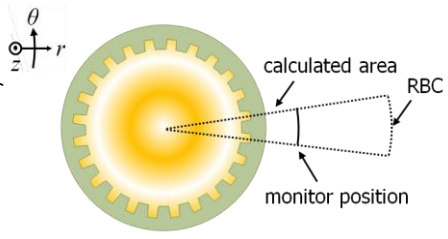


Fig. 3 Schematic view of the cylindrical cavity surrounded by CRG.

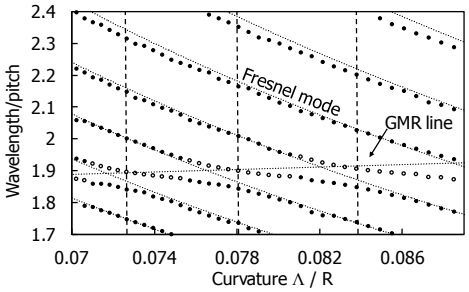


Fig.4 Curvature vs cavity wavelength.

we evaluated it using the intensities of the pointing power flowing across a monitoring point outside the cavity.

The result is shown in Fig. 5. The horizontal line and the dashed lines in the figure correspond to those in Fig. 4. This result clearly indicates that Q-factor is strongly dependent on the difference between the resonance wavelengths of CRG and the cavity. $\Lambda/R = 0.0781$, the maximum Q-factor was found to reach more than 1.0×10^8.

Next, we calculated Q-factors of various resonance modes for a specific structure ($\Lambda/R = 0.073$, left dashed line in Fig. 4 and Fig. 5).

The result is shown in Fig. 6. This figure indicates that Q-factor also reaches very high only if the cavity wavelength and the resonance wavelength coincide. These results imply that by appropriately designing the periodic structure of CRG, we can create nearly single-moded, wide-area circular cavity which has only one high-Q resonance mode in a wide range of wavelength.

Next, we calculated electric field distributions of GMR modes and Fresnel modes of the structure of $\Lambda/R = 0.073$.

The results for a Fresnel mode at $\lambda = 1.992\Lambda$ and GMR mode at $\lambda = 1.893\Lambda$ are shown in Fig. 7 by solid and dotted lines, respectively.

In Fig. 7, the nodes of electric field of Fresnel mode were not clear because of the low reflectivity of CRG. Also the field of Fresnel mode was found to leak to outside of the cavity. On the other hand, the nodes of electric field of GMR mode are clearly found and the leakage is small. The field of GMR mode is also strongly confined in the guided layer. This indicates that the CRG itself is also under resonance.

IV. Conclution

We analyzed the characteristics of the curved resonant grating, the resonance wavelengths and the reflectivity with various curvature radii. According to the results, we verified that the structure with 100% reflectivity exists if curvature is set larger than the adequate parameter. Next, we investigated the characteristics of cylindrical cavities surrounded by curved resonant gratings, the cavity wavelengths and Q-factor. According to these results, we demonstrated the possibility of single-moded, wide-area, high-Q circular cavity resonator thorough numerical simulations.

V. References

[1] D. Rosenblatt et al., “Resonant grating waveguide structures” IEEE J. Quantum Electron. **33**(11), 2038-2059 (1997)

[2] P. Vincent et al., ”Corrugated Dielectric Waveguides: A Numerical Study of the Second-Order Stop Bands,” Appl. Phys. **20**, 345-351 (1979).

[3] D. W. Peters et al., ”Effect of finite grating, waveguide width, and end-facet geometry on resonant subwavelength grating reflectivity,” J. Opt. Soc. Am. **21**(6), 981-987 (2004).

[4] R. Magnusson et al., ”Physical basis for wideband resonant reflectors,” Opt. Express **16**(5), 3456-3462 (2008)..

[5] Y. Ding et al., ”Resonant leaky-mode spectral-band engineering and device applications,” Opt. Express **12**(23), 5661-5674 (2004).

[6] S. Iijima et al., IEICE Gen. Conf. ‘11, C-3-79.

[7] S. Iijima et al., IEICE Ele. Conf. ‘11, C-3-64.

[8] Y. Ohtera et al., “Guided-mode resonance in curved geometry” Opt. Lett. **36**(9), 1689-1691 (2011).

[9] Y. Xu et al., “Finite-difference time-domain calculation of spontaneous emission lifetime in a microcavity” J. Opt. Soc. Am. B **16**(3), 465-474 (1999).

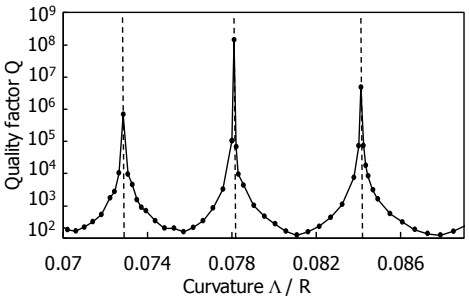


Fig. 5 Curvature vs Quality factor.

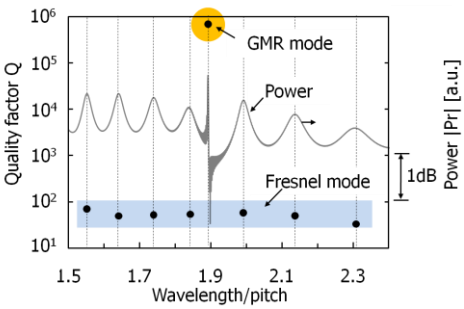


Fig. 6 Wavelength vs Quality factor and Power

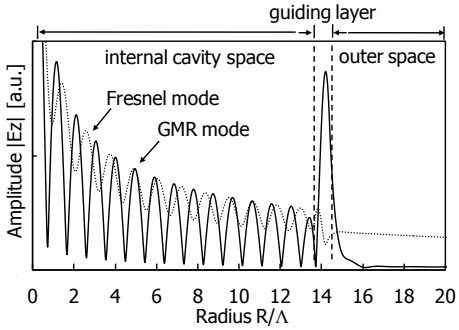


Fig . 7 Electric field of Fresnel mode and GMR mode

Design of 32W Power Amplifier Based on packaged SiC MESFET

Yijie Qiu, Yunchuan Guo, Yuehang Xu, Yongbo Chen, Ruimin Xu,Weigan Lin
EHF Key Laboratory of Fundamental Science
University of Electronic Science and Technology of China
Chengdu 611731, China

Abstract—This paper reports on the design methodology and realization of a power amplifier using a packaged Silicon Carbide metal–semiconductor field-effect transistor (SiC MESFET）. The power amplifier was measured to have 32.4W（45.1dBm） output power at 1000MHz. The large signal model for packaged SiC MESFET is established with the consideration of parasitic effects introduced by package. The simulated results by implementing large signal model into commercial microwave simulator shows quite consistently with measurement results.

Keywords-SiC; FET; Power amplifier; Packaged transistor

I. INTRODUCTION

Wireless communications, medical microwave imaging, radars and other traditional military applications require superior performance of microwave transistors in power amplifiers to improve the performance of the system. The essential challenge to the designer of the power amplifier is applying the right circuit design topology and choosing a suitable semiconductor technology to intensify power capability over a broad bandwidth.

With fast growing of the wide band-gap semiconductors - especially SiC and Gallium Nitride (GaN) - SiC and GaN transistors have achieved popularity due to their distinct high breakdown voltage, high electron velocity and high thermal conductivity [1].

The much higher output power density of wide band-gap transistors allows manufacturing much smaller size devices with the similar output power. The wide band-gap also enables the transistor to operate at higher temperatures. These attractive merits in power amplifier applications make SiC devices be widely applied in microwave, radio frequency (RF), radar and wireless communication power amplifiers. Numerous kinds of power amplifiers which utilized SiC MESFET have been reported [2]-[6].

SiC power amplifier MMIC technology is rapidly growing [7]. However it is expensive because of a large chip size particularly in the L and S frequency band. In addition passive components such as capacitor in MMIC should have a high breakdown voltage. According to these factors hybrid SiC RF power amplifier is still an important option which needs careful design considerations.

This paper demonstrates the use of design procedures to manufacture a power amplifier from 950 to 1000 MHz using a SiC MESFET. 32.4W output power is obtained with 5.1 dB power gain at frequency 1000MHz.

Section I is a brief introduction of SiC MESFET and its application in power amplifiers. In Section II, the power amplifier design method is presented. In section III the simulated and measured results of the implemented SiC MESFET power amplifier are discussed. Conclusions are presented in Section IV.

II. DESIGN

A. *Modeling of the SiC MESFET*

Empirical nonlinear models have been extensively used in the modeling of field effect transistors (FET). Numerous kinds of methods have been used to improve the accuracy of the empirical nonlinear models of the SiC MESFETS [8].

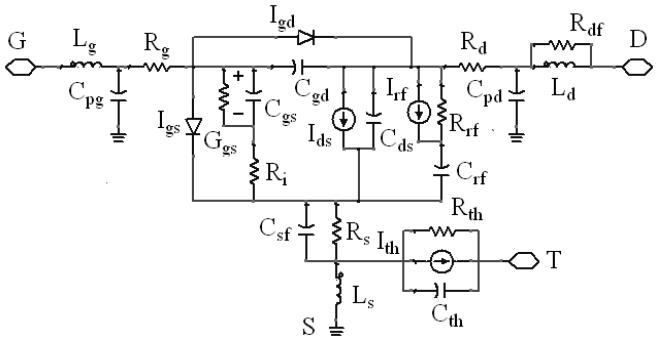


Figure 1. Large-Signal Equivalent Circuit for SiC MESFETs

The SiC MESFET used in the power amplifier consists of 30-mm gate periphery that composed of thirty 1-mm unit cells whose gate periphery has two 500-um gate fingers. The critical point in modeling the SiC MESFET is to have an accurate large signal equivalent circuit of the unit cell. Then the model of the whole SiC MESFET can be constructed by paralleling the unit cells.

The empirical nonlinear model of the SiC MESFET used to simulate the performance of the power amplifier is shown in Fig. 1 [9]. Angelov model was used to model the non-linear gate capacitance of SiC MESFET. The gate-source capacitance can be written as:

$$C_{gs} = C_{gs0}\left[1-\tanh(\Phi_1\right]\left[1-\tanh(\Phi_2\right] \qquad (1)$$

This work is supported by National Natural Science Foundation of China (Grant No.61106115, 60971037, and 60876052) and the Central University Basic Research Funding (103.1.2E022050205)

Where

$$\Phi_1 = A + BV_{ds} + CV_{ds}^{\ 2} \qquad \Phi_2 = A_1 + B_1 V_{gs} \tag{2}$$

The gate-drain capacitance is described as:

$$C_{gd} = C_{gd0} + C_{gd1}[1 + \exp(\Phi_1)]\exp(\Phi_2) \tag{3}$$

Where

$$\Phi_1 = AV_{ds} \quad \Phi_2 = A_1 V_{gs} \tag{4}$$

The above nonlinear model has been integrated into Agilent Advanced design system (ADS) as a user-defined model using symbolic defined devices (SDD). SDD of the SiC MESFET integrated in ADS is shown in the Fig.2 below.

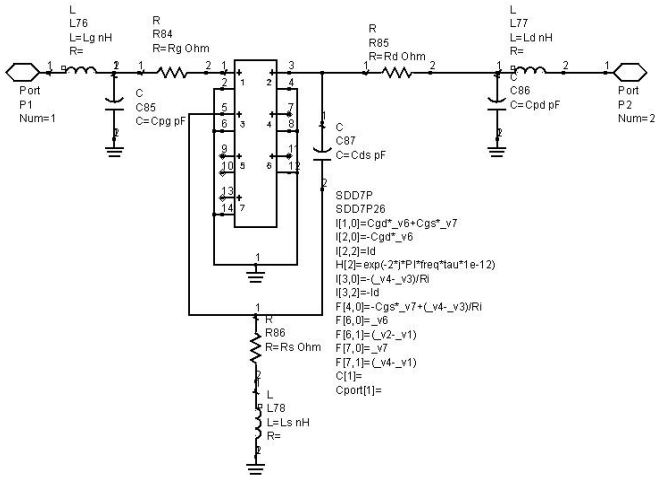


Figure 2. The SDD model of the Single Cell of SiC MESFET

B. *Simulation of the Package*

The Packaged SiC MESFETs are increasingly applied in communication systems in L and S band. Diversify approaches were proposed to extract package model [10].

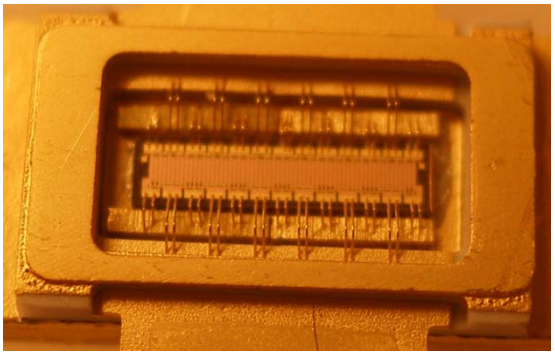


Figure 3. Packaged SiC MESFET

The packaged SiC MESFET used in the power amplifier is shown in Fig.3 above. Due to the complexity of precisely modeling parasitic effects introduced by the three dimensional package and mutual coupling of the gold wires and various elements within the package, development of modeling techniques to simulate the package is complicated since the package can contain tens of bonding wires.

Consequently commercial three dimensional full-wave Electromagnetic field simulation software such as HFSS was often used to simulate the three dimensional package.

Considering the size and complexity of a physically large packaged transistor and the complication and the computation time consumed in HFSS to obtain the S parameters, we segment the package into two symmetrical parts one of which is shown in the Fig.4 below.

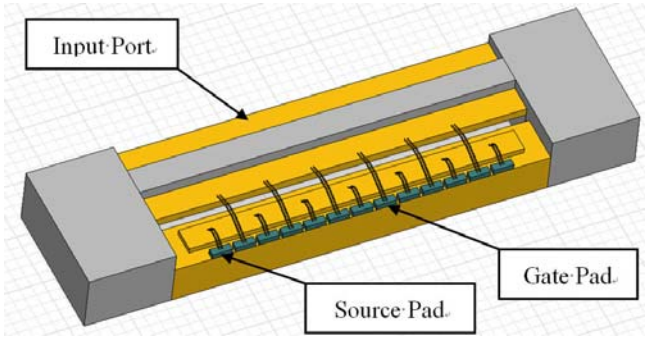


Figure 4. The model of the Packaged SiC MESFET in HFSS

The gold wires that were bonded to the gate and drain pads of the SiC MESFET in the package are often within close proximity, as illustrated in Fig. 4. Through simulation S parameter results can be obtained. The half package can be looked upon as a black box which has fourteen ports. The fourteen ports are described as: one port is the input of the half package, six ports connected to the gate pads of the SiC MESFET and seven ports connected to the source pads of the SiC MESFET.

C. *Design of the Power Amplifier*

The design steps of the power amplifier are described below: First the SDD model of the SiC MESFET is obtained and integrated into ADS. Then the package is analyzed by HFSS and we can get the S parameters of the package. Finally we use ADS to optimize the input and output matching circuits of the power amplifier. Fig.5 below is the detailed design strategy of the SiC MESFET power amplifier.

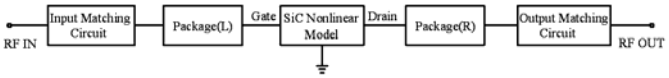


Figure 5. Designing strategy of the SiC MESFET power amplifier

The simplified circuit schematic of the SiC MESFET power amplifier is shown in the Fig.6 below. The matching networks consist of microstrips and capacitors. The resistor RS1 and RS2 are used to improve the stability of the circuit.

The substrate used is RO4003C made by Rogers Corporation. The thickness of the substrate is 0.813mm.

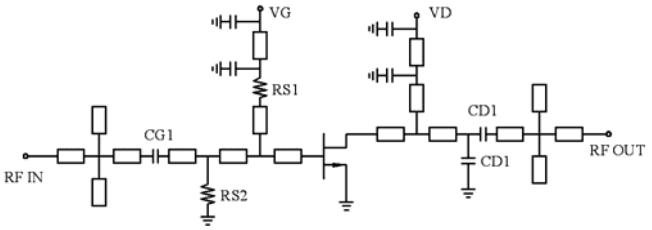


Figure 6. Circuit schematic of the SiC MESFET power amplifier

III. EXPERIMENTAL RESULTS

The measurement setup of the SiC MESFET power amplifier (DUT) is shown in Fig.7. The input power signal was supplied by a driver power amplifier impelled by Agilent microwave signal generator E8257D. The output of the SiC MESFET power amplifier was attenuated by a 40dB attenuator and then measured through Agilent power meter N1912A. Fig.8 shows the simulated power gain and measured power gain of the power amplifier from 950MHz to 1000MHz. The measure power gain is between 5.0-5.5dB at the fixed input power of 40dBm across the band. The measured result fits the simulated result well, but with less measured gain in the low frequency band. The SiC MESFET was biased at V_{GS}=-8V and V_{DS}=+48V.

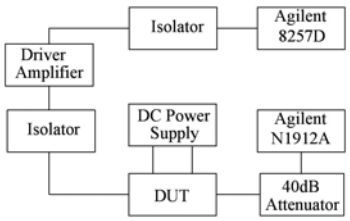


Figure 7. Measurement setup for power amplifier

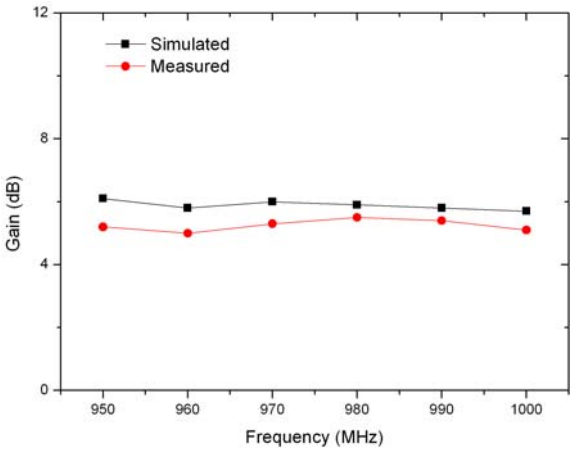


Figure 8. Measured and simulated power gain of the SiC MESFET power amplifier versus frequency at a fixed input power of 40dBm

The measured and simulated output power characteristics of the power amplifier at 1000 MHz is shown in the Fig.9. A maximum output power of 32.4W（45.1dBm）was obtained. Fig.10 is a photograph of the fabricated SiC MESFET power amplifier.

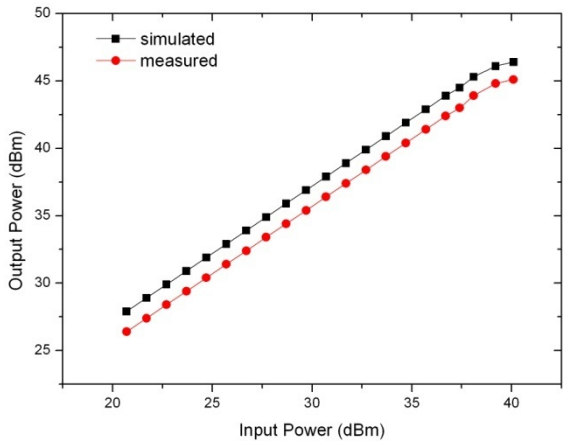


Figure 9. Measured and simulated output power of the SiC MESFET power amplifier at 1000 MHz

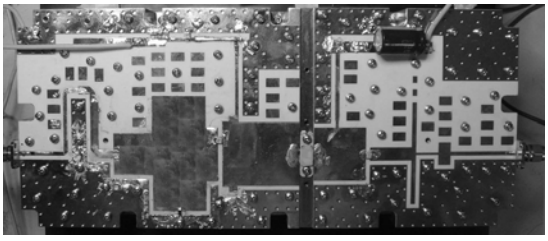


Figure 10. Photograph of the fabricated SiC MESFET power amplifier

IV. CONCLUSION

In this paper, an extensive design procedure for packaged SiC MESFET power amplifier has been presented. The proposed procedure has been demonstrated by implementing a power amplifier based on a packaged SiC MESFET. Experiment results verify the success of the submitted method. The measurement result show that 32.4W output power at 1000MHz, about 5dB power gain from 950MHz to 1000MHz. Good agreement has been obtained between simulations and measurements.

REFERENCES

[1] R.C.Clarke, John W.Palmour, “SiC Microwave Power Technologies,” Proceedings of the IEEE, VOL. 90, NO. 6, pp. 987-992, JUNE 2002.

[2] Chen Wenhua, Li Xiang, Wang Li, Feng Zhenghe, Xue Xing, “A Novel Broadband VHF SiC MESFET Class-E High Power Amplifier,” Microwave and Optical Technology Letters, VOL.52, No. 2, pp. 272-276, February 2010.

[3] H. George Henry, Godfrey Augustine, Gregory C. DeSalvo, Ronald C. Brooks, Robert R. Barron, James D. Oliver, Jr., et al, “S-Band Operation of SiC Power MESFET With 20 W (4.4 W/mm) Output Power and 60% power amplifierE,” IEEE Transaction On Electron Devices, VOL. 51, NO. 6, pp. 839–447, June 2004.

[4] Yong-Sub Lee, Yoon-Ha Jeong, “A High Efficiency Class E Power Amplifier Using SiC MESFET,” Microwave and Optical Technology Letters, Vol. 49, No. 6, pp. 1447–1449, June 2007.

[5] S. Azam, R. Jonsson and Q. Wahab, “Designing, Fabrication and Characterization of Power Amplifiers Based on 10-Watt SiC MESFET & GaN HEMT at Microwave Frequencies,” Pcoceedings of the 38th European Microwave Conference, pp. 444–447, October 2008.

[6] Ahmed Sayed, Stefan von der Mark and Georg Boeck, “An ultra wideband 5W power amplifier using SiC MESFETs,” Microwave Conference, 34th European , Volume: 1, pp. 57-60, 2004.

[7] Mattias Südow, Kristoffer Andersson, et al, “An SiC MESFET-Based MMIC Process,” IEEE Transactions On Microwave Theory and Techniques, VOL. 54, NO. 12, pp.4072-4078, DECEMBER 2006.

[8] I. Angelov, K. Andersson, D. Schreurs, D. Xiao, N. Rorsman, V. Desmaris, et al, ”Large-Signal Modelling and Comparison of AlGaN/GaN Large-Signal Modelling and Comparison of AlGaN/GaN,” Proceedings of Asia-Pacific Microwave Conference, pp.279-282, 2006

[9] Xu, Y., Y. Guo, R. Xu, B. Yan, and Y. Wu, "An support vector regression based nonlinear modeling method for SiC MESFET," Progress In Electromagnetics Research Letter, Vol. 2, pp. 103-114, 2008.

[10] Peter H. Aaen, Jaime A. Plá, and Constantine A. Balanis, “On the Development of CAD Techniques Suitable for the Design of High-Power RF Transistors,” IEEE Transactions On Microwave Theory and Techniques, VOL. 53, NO. 10, pp. 3067-3074, OCTOBER 2005.

The Analysis of Maximum-Distance SW Update for WSNs

HyeYeong Jeong, SeungHui Cha, YeungMoon Kwon, ByoungChul Ahn
Dept. of Computer Engineering
Yeungnam University
Gyungsan, Korea
b.ahn@yu.ac.kr

***Abstract*— Wireless Sensor Networks (WSNs) are applied to many monitoring areas and are deployed for long periods of time using batteries. Present sensor nodes can perform many functions at the same time and are programmed by complex software. During the lifetime of sensor nodes, they are required to reprogram their software because of their new functions, software upgrade, and software bug fixes. The nodes are inaccessible physically or it is very difficult to upgrade their software by one by one. To upgrade the software of sensor nodes in WSNs remotely, this paper analyzes the efficiency and performance of software updates and protocols using the maximum distance method. Their performances are measured by upgrade times, the number of relay nodes, energy consumptions and error rates according to packet sizes. By the simulation results, 160 bytes or 192 bytes of packet size are fast update time and low error rates.**

Keywords- software upgrade; data relay; node upgrade; node relay method

I. Introduction

Wireless Sensor Networks (WSNs) have been applied to many areas and are used for a long period of time until their batteries were exhausted. A few years ago, their performance and functions were very limited and were discarded without reusing or reprogramming them. But present sensor nodes for WSNs have been developed to reduce power consumption with the recent rapid advance of semiconductor technology.

During the lifetime of nodes, sensor nodes needs to update software, fix software bugs, or reprogram new functions without discarding or removing them. But in many cases the nodes are not inaccessible physically and they are not modified by hand. In this case, a lot of costs and efforts are wasted if they are recollected or replaced them with new nodes. Therefore, it is required to upgrade their software remotely.

In this paper, the performance of Maximum-Distance method is analyzed to find the size of packets to update efficiently.

II. Related Works

Many researches for WSNs are studied for low power consumption or data collection methods from nodes to their sink node efficiently. For efficient data transmissions, Intanagonwiwat *et al*. suggest direct-diffusion method [1]. But this research is focused on data aggregation and data transmission path. So it is not suitable for software updates since the data flow direction is reversed. Stephen *et al*. suggest a model for upgrading software in WSNs [8]. This model presents the theoretical approaches to upgrade software. They have not presented any simulation results or experiments to verify their model. They validate their model by comparing three different systems, representing three classes of software update: static/monolithic updating (MOAP), dynamic/mobile agent-based updating (Mate) and dynamic/component-based updating (Impala)

Since control software contains execution code for a processor of sensor nodes, it is very important to maintain reliable data transfer. A method for reliable data transfer in WSNs is developed for 1:1 communication such as S-TCP [3] and RMTS [4]. But 1:1 communication methods are inefficient to upgrade many nodes of WSNs. If these methods are used for re-programming sensor nodes of WSNs, each node must be updated first and retransmit the software upgrade data to another node one by one. Therefore it is necessary to develop an efficient upgrade method for sensor nodes with fast upgrade time and small data retransmission.

The direction to upgrade control software is the opposite direction of normal data transfer [5][6]. It is necessary to study for large data transfer from one node to many nodes efficiently. There are some researches about upgrades for sensor nodes. But they are focused on system management, not an upgrade it [7].

III. Upgrade Model

All sensor nodes of WSNs are assumed to be the same configured such as the memory size, the same processor and *etc*. It means that all sensor nodes use the same software version. And a distance between two nodes is the same and the location of nodes is fixed.

A. Data Relay

Generally, the normal mode of a node is operating in sensing data mode with low power consumption. Each node transmits its sensed data to the sink node. When a sink node starts to transfer its software update data to others, all nodes stop their sensing operation and switch their mode to the software upgrade mode. In case of software data transmission, data size is very large and must be transmitted very fast and continuously. If a node detects the software upgrade protocol, it should switch the normal sensing mode to the software upgrade

mode and prepare for the software upgrade. When the node finishes receiving all software upgrade data, it requests lost packets to its source node. After lost data are received again, the node reprograms its own flash memory and restarts its operation again. After upgrading its software, the node must relay the software upgrade data to other nodes. But it is not necessary for all nodes to participate in relaying the software upgrade process to another node in WSNs. Only a few nodes need to relay the software upgrade data to other nodes. Therefore, it is very important to choose relay nodes because of overall performance.

Figure 1 (a) shows a data relay model when nodes are placed in-line. Black nodes in Figure 1 (a) represent nodes that participate in relaying the software upgrade data. And the most left black node is the start node. In Figure 1, "*r*" means radio radius. All relay nodes are placed in multiples of radio radius of the location. The other nodes only receive the software upgrade data and reprogram themselves. So the number of relay nodes, *N*, is calculated by (1).

$$N = \left\lceil \frac{l}{r} \right\rceil \tag{1}$$

r is a radio radius of a node

l is a distance between the start node and the last node

If nodes are placed in line, the number of relay nodes is proportional to the distance of the start node and the last node. But in real WSNs, each node is located on 2-dimensional plain rather than in line. Figure 1(b) and Figure 1(c) show the software upgrade data relay nodes that are located on 2-dimensional plain at each step. If a node is located on the boundary of radio radius, it is the ideal the software upgrade. So the total number of relay nodes, *N*, is calculated by (2).

$$N = \begin{cases} 1, & (l <= r) \\ 1 + \sum_{n=1}^{\lfloor d \rfloor} \left\lceil \frac{2\pi \cdot n \cdot r}{r} \cdot \frac{1}{4} \right\rceil = 1 + \sum_{n=1}^{\lfloor d \rfloor} \left\lceil \frac{\pi \cdot n}{2} \right\rceil & (l > r) \end{cases} \tag{2}$$

r is a radio radius of a node

l is distance between start node and last node

$d = l/r$

The minimum number of relay nodes is calculated by (2). To upgrade all nodes in the field at least N nodes should participate in relay operation. When the start node is located in the corner of WSNs, Equation (2) is acceptable. If the start node is located in the center of WSNs, the number of relay nodes, *N*, is increased up to four times

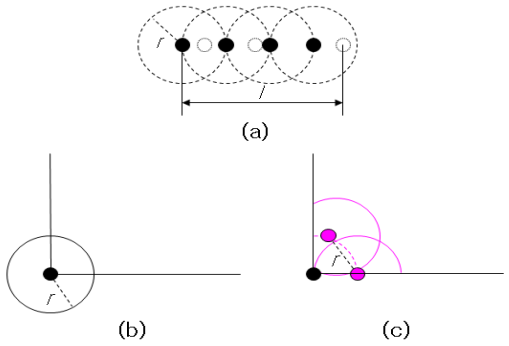


Figure 1. Data relay model

B. Power Consumption and Upgrade times

The power consumption of relay nodes is calculated by the sum of receiving data and transmitting data, and is calculated as below.

$$J_{nr} = J_r \cdot file\_size(1+e)$$
$$J_{ns} = J_s \cdot file\_size(1+e) + J_{nr} \tag{3}$$

J_r is the energe for datas end

J_S is the energe for datas send

J_{nr} is the energy consumed by receiving node

J_{ns} is the energy consumed by relaying node

e is trasnmission error rate

$file\_size$ is the size of firmware data

Equation (3) represents that J_{nr} and J_{ns} are the power consumption of receiving data and transmitting data to other nodes.

Some nodes located in duplicated radio area are received a few multiple times of data size of the software upgrade data. Therefore, the total energy consumption of all nodes is *J* in (4).

$$J = (J_s + J_r \frac{N_t}{Area\_of\_field} \cdot \pi \cdot r^2 + e \cdot J_s) file\_size \cdot N \tag{4}$$

N_t is total number of nodes in a filed

N is a number of relaying nodes

Equation (4) represents that all parameters are fixed except *e* and *N*. It is very important to reduce transmission error and the number of relaying nodes.

The time to upgrade all nodes of WSNs depends upon the hop count of relay. The time is calculated by Equation (5).

$$T = \lceil d \rceil \cdot (t_s + e \cdot t_s + t_u) \tag{5}$$

d is relay step count (l/r)

t_S is firmware filesned time (file_size/ bandwidth)

t_u is update time

e is transmission error rate

From (5), the total software update time of WSNs depends upon relay step count, *d*, and transmission error rate, *e*. If all relay nodes send software upgrade data to other nodes at the same time, they make a lot of collision or interference by radio signals and the transmission error rate is increased. It is necessary to prevent all nodes from relaying data at the same time in each step. To reduce energy consumption and upgrade time, it is very important to select a relay node at each step.

IV. Relay Methods and Simulation

In this section, three relay methods to transmit the software upgrade data are described. Simulation results of their performances are discussed.

A. Relay Methods

It is very important to reduce the number of relay nodes and prevent neighbor nodes from retransmitting data at the same time. As soon as a relay node transmits the software upgrade

data, it should select the next relay node to propagate the software update operation.

If the next relay node is located on the boundary line of radio radius of the present relay node, it is the most effective relay node. But in real WSNs, there are a few nodes on the boundary line of radio radius. It is hard to recognize if the nodes are located on boundary line of radio radius. Selection methods of relay nodes are carefully considered.

The software upgrade protocol has two steps. The first step is pre-transmission step. A relay node should know the status of all neighbor nodes before it starts sending the software upgrade data. All neighbor nodes broadcast their status to the relay node periodically in the first step. The status data includes the upgrade version information, the number of neighbor nodes and the node status.

The next step is to receive the software upgrade data and update it. Before the relay node starts to send the software upgrade data, neighbor nodes respond by sending a "*receive-start*" message. And neighbor nodes start to receive the software upgrade data and store it to their memory. After the upgrade process, neighbor nodes must request the relay node to retransmit lost packets. After they finish receiving the lost packets, they reprogram by themselves and send back "*reprogram-done*" message to the relay node. If a time-out occurs or it receives messages from all participated neighbor nodes, the relay node should select the next relay node from neighbor nodes and sends "*relay-start*" message.

To select a next relay node, Maximum-Distance method is chosen. This method is that the relay node is the longest distance from relay node.

B. Simulation Environments

NS-2 network simulator is used for the simulation. 100 nodes are deployed in the field uniformly. The distance between two nodes is 40*m* and each node has a 60*m* radio radius. Radio bandwidth is 256Kbps and software file size is 128 Kbytes. The energy consumption is 75.9*mW* to send data and 62.7*mW* to receive data. The time to upgrade a node is 1.5*sec*. The start node is located in the left bottom corner. Three selection methods are simulated to compare their performances.

C. Performance Metrics

In order to evaluate the proposed methods, following metrics are used.

The number of relay nodes - Depending on the selection method of relay nodes, the number of relay nodes might be varied. It is the most important to measure the upgrade performance.

Energy consumption - Since sensor nodes in WSN are operated by battery power, the energy consumption is important factor with total upgrade time.

Total upgrade time to all nodes - This factor shows the effectiveness of the software upgrade in the WSNs. Depending upon relaying nodes, the total upgrade time is much different.

Data loss rate – This factor shows the performance of each selection method. Typically, collision and interference leads to data packet loss in WSNs.

D. Simulation Results

The number of relay nodes to upgrade software is calculated by (2). Sensor nodes in WSNs are deployed in rectangular. It is very difficult to calculate exact relay nodes for the Maximum-Distance method. By Equation (2), the number of relay nodes is about 37. Figure 2 shows the number of relay nodes. In all case, the number of relay nodes is between 33 ~ 39 nodes. When the packet size is 96bytes, the total number of relay nodes is 33 nodes. When the packet size are 128bytes and 160bytes, the total number of relay nodes are 36 nodes and 37nodes respectively. The difference of relay node is six among variable packet sizes.

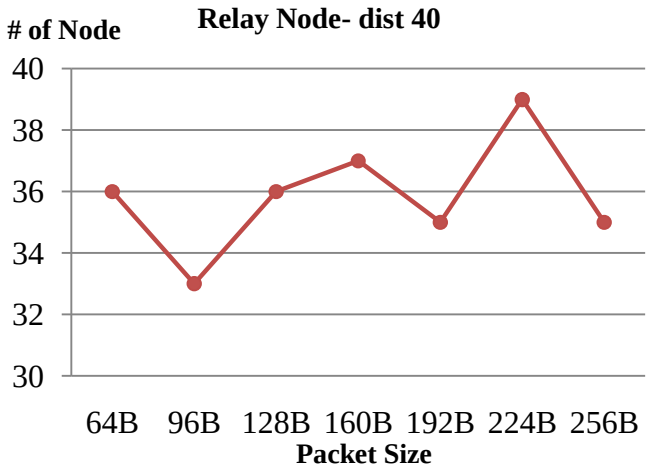


Figure 2. The Number of Relay Nodes

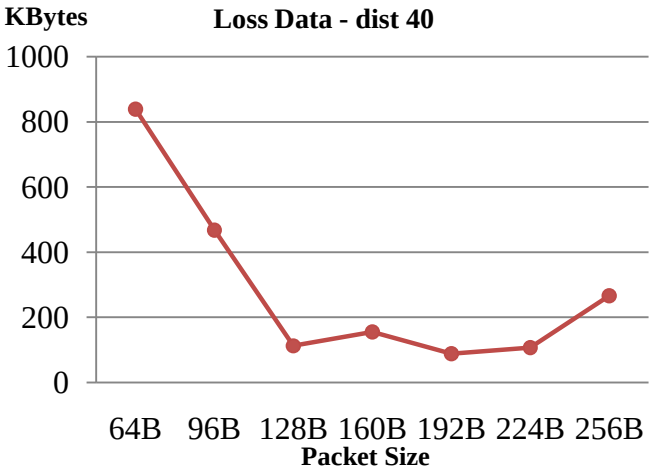


Figure 3. Data Loss

Figure 3 shows the size of lost data. When the packet sizes are 64bytes or 96 bytes, the data loss are 840Kbyes and 460Kbyets updating 128Kbyets of software. When the packet size is small, nodes transmit many data packets and there are a lot of collision occurred by sending packets. And the collision induces energy consumption and total upgrade time. When

packet sizes are between 160 bytes and 256 bytes, data loss is less than 150KB.

The energy consumed by each node is calculated with (3). If there is no lost packet, error rate, *e*, is zero. Ignoring control and status packets, the energy consumed by each data receiving node, J_{nr}, is 257*mJ*. And the energy consumed by each relay node, J_{ns}, is the sum of energy consumed by receiving data and transmitting data. The energy consumed by relay nodes, J_{ns}, is calculated as 568*mJ*. The total energy consumption of WSNs is calculated by (4) and is about 103.9*J*. Figure 4 shows total energy consumption to upgrade all nodes. When the packet size is 256 bytes, the energy consumption is low. When the packet sizes are between 96 bytes and 224 bytes, the energy consumptions are between 115 *J* and 105 *J*.

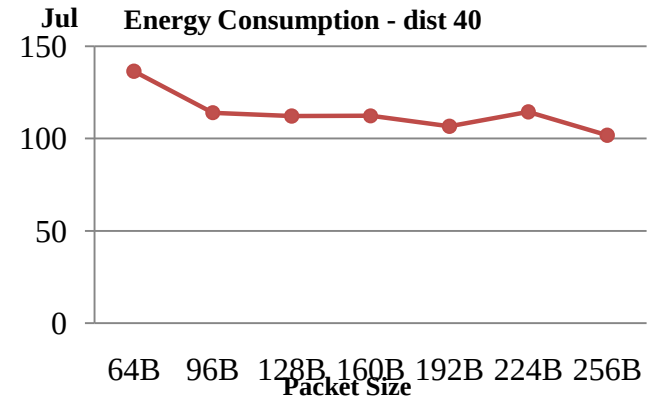


Figure 4. Energy Consumption

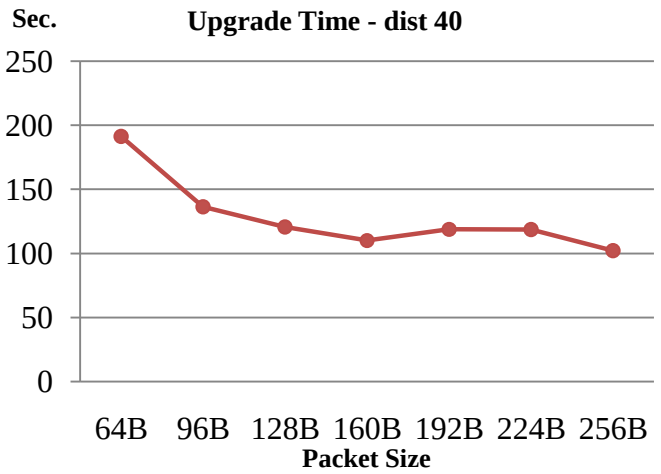


Figure 5. Upgrade Time

Total time to upgrade software is estimated as 51.2 sec by (5). This value is the minimum value. But in the simulations, some additional time are considered such as error recovery time, message wait time and so on. Figure 5 shows the time to upgrade software to all nodes. The results follow the similar pattern to Figure 4. When the packet size is 256 bytes, the upgrade time is 102*sec*. When the packet sizes are between 96 bytes and 224 bytes, the upgrade times are between 135*sec* and 105 *sec*.

V. Conclusion

In this paper, the packet sizes are simulated to find the proper packet size for upgrading software using the Maximum-Distance method. The number of relay nodes, data loss, energy consumption and upgrade time are analyzed by the packet sizes for WSNs. The simulation results show that 128 bytes or 160 bytes are fast update time, low data loss and low energy consumption. But the number of their relay nodes is higher than that of 96 bytes.

References

[1] C. Intanagonwiwat, R. Govindan, and D. Estrin "Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks," Proc. of the 6th annual international conference on Mobile computing and networking(Mobicom '00), pp.56-67, 2000

Wei Ye, J. Heidemann, and D. Estrin, "Sensor-MAC (S-MAC): Medium Access Control for Wireless Sensor Networks," Proc. of the 21st International Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2002), vol.3, pp.1567-1576, 2002.

Y. G. Iyer, S. Gandham, and S. Venkatesan, "STCP: A Generic Transport Layer Protocol for Wireless Sensor Networks," Proc. of 14th International Conference on Computer Communications and Networks, pp.449-454, 2005.

F. Stann, and J. Heidemann, "RMST: Reliable data transport in sensor networks," Proc. of the First IEEE. 2003 IEEE International Workshop on Sensor Network Protocols and Applications, pp. 102 -112, 2003.

W. Chen, P. Chen, W. Lee, and C. Huang, "Design and Implementation of a Real Time Video Surveillance System with Wireless Sensor Networks," Proc. of Vehicular Technology Conference, 2008, pp.218-222. 2008.

Honggang Wang , Dongming Peng , Wei Wang , Hamid Sharif , Hsiao-Hwa Chen , "Image transmissions with security enhancement based on region and path diversity in wireless sensor networks", IEEE Transactions on Wireless Communications, vol. 8, no. 2, pp.757-765, 2009.

C-C. Han, R. Kumar, R. Shea, M. Srivastavam, "Sensor Network Software Update Management: a Survey," Intl. Journal of Network Management, no. 15, No. 4, John Wiley & Sons, pp. 283-294, 2005.

S. Brown and C. Sreenan, "A New Nodel for Updating Softeare in Wireless Sensor Networks," IEEE Network Nov/Dec. pp.42-47, 2006.

Y. Kwon, Ph.D. Dissertation, "A Study on Real Time Software Upgrade Methods on WSNs," 2011, Youngnam University.

Complexity Reduced Near ML Block Signal Detection for Single-carrier MIMO Transmission

Tetsuya YAMAMOTO† and Fumiyuki ADACHI‡
Dept. of Electrical and Communication Engineering, Graduate School of Engineering, Tohoku University
6-6-05 Aza-Aoba, Aramaki, Aoba-ku, Sendai, 980-8579 Japan
†yamamoto@mobile.ecei.tohoku.ac.jp, ‡adachi@ecei.tohoku.ac.jp

Abstract—**For high-speed data transmissions, multi-input multi-output (MIMO) spatial multiplexing is very effective. However, if single-carrier (SC) transmission is used, the transmission performance significantly degrades due to not only an inter-antenna interference (IAI) but also a large inter-symbol interference (ISI) resulting from a severe frequency-selective fading. Various signal detection schemes for SC-MIMO spatial multiplexing have been proposed, among which maximum likelihood (ML) detection is known as an optimum signal detection scheme. However, the computational complexity of ML detection is extremely high. In this paper, we present a complexity reduced near ML block signal detection which jointly use a training sequence (TS) aided SC block transmission and QR decomposition and M-algorithm. We evaluate, by computer simulation, the achievable bit error rate (BER) and throughput performance of turbo-coded hybrid automatic repeat request (HARQ) and show that the proposed near ML block signal detection significantly improves the BER and throughput performances at a much lower computational complexity compared with the conventional near ML detection. We also compare the achievable PER, throughput, and computational complexity of the proposed near ML block signal detection to that of the conventional near ML block detection and linear detection based on the minimum mean square error (MMSE) criterion.**

Keywords-component; Single-carrier, MIMO, near maximum likelihood detection, QR decomposition, M-algorithm, training sequence

I. INTRODUCTION

Recently, there have been tremendous demands for high speed data transmissions of close to or even above 1Gbps in mobile communications [1]. However, since the mobile wireless channel is composed of many propagation paths with different time delays, the channel becomes severely frequency-selective as the transmission data rate increases [2]. Orthogonal frequency-division multiplexing (OFDM) [3] has been adopted in several wireless communication standards because of its robustness against frequency-selective fading. However, OFDM has higher peak-to-average power ratio (PAPR) than the single-carrier (SC) transmission. Recently, SC transmission with frequency-domain equalization has been gaining an increasing popularity because of its lower PAPR property [4, 5].

To achieve very high speed data transmission with a limited bandwidth, multi-input multi-output (MIMO) spatial multiplexing has been attracting considerable attention [6]. The use of high level modulation is also essential. However, the SC transmission using MIMO spatial multiplexing and high level modulation is vulnerable not only to the inter-antenna interference (IAI) but also to the inter-symbol interference (ISI) caused due to the frequency-selective fading channel, and therefore, the use of powerful equalization and signal detection scheme is indispensable.

Various signal detection schemes have been developed for SC-MIMO spatial multiplexing, among which maximum likelihood (ML) detection [7] is known as an optimum signal detection scheme. However, its computational complexity becomes extremely high because the number of symbol candidate sequence is exponentially increased to $X^{N_t N_c}$ for X-QAM, where N_t is the number of transmit antennas and N_c is the block size which equals the number of symbols in a block. Therefore, the use of ML detection is not realistic and a low-complexity signal detection scheme such as a frequency-domain linear detection based on the minimum mean squared error (MMSE) criterion is often used [8]. However, the performance improvement is limited and a big performance gap still exists from the ML performance due to the presence of residual ISI and IAI, particularly when higher level modulation (e.g., 16QAM, 64QAM) is used. Therefore, a lot of signal detection schemes, which provide a close-to-ML performance but with reduced complexity, have been studied.

Complexity reduced near ML detection based on the QR decomposition and M algorithm [9] is one of the promising schemes and the application of the QR decomposition and M algorithm for the SC-MIMO block signal detection was proposed [10] (we call this detection scheme as the QRM-MLBD). QRM-MLBD can significantly improve the bit error rate (BER) performance of SC-MIMO spatial multiplexing when compared to the MMSE detection (MMSED). However, the use of a fairly large number M of surviving paths in the M-algorithm is required, and therefore, the computational complexity of the conventional QRM-MLBD for SC-MIMO is still very high. This is because if smaller M is used, the probability of removing the correct path at early stages increases.

In this paper, to further reduce the computational complexity, we present to jointly use the training sequence (TS) aided SC block transmission and the QRM-MLBD for SC-MIMO spatial multiplexing. TS aided SC (TS-SC) block transmission [11] is used instead of an often used cyclic prefix inserted SC (CP-SC) block transmission and its TS is utilized to reduce the probability of removing the correct path at early stages [12, 13]. The achievable BER performance of SC-MIMO using the proposed TS aided QRM-MLBD is evaluated by computer simulation to show that TS aided QRM-MLBD achieves better BER performance than the conventional QRM-MLBD while significantly reducing the computational complexity.

Packet access will be the core technology of the next generation mobile communication systems. Very high-speed and high-quality packet transmissions in a limited bandwidth can be achieved by the joint use of MIMO multiplexing and hybrid automatic repeat request (HARQ) [14]. However, the throughput performance of SC-MIMO HARQ using near ML detection has not yet been fully investigated. Then, we evaluate, by computer simulation, the throughput performance of turbo-coded SC-MIMO HARQ using the proposed near ML detection and show that the proposed TS aided QRM-MLD provides significantly higher throughput performance than MMSED especially when high level modulation is used.

The remainder of this paper is organized as follows. Sect. II describes CP-SC MIMO and TS-SC MIMO transmission system model. Sect. III presents QRM-MLBD. In Sect. IV, we show the computer simulation results of the achievable BER and throughput performance. Sect. V offers some concluding remarks.

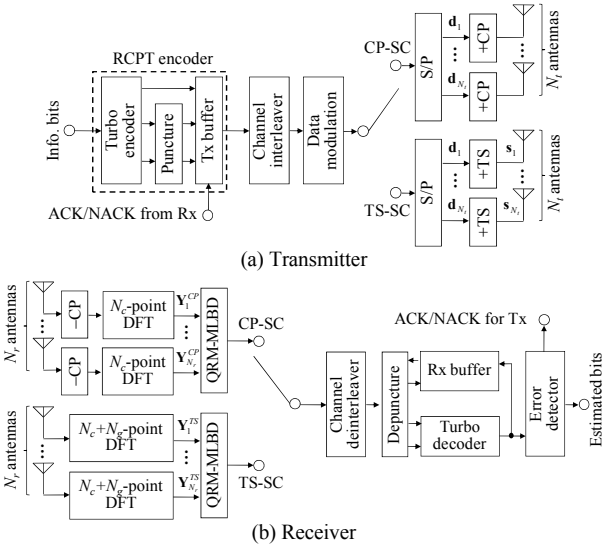


Figure 1. Transmitter and receiver structure.

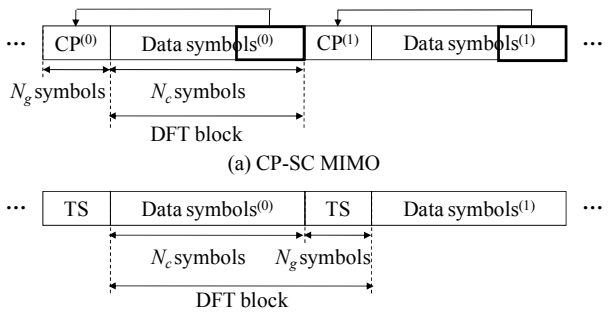


Figure 2. Block structure.

II. CP- AND TS-SC MIMO TRANSMISSION SYETEM MODEL

A. CP-SC MIMO

The CP-SC MIMO block transmission system model and the block structure are illustrated in Fig. 1 and Fig. 2 (a), respectively. Throughout the paper, the symbol-spaced discrete time representation is used. At the transmitter, the coded bit sequence is transformed into a data-modulated symbol sequence. Then, the data-modulated symbol sequence is serial-to-parallel (S/P) converted to N_t parallel symbol sequence, each to be transmitted from a different transmit antenna and each parallel symbol sequence is divided into a sequence of symbol blocks of N_c symbols each. The data symbol block of n_t-th transmit antenna can be expressed using the vector form as $\mathbf{d}_{n_t} = [d_{n_t}(0),...,d_{n_t}(t),...,d_{n_t}(N_c-1)]^T$, where $(.)^T$ expresses the transposition. The last N_g symbols of each block are copied as a CP and inserted into the guard interval (GI) placed at the beginning of each block and a CP-inserted data block of N_c+N_g symbols is transmitted.

The signal block is transmitted over a frequency-selective fading channel. The received signal block after CP removal is transformed by N_c-point discrete Fourier transform (DFT) into the frequency-domain signal. The frequency-domain received signal vector at the n_r-th receive antenna $\mathbf{Y}_{n_r}^{CP} = [Y_{n_r}^{CP}(0),...,Y_{n_r}^{CP}(k),...,Y_{n_r}^{CP}(N_c-1)]^T$ is expressed as

$$\mathbf{Y}_{n_r}^{CP} = \sqrt{\frac{2E_s}{T_s}} \sum_{n_t=1}^{N_t} \mathbf{H}_{n_r,n_t}^{CP} \mathbf{F}^{(N_c)} \mathbf{d}_{n_t} + \mathbf{N}_{n_r}^{CP} , \qquad (1)$$

where E_s and T_s are respectively the symbol energy and duration, $\mathbf{F}^{(J)}$ is the DFT matrix of size $J \times J$, $\mathbf{H}_{n_r,n_t}^{CP} = diag[H_{n_c,n_t}^{CP}(0),...,H_{n_r,n_t}^{CP}(k),...,H_{n_r,n_t}^{CP}(N_c-1)]$ is the frequency-domain channel matrix between the n_t-th transmit antenna and n_r-th receive antenna, where $H_{n_r,n_t}^{CP}(k)$ is the channel gain at the k-th frequency. $\mathbf{N}_{n_r}^{CP} = [N_{n_r}^{CP}(0),...,N_{n_r}^{CP}(k),...,N_{n_r}^{CP}(N_c-1)]^T$ is the frequency-domain noise vector. The k-th element of $\mathbf{N}_{n_r}^{CP}$ is the zero-mean additive white Gaussian noise (AWGN) having the variance $2N_0/T_s$ with N_0 being the one-sided noise power spectrum density.

From Eq. (1), the $N_rN_c \times 1$ overall frequency-domain received signal $\mathbf{Y}^{CP} = [\{\mathbf{Y}_1^{CP}\}^T,...,\{\mathbf{Y}_{N_r}^{CP}\}^T]^T$ is given by

$$\mathbf{Y}^{CP} = \sqrt{\frac{2E_s}{T_s}} \begin{bmatrix} \mathbf{H}_{1,1}^{CP}\mathbf{F}^{(N_c)} & \cdots & \mathbf{H}_{1,N_t}^{CP}\mathbf{F}^{(N_c)} \\ \vdots & \ddots & \vdots \\ \mathbf{H}_{N_r,1}^{CP}\mathbf{F}^{(N_c)} & \cdots & \mathbf{H}_{N_r,N_t}^{CP}\mathbf{F}^{(N_c)} \end{bmatrix} \begin{bmatrix} \mathbf{d}_1 \\ \vdots \\ \mathbf{d}_{N_t} \end{bmatrix} + \begin{bmatrix} \mathbf{N}_1^{CP} \\ \vdots \\ \mathbf{N}_{N_r}^{CP} \end{bmatrix},$$
$$= \sqrt{\frac{2E_s}{T_s}} \mathbf{H}^{CP} \mathbf{d}_{all} + \mathbf{N}^{CP} \qquad (2)$$

where $\mathbf{H}^{CP}$ is an equivalent channel matrix of size $N_rN_c \times N_tN_c$, which is a concatenation of the space and frequency-domain channel and DFT, $\mathbf{d}_{all}$ is the $N_tN_c \times 1$ overall data symbol vector, and $\mathbf{N}^{CP}$ is the $N_rN_c \times 1$ overall noise vector.

B. TS-SC MIMO

The TS-SC MIMO block transmission system model and the block structure are illustrated in Fig. 1 and Fig. 2 (b), respectively. CP is replaced by known TS. In order to let TS to play the role of CP, DFT size at the receiver must be the sum of the number of useful data symbols and the TS length. In this paper, to keep the same data rate as CP-SC, the data symbol block length and the TS length need to be set to N_c and N_g, respectively. The difference between TS-SC and CP-SC is the size of DFT to be used at the receiver; the DFT size is N_c+N_g symbols for TS-SC while it is N_c symbols for CP-SC.

The data symbol block of n_t-th transmit antenna can be expressed similar to CP-SC MIMO. Before the transmission,

the TS of length N_g symbols is appended at the end of each block. The block $\mathbf{s}_{n_t}=[s_{n_t}(0),...,s_{n_t}(t),...,s_{n_t}(N_c+N_g-1)]^T$ to be transmitted at the n_t-th transmit antenna is expressed using the vector form as

$$\mathbf{s}_{n_t}=\begin{bmatrix}\mathbf{d}_{n_t}\\ \mathbf{u}_{n_t}\end{bmatrix}, \tag{3}$$

where $\mathbf{u}_{n_t}=[u_{n_t}(0),...,u_{n_t}(t),...,u_{n_t}(N_g-1)]^T$ denotes the TS vector which is identical for all blocks.

In TS-SC MIMO, the frequency-domain received signal vector $\mathbf{Y}_{n_r}^{TS}=[Y_{n_r}^{TS}(0),...,Y_{n_r}^{TS}(k),...,Y_{n_r}^{TS}(N_c+N_g-1)]^T$ after an N_c+N_g-point DFT is expressed as

$$\mathbf{Y}_{n_r}^{TS}=\sqrt{\frac{2E_s}{T_s}}\sum_{n_t=1}^{N_t}\mathbf{H}_{n_r,n_t}^{TS}\mathbf{F}^{(N_c+N_g)}\mathbf{s}_{n_t}+\mathbf{N}_{n_r}^{TS}, \tag{4}$$

where $\mathbf{H}_{n_r,n_t}^{TS}=diag[H_{n_r,n_t}^{TS}(0),...,H_{n_r,n_t}^{TS}(k),...,H_{n_r,n_t}^{TS}(N_c+N_g-1)]$ is the frequency-domain channel matrix and $\mathbf{N}_{n_r}^{TS}=[N_{n_r}^{TS}(0),...,N_{n_r}^{TS}(k),...,N_{n_r}^{TS}(N_c+N_g-1)]^T$ is the frequency-domain noise vector. From Eq. (4), the $N_r(N_c+N_g)\times 1$ overall frequency-domain received signal $\mathbf{Y}^{TS}=[\{\mathbf{Y}_1^{TS}\}^T,...,\{\mathbf{Y}_{N_r}^{TS}\}^T]^T$ is given by

$$\begin{aligned}\mathbf{Y}^{TS}&=\sqrt{\frac{2E_s}{T_s}}\begin{bmatrix}\mathbf{H}_{1,1}^{TS}\mathbf{F}^{(N_c+N_g)} & \cdots & \mathbf{H}_{1,N_t}^{TS}\mathbf{F}^{(N_c+N_g)}\\ \vdots & \ddots & \vdots\\ \mathbf{H}_{N_r,1}^{TS}\mathbf{F}^{(N_c+N_g)} & \cdots & \mathbf{H}_{N_r,N_t}^{TS}\mathbf{F}^{(N_c+N_g)}\end{bmatrix}\begin{bmatrix}\mathbf{s}_1\\ \vdots\\ \mathbf{s}_{N_t}\end{bmatrix}+\begin{bmatrix}\mathbf{N}_1^{TS}\\ \vdots\\ \mathbf{N}_{N_r}^{TS}\end{bmatrix},\\ &=\sqrt{\frac{2E_s}{T_s}}\mathbf{H}^{TS}\mathbf{s}_{all}+\mathbf{N}^{TS}\end{aligned} \tag{5}$$

where $\mathbf{H}^{TS}$ is an equivalent channel matrix of size $N_r(N_c+N_g)\times N_t(N_c+N_g)$, $\mathbf{s}_{all}$ is the $N_t(N_c+N_g)\times 1$ overall transmit symbol vector, and $\mathbf{N}^{TS}$ is the $N_r(N_c+N_g)\times 1$ overall noise vector.

III. QRM-MLBD

A. Operation Principle of QRM-MLBD

QRM-MLBD consists of three steps; ordering, QR decomposition, and M-algorithm. The order of the symbols in the overall symbol vector $\mathbf{d}_{all}$ and $\mathbf{s}_{all}$ can be arbitrarily changed at the receiver by exchanging the corresponding columns in the equivalent channel matrix $\mathbf{H}^{CP}$ and $\mathbf{H}^{TS}$. In [10], the ordering method suitable for SC-MIMO using QRM-MLBD was proposed. The ordered overall data symbol vector $\mathbf{d}_{all}^{order}$ for CP-SC MIMO and the ordered overall transmit symbol vector $\mathbf{s}_{all}^{order}$ for TS-SC MIMO can be respectively expressed as

$$\begin{aligned}\mathbf{d}_{all}^{order}&=[d_1(0),...,d_{N_t}(0),...,d_1(N_c-1),...,d_{N_t}(N_c-1)]^T\\ &=[\mathbf{d}^T(0),\mathbf{d}^T(1),...,\mathbf{d}^T(N_c-1)]^T\end{aligned} \tag{6}$$

and

$$\begin{aligned}\mathbf{s}_{all}^{order}&=[s_1(0),...,s_{N_t}(0),...,s_1(N_c+N_g-1),...,s_{N_t}(N_c+N_g-1)]^T\\ &=[\mathbf{d}^T(0),\mathbf{d}^T(1),...,\mathbf{d}^T(N_c-1),\mathbf{u}^T(0),...,\mathbf{u}^T(N_g-1)]^T\end{aligned}, \tag{7}$$

where $\mathbf{d}(t)=[d_1(t),...,d_{N_t}(t)]^T$ and $\mathbf{u}(t)=[u_1(t),...,u_{N_t}(t)]^T$ denote the data symbol vector and TS vector at t-th symbol of size $N_t\times 1$, respectively.

After ordering, the QR decomposition is applied to the ordered equivalent channel matrix to obtain $\mathbf{H}^{CP(\text{or }TS)}=\mathbf{Q}^{CP(\text{or }TS)}\mathbf{R}^{CP(\text{or }TS)}$, where $\mathbf{Q}^{CP(\text{or }TS)}$ is a unitary matrix of size $N_rN_c\times N_tN_c$ ($N_r(N_c+N_g)\times N_t(N_c+N_g)$) and $\mathbf{R}^{CP(\text{or }TS)}$ is an upper triangular matrix of size $N_tN_c\times N_tN_c$ ($N_t(N_c+N_g)\times N_t(N_c+N_g)$). The overall frequency-domain received signal vector can be transformed by using $\mathbf{Q}^{CP(\text{or }TS)}$ as

$$\begin{cases}\hat{\mathbf{Y}}^{CP}=\{\mathbf{Q}^{CP}\}^H\mathbf{Y}^{CP}\\ \qquad=\sqrt{\dfrac{2E_s}{T_s}}\mathbf{R}^{CP}\mathbf{d}_{all}^{order}+\{\mathbf{Q}^{CP}\}^H\mathbf{N}^{CP}\\ \hat{\mathbf{Y}}^{TS}=\{\mathbf{Q}^{TS}\}^H\mathbf{Y}^{TS}\\ \qquad=\sqrt{\dfrac{2E_s}{T_s}}\mathbf{R}^{TS}\mathbf{s}_{all}^{order}+\{\mathbf{Q}^{TS}\}^H\mathbf{N}^{TS}\end{cases}, \tag{8}$$

where $(\cdot)^H$ denotes the Hermitian transpose operation. Vector $\hat{\mathbf{Y}}^{CP(\text{or }TS)}$ is the N_tN_c ($N_t(N_c+N_g)$)-dimensional transformed received signal vector.

From Eq. (8), the ML solution can be obtained by searching for the best path having the minimum Euclidean distance in the tree diagram. The number of stages in the tree diagram is N_tN_c for CP-SC MIMO and $N_t(N_c+N_g)$ for TS-SC MIMO as shown in Fig. 3. M-algorithm [15] can be applied to reduce the computational complexity. The M-algorithm keeps only M paths that have the highest reliability at each stage. In each stage, the best M paths are selected as surviving paths by comparing the path metrics based on the squared Euclidean distance for all surviving paths and are passed to the next stage.

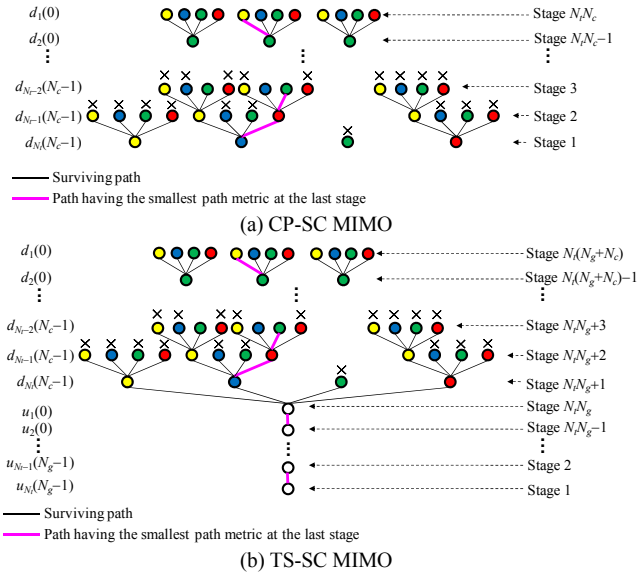


Figure 3. M-algorithm (M=3) with QPSK.

B. Advantage of TS-SC MIMO over CP-SC MIMO

In QRM-MLBD, the received signal power available at early stages likely becomes low since only the elements of upper triangular matrix $\mathbf{R}$ closer to the lower right positions are available. Furthermore, in the case of SC-MIMO block transmission, the magnitude of elements of matrix $\mathbf{R}$ closer to the lower right positions may drop with higher probability [16]. As a result, the received signal power available at early stages in the M-algorithm significantly drops. In the case of CP-SC

MIMO, it can be understood from Eqs. (6) and (8) that the lower right elements of $\mathbf{R}^{CP}$ are relevant to the selection of the surviving path and therefore, the probability of removing the correct path at early stages may increase when smaller M is used. Therefore, a fairly large M must be used to achieve the sufficiently improved performance. However, the use of larger M increases the computational complexity.

On the other hand, it can be understood from Eqs. (7) and (8) that in TS-SC MIMO, TSs are localized at the bottom of the overall transmit signal vector of size $N_t(N_c+N_g)\times 1$. This helps to improve the detection performance of the M-algorithm. Since the symbols to be detected at early stages in the M-algorithm are symbols belonging to the known TSs. Therefore, the probability of removing the correct path can be significantly reduced even if small M is used.

IV. Computer Simulation

The simulation condition is summarized in Table I. We assume N_c=64, N_g=16, and L=16-path frequency-selective block Rayleigh fading channel with uniform power delay profile. The ideal channel estimation is assumed.

TABLE I. Computer Simulation Condition

Channel coding	R=1/3 (13, 15) RSC encoder Log-MAP decoding with 8 iterations	
HARQ	HARQ type II S-P4	
Transmitter	Data modulation	QPSK, 16QAM, 64QAM
	Number of transmit antennas	N_t=2, 4
	Data symbol block length	N_c=64
	TS or CP lengths	N_g=16
Channel	Fading type	Frequency-selective block Rayleigh
	Power delay profile	L=16 path uniform power delay profile
Receiver	Number of receive antennas	N_r=2, 4
	Channel estimation	Ideal

A. Average BER Performance

The average BER performance of uncoded SC-MIMO spatial multiplexing using TS aided QRM-MLBD is plotted in Fig. 4 as a function of average received bit energy-to-noise power spectrum density ratio $E_b/N_0(=(E_s/N_0)(1+N_g/N_c)/\log_2 X)$ for M=1, 4, 16, 64, and 256. 16QAM is assumed as the data modulation scheme. For comparison, the BER performances of the conventional QRM-MLBD and the MMSED are also plotted. It can be seen form Fig. 4 that when small M is used, the achievable BER performance of the conventional QRM-MLBD degrades. On the other hand, TS aided QRM-MLBD achieves better BER performance even if small M is used. This is because the use of TS instead of CP significantly reduces the probability of removing the correct path at early stages in the M-algorithm even if small M is used. When N_t=N_r=2(4), TS aided QRM-MLBD requires M=4(1) to achieve the BER performance similar to the conventional QRM-MLBD with M=1024. Hence, TS aided QRM-MLBD significantly reduces the computational complexity. Computational complexity reduction is discussed in Sect. IV-C. Compared to the MMSED, the required average received E_b/N_0 for the average BER=10^{-3} is reduced by approximately 8.5 and 10dB when M=16.

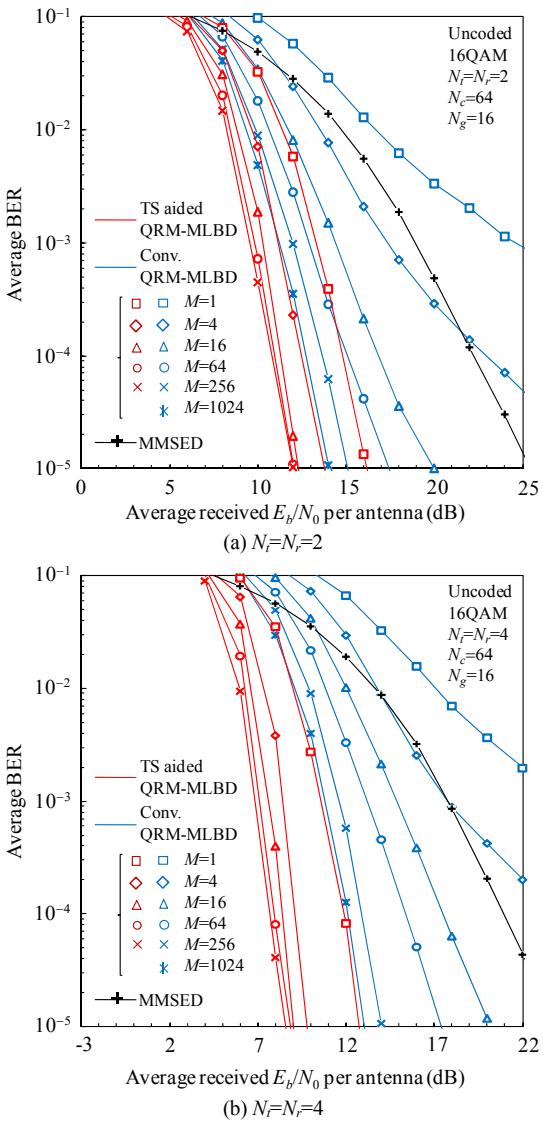


Figure 4. Average BER performance.

B. Throughput Performance

In this paper, we consider the HARQ type II S-P4 [17] and turbo coding. The detail of the HARQ is described in Appendix. We employ a rate 1/3 turbo encoder using two (13, 15) recursive systematic convolutional (RSC) component encoders. Log-MAP decoding with 8 iterations is assumed. The packet size is set to K=3072. The ideal error detection and no transmission error in ACK/NACK are assumed.

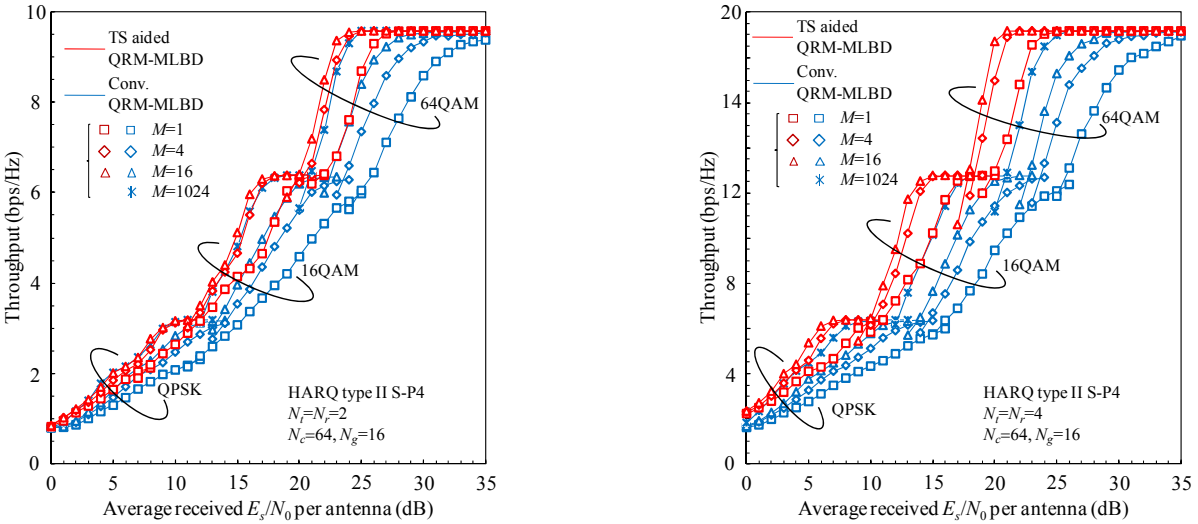
The throughput performance of SC-MIMO HARQ using TS aided QRM-MLBD is plotted in Fig. 5 as a function of average received E_s/N_0 for M=1, 4, and 16. The modulation level which gives the best throughput is selected for each E_s/N_0. For comparison, the throughput performance of SC-MIMO HARQ using the conventional QRM-MLBD is also plotted. It can be seen from Fig. 5 that TS aided QRM-MLBD provides better throughput performance than the conventional QRM-MLBD even if the smaller M is used. When N_t=N_r=2(4), TS aided QRM-MLBD with M=16(4) achieves better throughput performance than the conventional QRM-MLBD with M=1024 in all E_s/N_0 region.

Figure 6 compares the throughput performances of SC-MIMO HARQ using TS aided QRM-MLBD, the conventional QRM-MLBD, and MMSED. It can be seen from Fig. 6 that the throughput performance is significantly improved in a high E_s/N_0 region by using TS aided QRM-MLBD. When N_t=N_r=2 and M=16 is used, the E_s/N_0 reduction from MMSED is about 3.3dB and 9.5dB for the throughput=5.0bps/Hz and 7.5bps/Hz, respectively. When N_t=N_r=4 and M=16 is used, the E_s/N_0 reduction from MMSED is about 5.0dB and 13.2dB for the throughput=10bps/Hz and 15bps/Hz, respectively. It can be seen from the results that the use of TS-aided QRM-MLBD is more effective when high level modulation is used.

C. *Computational Complexity*

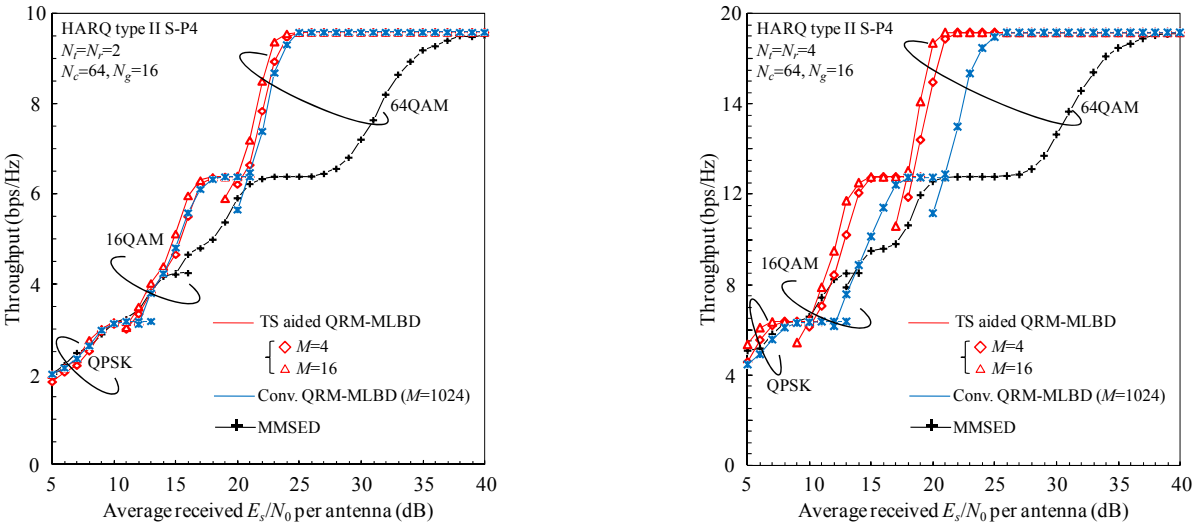
The computational complexity of TS aided QRM-MLBD is discussed. The complexity here is defined as the number of complex multiplications. The required number of multiplications is shown in Table II. First, we discuss the complexity comparison between TS aided QRM-MLBD and the conventional QRM-MLBD. TS aided QRM-MLBD achieves better throughput performance even if small M is used. Hence, TS aided QRM-MLBD reduces significantly the computational complexity required for the squared Euclidean distance calculations in the M-algorithm. It can be seen from Figs. 4 and 5 that when N_t=N_r=2(4), TS aided QRM-MLBD with M=4(1) achieves almost the same BER and throughput performances as the conventional QRM-MLBD with M=1024. When N_t=N_r=2(4), the overall computational complexity of TS aided QRM-MLBD with M=4(1) is about 12(22) %, 3.4(6.0) %, and 1.1(1.6) % of that of the conventional QRM-MLBD with M=1024 for QPSK, 16QAM, and 64QAM, respectively (N_c=64 and N_g=16).

Next, we discuss the complexity comparison between TS aided QRM-MLBD and MMSED. TS aided QRM-MLBD provides significantly higher throughput performance at the cost of increased complexity. When N_t=N_r=2(4), TS aided QRM-MLBD with M=4 requires about 20(58), and 39(102) times higher computational complexity than MMSED at E_s/N_0=17 and 27dB, respectively (N_c=64 and N_g=16).



(a) N_t=N_r=2 (b) N_t=N_r=4

Figure 5. Throughput performance of SC-MIMO HARQ using TS aided QRM-MLBD and the conventional QRM-MLBD.



(a) N_t=N_r=2 (b) N_t=N_r=4

Figure 6. Throughput performance comparison between TS aided QRM-MLBD, conventional QRM-MLBD, and MMSED.

TABLE II. NUMBER OF MULTIPLICATIONS

TS aided QRM-MLBD	DFT	$N_r(N_c+N_g)^2$
	QR decomposition	$N_rN_t^2(N_c+N_g)^3+N_rN_t(N_c+N_g)^2$
	Multiplication of $\mathbf{Q}^H$	$N_rN_t(N_c+N_g)^2$
	Squared Euclidian distance calculations	$X\{2+(M/2)(N_tN_c+4)(N_tN_c-1)\}$ $+N_t^2N_g(N_c+N_g)$
Conventional QRM-MLBD	DFT	$N_rN_c^2$
	QR decomposition	$N_rN_t^2N_c^3+N_rN_tN_c^2$
	Multiplication of $\mathbf{Q}^H$	$N_rN_tN_c^2$
	Squared Euclidian distance calculations	$X\{2+(M/2)(N_tN_c+4)(N_tN_c-1)\}$
MMSED	DFT/IDFT	$N_r(N_c+N_g)^2+N_t(N_c+N_g)^2$
	Weight generation	$(N_r^3+2N_r^2N_t)(N_c+N_g)$
	Weight multiplication	$N_tN_r(N_c+N_g)$
	LLR calculation	$N_t\{N_t(N_r+1)+N_r+2\}(N_c+N_g)$ $+2XN_tN_c$

V. CONCLUSION

To realize the high-speed packet access, the use of MIMO multiplexing and high level modulation is essential. However, with MMSED, the transmission performance of SC-MIMO significantly degrades due to the ISI and IAI. In this paper, to improve the transmission performance of SC-MIMO, we presented to use a combination of TS-SC and QRM-MLBD. We evaluated, by computer simulation, the BER and HARQ throughput performances of SC-MIMO using TS aided QRM-MLBD. We showed that the TS aided QRM-MLBD improves the BER and throughput performances of SC-MIMO compared to the conventional QRM-MLBD while significantly reducing the computational complexity. We also showed that TS aided QRM-MLD provides significantly higher throughput performance than MMSED especially when high level modulation is used. Hence, the TS-aided QRM-MLBD is promising for high-speed packet access.

REFERENCES

[1] Y. Kim, B. J. Jeong, J. Chung, C.S. Hwang, J.S. Ryu, K. H. Kim, and Y. J. Kim, "Beyound 3G; vision, requirements, and enabling technologies," IEEE Commun. Mag., Vol. 41, No. 3, pp. 120-124, Mar. 2003.

[2] J. G. Proakis and M. Salehi, *Digital communications*, 5th ed., McGraw-Hill, 2008.

[3] K. Higuchi, H. Kawai, N. Maeda, H. Taoka, and M, Sawahashi, "Experiments on real-time 1-Gb/s packet transmission using MLD-based signal detection in MIMO-OFDM broadband radio access," IEEE Journal on Selected Areas in Commun., Vol. 24, No. 6, pp. 1141-1153, June 2006.

[4] D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar, and B. Edison, "Frequency domain equalization for single-carrier broadband wireless systems," IEEE Commun. Mag., Vol. 40, No. 4, pp. 58-66. Apr. 2002.

[5] F. Adachi, H. Tomeba, and K. Takeda, "Introduction of frequency-domain signal processing to broadband single-carrier transmissions in a wireless channel," IEICE Trans. Commun., Vol.E92-B, No.09, pp. 2789-2808, Sept. 2009.

[6] G. J. Foschini and M. J. Gans, "On limit of wireless communications ini a fading environment when using multiple antennas," Wireless Personal Commun., Vol. 6, No. 3, pp. 311-335, 1998.

[7] A. van Zelst, R. van Nee, and G. A. Awater, "Space division multiplexing (SDM) for OFDM systems," Proc. IEEE 51st Vehicular Technology Conference (VTC2000-Spring), Vol. 2, pp. 1070-1074, May 2000.

[8] A. Nakajima, D. Garg, and F. Adachi, "Throughput of turbo coded hybrid ARQ using single-carrier MIMO multiplexing," Proc. IEEE 61st Vehicular Technology Conference (VTC2005-Spring), Vol. 1, pp. 610-614, 30 May-1 June 2005.

[9] L. J. Kim and J. Yue, "Joint channel estimation and data detection algorithms for MIMO-OFDM systems," Proc. Thirty-Sixth Asilomar Conference on Signals, System and Computers, pp. 1857-1861, Nov. 2002.

[10] K. Nagatomi, K. Higuchi, and H. Kawai, "Complexity reduced MLD based on QR decomposition in OFDM MIMO multiplexing with frequency domain spreading and code multiplexing," Proc. IEEE Wireless Communications and Networking Conference (WCNC 2009), pp. 1-6, Apr. 2009.

[11] L. Deneire, B. Gyselinckx, and M. Engels, "Training sequence versus cyclic prefix - a new look on single carrier communication," IEEE Commun. Lett., Vol. 5, No. 7, pp. 292-294, July, 2001.

[12] T. Yamamoto, K Takeda and F. Adachi, "Frequency-domain block signal detection with QRM-MLD for training sequence-aided single-carrier transmission," EURASIP Journal on Advances in Signal Processing, Vol. 2011, Article ID 575706, 12 pages, 2010.

[13] T. Yamamoto, K. Takeda, and F. Adachi, "Training sequence-aided QRM-MLD block signal detection for single-carrier MIMO spatial multiplexing," Proc. IEEE International Conference on Communications (ICC 2011), Jun. 2011.

[14] S. Lin and D. J. Costello, *Error control coding: Fundamentals and Aplications*, Prentice Hall, 1983.

[15] J. B. Anderson and S. Mohan, "Sequential coding algorithms: A survey and cost analysis," IEEE Trans. on Commun., Vol. 32, pp. 169-176, Feb. 1984.

[16] K. Takeda, H. Tomeba, and F. Adachi, "Joint Tomlinson-Harashima precoding and frequency-domain equalization for broadband single-carrier transmission," IEICE Trans. Commun., Vol. E91-B, No. 1, pp. 258-266, Jan. 2008.

[17] D. Garg and F. Adachi, "Throughput comparison of turbo-coded HARQ in OFDM, MC-CDMA and DS-CDMA with frequency-domain equalization," IEICE Trans. Commun., Vol. E88-B, No.2, pp.664-677, Feb. 2005.

APPENDIX

In this paper, we consider the HARQ type II S-P4 and turbo coding with rate R=1/3, as illustrated in Fig. 7. The turbo encoder outputs the systematic bit sequence and two parity bit sequences. These sequences are punctured into five sequences (including systematic bit sequence) by the puncturing matrices given by

$$\begin{bmatrix}1&1&1&1\\0&0&0&0\\0&0&0&0\end{bmatrix}\begin{bmatrix}0&0&0&0\\1&0&0&0\\0&0&1&0\end{bmatrix}\begin{bmatrix}0&0&0&0\\0&1&0&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}0&0&0&0\\0&0&1&0\\1&0&0&0\end{bmatrix}\begin{bmatrix}0&0&0&0\\0&0&0&1\\0&1&0&0\end{bmatrix}, \quad \text{(A.1)}$$

For the first transmission, only the systematic bit sequence is transmitted. At the receiver, data decision and error detection are performed. If any error is detected in the received packet, second transmission is requested from the receiver by sending an NACK signal. When the NACK signal is received at the transmitter, the second packet (consisting of the punctured parity bit sequence) is transmitted. At the receiver, turbo decoding is carried out by using the first and second received packets. If any error is detected after turbo decoding, the NACK signal is transmitted again. One of the punctured parity bit sequences is transmitted each time the NACK signal is received at the transmitter until the 5th packet transmission. After the 5th packet transmission, the same packet is retransmitted.

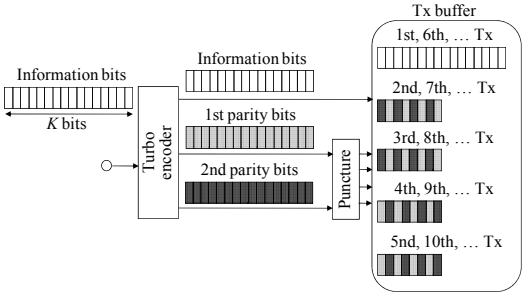


Figure 7. HARQ type II S-P4.

Study about the Parameters of the Holographic Impedance Modulated Surface Antenna

Jian.Zhang Jun.Ouyang Feng.Yang
Department of microwave
University of electronic science technology of china
Chengdu, China
sczjzhx@126.com

Abstract—**In this paper, holographic impedance modulated surface antenna is simplified to a 1 - D model which is equal to a leaky-wave antenna with sinusoidally modulated surface reactance. Comparing with leaky-wave antenna with sinusoidally modulated surface reactance after mathematical derivation, we get to the conclusion that both modulation depth and average surface impedance control leakage speed of holographic impedance modulated surface, and further affect antenna gain. Later, in order to verify this conclusion, three examples are simulated by full- wave numerical simulation software.**

Keywords-Holographic antenna; holographic impedance modulated surface antenna; leak waves; Parameters;

I. INTRODUCTION

The first holographic image was created by Dennis Gabor in 1948 [1, 2], and the holographic technique had obtained widely application in the optical region from that moment on. Consequently, Dennis Gabor won the noble prize in 1971 due to this great invention. Following the intensification of holographic activity in the optical community, some professors and experts in microwave attempted to extend the concept of holography from optical band to microwave band while plenty of papers focusing on this topic were published in 1960s-1970s. In 1965, R. P. DOOLE employed the holographic two-step process to record the images of objects at X-band, and he reconstructed these objects at optical frequency [3]. As a new branch of electromagnetics, microwave holography was really established by Kock later in 1968 when he verified the feasibility to combine this theory with antenna to design holographic antenna [4]. Influenced by this new progress ,many researchers had investigated about electromagnetic characteristics of holographic antenna and discussed their operation mechanism, what's more, plenty of types of holographic antennas were fabricated and tested[5-13]: encountered with the practical problem that there is no recording materials at microwaves as photographic plate in optics, different kinds of approximations of the interference pattern for holography were constructed and tested by Checcacci and Russo etc. in 1970 [5]. On the other hand, in order to overcome the disadvantage of bidirectional radiation property of planar holographic antenna, Iizuka, K. and Mizusawa, M introduced a novel volume-type holographic which consisted two or more hologram plane to cancel the beam in one direction and concentrate the power in the opposite direction[7]. In 2004, ElSherbiny and Fathy utilized the surface wave to feed holographic antenna to remedy some drawbacks of traditional holographic antenna such as: low coupling efficiency and grating-lobe problems caused by sparse space between neighbouring elements [9]. Meanwhile, with the rapid development of artificial electricmagnetic material from 1950s, the artificial impedance surface has became the hot topic among antenna experts and was discussed frequently these years. It is notable that Thomas, Zucker[14] and Pease[15] told us that sinusoidally modulated grounded dielectric slab with metallic strips can produce beams at arbitrary angles and systematic analysis was made by A. Oliner in 1959 along with rigorous derivations and explicit equations[16].On the basis of modulated reactance surface and leaky wave antenna, Bryan H. Fong and his group offered a fresh holographic antenna concept called holographic impedance modulated surface antenna[17]. General speaking, we may treat aperture of traditional holographic antenna as antenna arrays and the elements lie along the zero contour line of the holographic interference pattern, and then the holographic elements interact with the space wave feed to scatter the power to form the farfield radiation. Compared to traditional holographic antenna, the holographic impedance modulated surface antenna possesses a holographic impedance surface with surface impedance distribution determined by holographic interference. When the surface wave meets the periodic modulation, some leaky wave modes arise even though most of other discrete spatial harmonics are bound wave modes. So, in a sense, holographic impedance modulated surface antenna is a type of leaky wave antenna. In paper [17], the authors provide us three methods to obtain the effective surface impedance and later modulate an impedance surface in a holographic fashion as follow:

$$Z = j[X_s + M\,\mathrm{Re}(\psi_{rad}\psi_{ref}^{*})] \tag{1}$$

But unfortunately, the paper lacks necessary discussions about the key parameters— M and X_s of holographic impedance modulated surface antenna, especially for X_s . For this paper, we plan to further study how these parameters, including X_s and M , which refers to average surface impedance and modulation depth respectively, affect the performance of holographic impedance modulated surface antenna. Especially, we finally get to the conclusion that X_s will also affect the radiation efficiency as well as M through mathematic derivation, and we will use some numerical simulated examples resulted from commercial full wave simulation software (cst) to demonstrate our deduction.

II. STUDY OF THE PARAMETERS

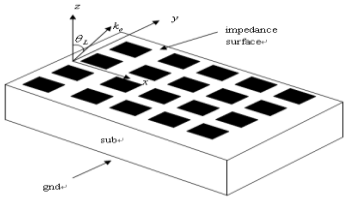


Figure.1 part of holographic impedance surface

$$\begin{cases} \psi_{rad} = e^{jk_0 x \sin\theta_L} \\ \psi_{ref} = e^{-jk_0 nr} \end{cases} \quad (2)$$

Where k_0 is the wave number of free space, n is effective refraction index of impedance surface, $r = \sqrt{x^2+y^2}$, and θ_L is the desired radiation angle from the broadside of the impedance . As a result,

$$Z(x,y) = j[X_s + M\cos k_0(x\sin\theta_L + n\sqrt{x^2+y^2})] \quad (3)$$

Here, we simplify this problem to a 1-D problem which mean we just only observe the cross section (x-z plane, see in Figure. 2) of holographic impedance modulated surface antenna,

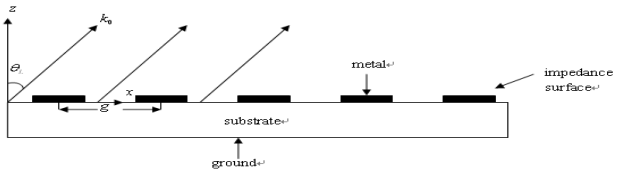


Figure.2

and we make y=0 at last, so,

$$Z(x) = j[X_s + M\cos k_0(\sin\theta_L + n)x] \quad (4)$$

Obviously, we can easily find that not only the structure showed in Figure.2 is quite similar to that in the Figure. 1(b) in [18], but also the equation above is nothing different from (1) in [18] except which symbols are employed. So, it is quite reasonable that we apply the conclusions about the sinusoidally modulated reactance Surface provided in [18] to holographic impedance modulated surface antenna here. Further, we separate X_S out:

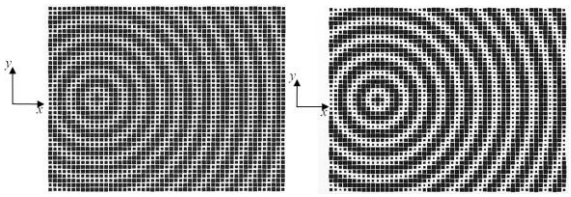
$$Z(x) = j\{X_s[1+\frac{M}{X_S}\cos\frac{2\pi x}{\frac{2\pi}{k_0(\sin\theta_L+n)}}]\} \quad (5)$$

Comparing this equation to (1) in [18], we know that $\frac{2\pi}{k_0(\sin\theta_L+n)}$ determines the modulation period directly, and when k_0 and n are fixed, θ_L will be the dominating factor. On the other hand, as the attenuation constant—α of guided wave along the impedance surface is depicted in [18], here we observe that not only the modulation –depth(M), but also the average surface impedance(X_S) control the leakage speed from the guided surface wave from holographic modulated impedance surface, and we have enough reasons to believe that the lager $\frac{M}{X_S}$ is , the faster the leakage occurs and then different antenna gains emerge. So in this way, different aperture efficiency or antenna gains are obtained through different choices of $\frac{M}{X_S}$. In the following section, some simulation examples are given out to verify our viewpoints.

III. SIMULATION

Example.1

Firstly, on the purpose of verifying the truth that the gains of holographic impedance modulated surface antennas are controlled by the modulation-depth (M), two simulation examples are implemented with macro modeling process in cst microwave studio, and the antenna structures are showed in Figure. 3:



(a) M=18.25 (b) M=36.5

Figure.3 structures of these impedance surfaces with different modulation depths

As demonstrated in Figure. 3, these two holographic impedance modulated surfaces are fabricated on the 1.6mm thick grounded substrate, the dielectric constant of which is about 2.2(Rogers 5880), and we can easily find that the surfaces are also consisted by square patches and the gaps between them. Meanwhile, the same lattice constant—3mm are chosen for these surfaces as that described in [17]. The modulate depth (M) for the impedance surface in Figure. 3(a), (b) is 18.25 and 36.5 respectively with fixed average surface impedance (X_S) 197.5. The size of these holographic impedance modulated surface antennas is 153mm by (x-dimension, 51 patches) 123 mm(y-dimension, 41 patches).

Series of concentric circles are observed due to broadside radiation property, and these antennas work at 22 GHz, and we place the 3mm monopole antenna around the center of these concentric circles to feed these impedance surfaces.

Figure.5 and Figure.6 are the radiation pattern for these two holographic impedance modulated surface antennas at x-z plane ($phi = 0^o$) while the max gains of these two cases are listed in Table.1 and Table.2. When these antennas work at 22GHz, the M=36.5 case has the gain about 17.74dB, approximately 3dB than M=18.25 case. Similar phenomenon happens when operation frequency remove to 24GHz. So, it is definitely reasonable to believe that the modulation depth (M) affects the speed of leakage from holographic impedance modulated surface.

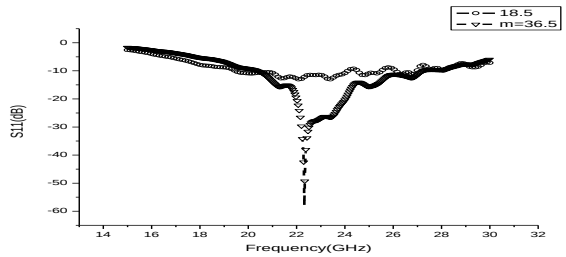
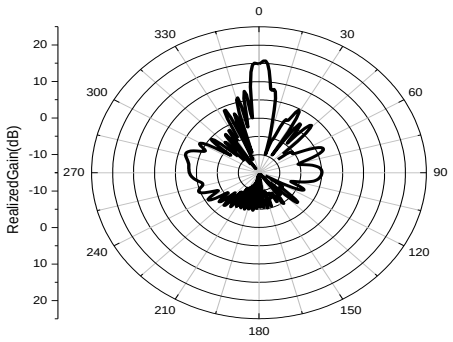
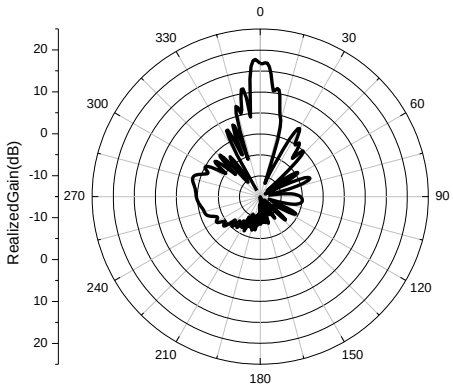


Figure.4 return loss



(a) M=18.25

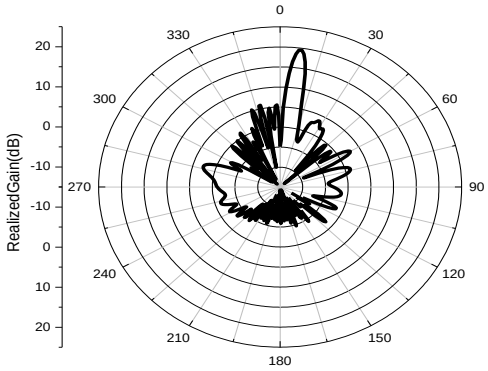


(b) M=36.5

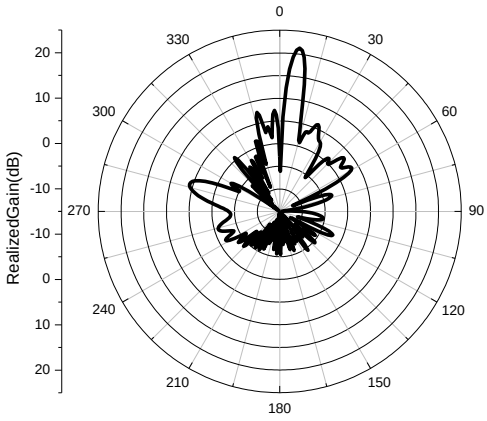
Figure.5 Radiation pattern at 22GHz

M=18.25	15.14dB
M=36.5	17.74dB

Table.1 max gains



(a) M=18.25



(b) M=36.5

Figure.6 Radiation pattern at 24GHz

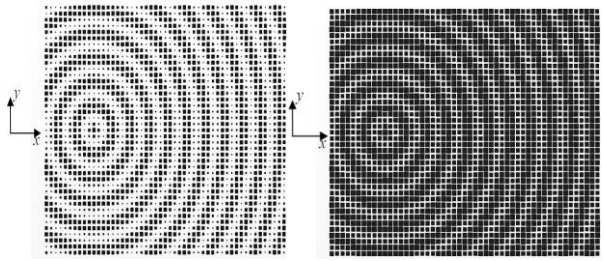
M=18.25	19.49dB
M=36.5	21.27dB

Table.2 max gains

Figure.5 and Figure.6 are the radiation pattern for these two holographic impedance modulated surface antennas at x-z plane ($phi = 0^o$) while the max gains of these two cases are listed in Table.1 and Table.2. When these antennas work at 22GHz, the M=36.5 case has the gain about 17.74dB, approximately 3dB than M=18.25 case. Similar phenomenon happens when operation frequency remove to 24GHz. So, it is definitely reasonable to believe that the modulation depth (M) affects the speed of leakage from holographic impedance modulated surface.

Example.2

Later, we set modulation depth (M) to be 20 as a fixed constant and choose different average surface impedances (X_S) to investigate our deduction:



(a) $X_S = 170$ (b) $X_S = 210$

Figure.7 structures of these impedance surfaces with different average impedances

Comparing these two figures above, great distinguish can be immediately found that the impedance surface with $X_S = 170$ looks much more sparse while the other one is quite dense. The simulated return losses are pictured in the Figure.7. It turns out that the holographic impedance modulated surface antenna achieves good matching (S11-curves remain below -10dB) for both cases from 22GHz to 24GHz, and we pick the 22GHz, 23GHz to draw radiation patterns in x-z plane ($phi = 0^o$) in Figure.8 and Figure.9.

From Table.3, it can be found that great distinction of gain turns up at 22GHz, almost 5.2dB. Although this distinction reduced to about 3.5dB (Table.4), we must say that good agreement between simulated results and our deduction about X_S is acquired. To further validate our ideas and to avoid some accidental, two holographic impedance modulated surface antennas which radiate at about 30^o off broadside have been simulated in example.3:

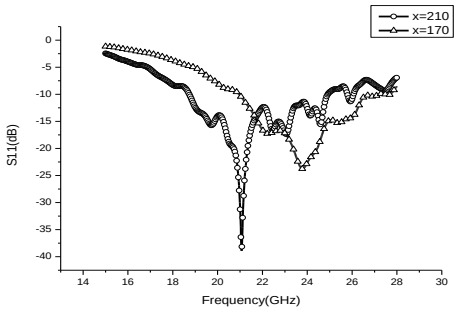
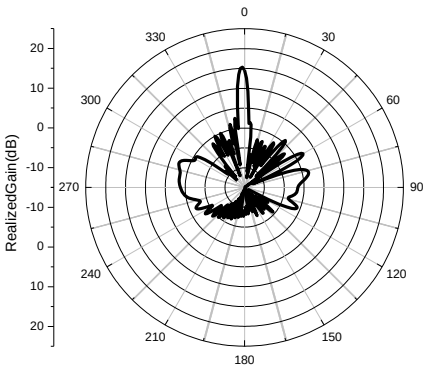
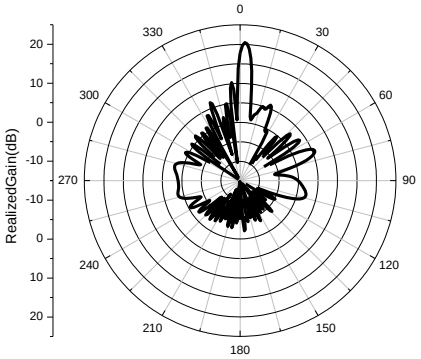


Figure.7 return loss



(a) $X_S = 170$

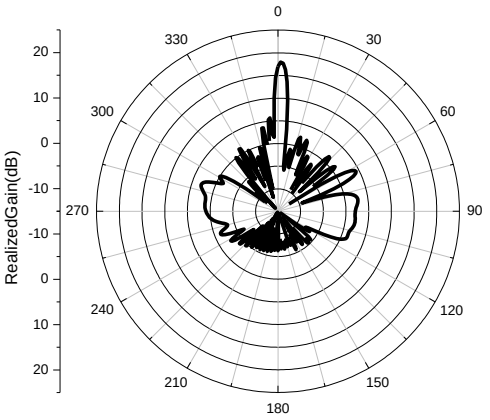


(b) $X_S = 210$

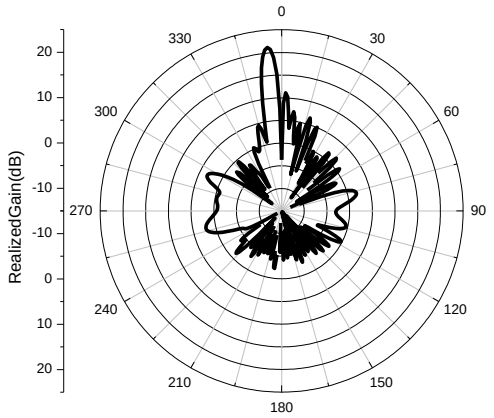
Figure.8 Radiation pattern at 22GHz

X_S =170	15.29dB
X_S =210	20.43dB

Table.3 max gains



(a) $X_S = 170$



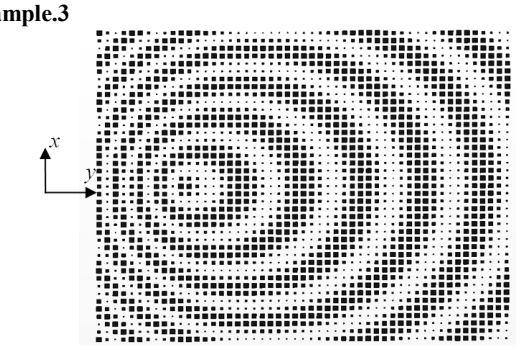
(b) $X_S = 210$

Figure.9 Radiation pattern at 24GHz

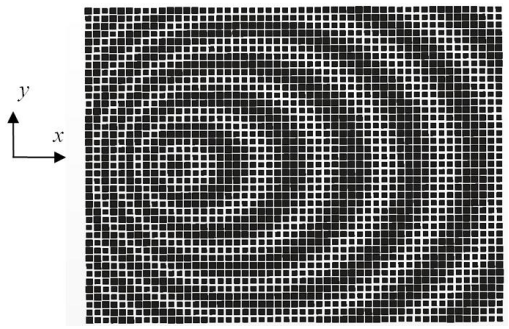
X_S =170	17.88dB
X_S =210	21.2dB

Table4 max gains

Example.3



(a) $X_S = 170$



(b) $X_S = 210$

Figure.10 structures of these impedance surfaces with different average impedances

On account of purposed radiation at about 30^o off broadside, these two impedance surfaces present to be series of ellipses, and the monopole antennas are inserted at the focuses as feeder similarly.

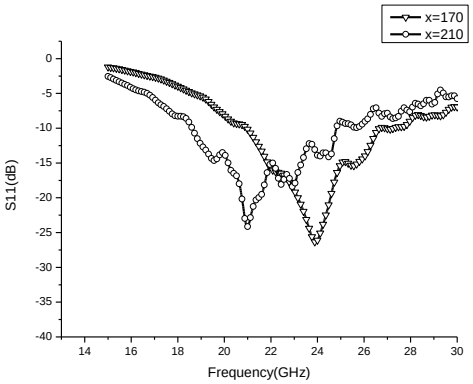
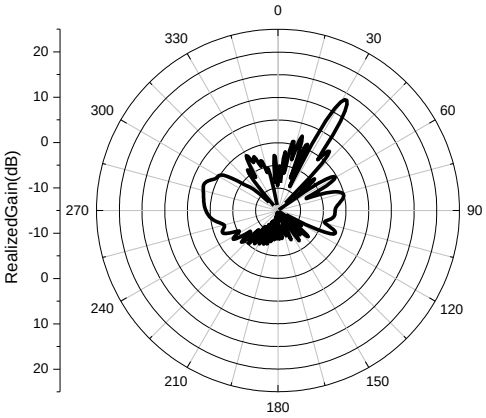
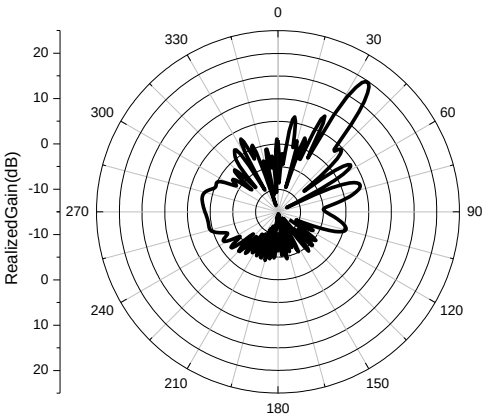


Figure.11 return loss



(a) $X_S = 170$

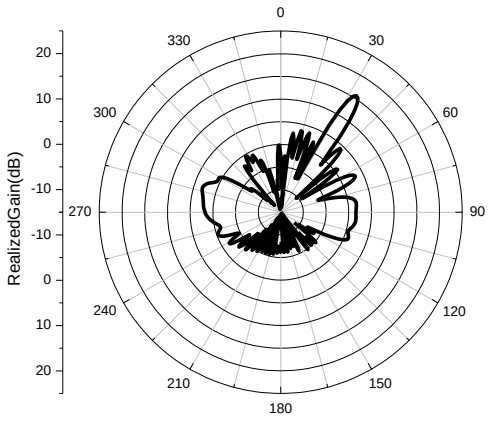


(b) $X_S = 210$

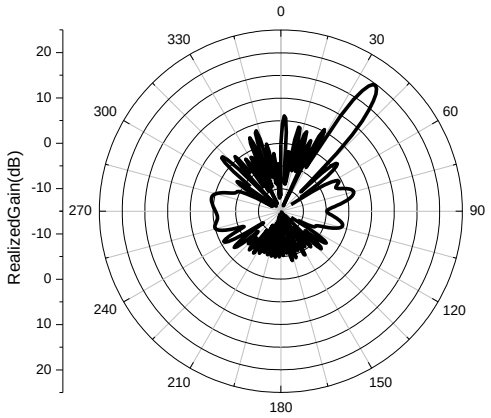
Figure.12 Radiation pattern at 22GHz

X_S =170	13.61dB
X_S =210	19.7 dB

Table.5 max gains



(a) $X_S = 170$



(b) $X_S = 210$

Figure.13 Radiation pattern at 24GHz

X_S =170	15.64dB
X_S =210	19.68dB

Table.6 max gains

Sure enough, different antenna gains can be observed from the results given above.

ACKNOWLEDGMENT

This work was supported by the postdoctoral Science Foundation of China (No. 20090461325, No. 201003690), the Natural Science Foundation of China (No.61001029), the Fundamental Research Funds for the Central Universities(No. 103.1.2 E022050205);

REFERENCES

[1] D. Gabor, "A new microscopic principle," Nature. vol. 161, p. 777, January 1948.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] G. L. Rogers, "Gabor diffraction microscopy: The hologram as a generalizedzone plate," Nature. vol. 166, p. 237, August 1950.

[3] R. P. Dooly, "X-Band Holography," Proceedings of the IEEE. vol. 53, p. 1733, November 1965.

[4] W. E. KOCK, "Microwave holography", Microwaves, 1968, pp. 46-54

[5] Checcacci. P., Russo. V., Scheggi. A., "Holographic antennas," Antennas and Propagation, IEEE Transactions , vol. 18, p. 811, November 1970.

[6] Checcacci. P., Papi. G., Russo, V., "A Holographic VHF Antenna," Antennas and Propagation, IEEE Transactions , vol. 19, p. 278-279, September 1970.

[7] Iizuka. K., Mizusawa. M., Urasaki. S., Ushigome. H., "Volume-type holographic antenna," IEEE Trans. Antennas Propagation.,1975, 23, pp. 807–810 November 1975.

[8] Levis. K. , Ittipiboon. A. , Petosa. A. , Roy. L. , Berini. P. , "Ka-band dipole holographic antennas," Microwaves, Antennas and Propagation, IEE Proceedings, vol 148, p.129, April 2001.

[9] ElSherbiny. M., Fathy. A. E., Rosen. A. ,Ayers. G., Perlow. S. M., "Holographic antenna concept, analysis, and parameters,"Antennas and Propagation, IEEE Transactions ,vol. 52, p. 830-839, March 2004.

[10] Petosa, A., Thirakoune, S., Levis, K., Ittipiboon, A., "Microwave holographic antenna with integrated printed dipole feed," Electronics Letters, vol. 40, pp. 1162, September. 2004.

[11] T. Quach, D.A. McNamara , A. Petosa , "Holographic antenna realised using interference patterns determined in presence of dielectric substrate," Electronics Letters, vol. 41, pp. 724, June. 2005.

[12] P. Sooriyadevan, D.A. McNamara, A. Petosa, A. Ittipiboon, "Electromagnetic modelling and optimisation of a planar holographic antenna," IET Microw. Antennas Propag., vol.1, pp. 693–699, 2007.

[13] A. E. Fathy, M. ElSherbiny, A. Rosen, G. Ayers, S. Perlow, "EXPERIMENTAL DEMONSTRATION OF A 35 GHz HOLOGRAPHIC ANTENNA," Microwave Symposium Digest, 2003 IEEE MTT-S International, vol.3, pp. 1833-1836, 2003.

[14] A. S. Thomas , F. J. Zucker, "Radiation from Modulated Surface Wave Structures-I," IRE NATIONAL COKVENTION RECORD, Pt. 1, pp. 153-160. 1957.

[15] R. L. Pease, "Radiation from modulated surface wave structures- II," IRE CONVENTION RECORD, Pt. 2, pp. 161-165. 1957.

[16] A. Oliner, A. Hessel, "Guided Waves on Sinusoidally-Modulated Reactance Surfaces," IEEE Trans. Antennas Propagation., vol. 7, no. 5, pp. 201-208, December 1959.

[17] Bryan H. Fong, Joseph S. Colburn, John J. Ottusch, John L. Visher, Daniel F. Sievenpiper, "Scalar and Tensor Holographic Artificial Impedance Surfaces," IEEE Trans. Antennas Propagation., vol. 58, pp. 3212-3221, October 2010.

[18] Pate. A. M., Grbic. A., "A Printed Leaky-Wave Antenna with a Sinusoidally Modulated Surface Reactance," Antennas and Propagation Society International Symposium, 2009. APSURSI '09. IEEE, pp.1, June 2009.

Separation of Decorative Characters into Skeleton Parts and Decoration Parts

Yuji Soma
Graduate School of Engineering
Tohoku University
Sendai, Japan
soma@iic.ecei.tohoku.ac.jp

Yoshihiro Sugaya
Graduate School of Engineering
Tohoku University
Sendai, Japan
sugaya@ecei.tohoku.ac.jp

Shinichiro Omachi
Graduate School of Engineering
Tohoku University
Sendai, Japan
machi@ecei.tohoku.ac.jp

Abstract— **Various kinds of decorative characters are used in our daily life. If an optical character reader (OCR) can recognize decorative characters, the system can be used in various fields. However, because the shapes of decorative characters are very complicated, most OCR cannot deal with decorative characters. In this paper, we propose a structure extraction method as a preprocessing of OCR. We also propose a method for generating decorative characters.**

Keywords-structure extraction; decorative character; graph spectral decomposition)

I. INTRODUCTION

As shown in Fig.1, decorative characters are used in our daily life for advertising display, headline, and so on. To attract readers' attention, decorative characters have unusual and complicated appearances.

Figure 1. Decorative characters in daily life

If we can take a picture of decorative characters, recognize the words, and investigate them on the Internet, our life become more comfortable. However, it is difficult to deal with decorative characters by practical optical character readers (OCR) because many OCR systems assume that the characters are described by common fonts. For common OCR systems, the essential structure of the decorative characters should be extracted before inputting into the OCR systems.

Some approaches of structure extraction have been proposed [1] [2]. However, in these methods, decoration features will be lost. If we can extract not only essential structure but also decoration features, we can use them for another purpose. In this paper, we propose a method to separate decorative characters into skeleton parts and decoration parts using graph spectral decomposition. We also propose generating decorative characters using decoration features.

II. GRAPH SPECTRAL DECOMPOSITION

Skeletonization method by graph spectral decomposition has been proposed [2]. The algorithm is composed of following steps:

1) *Translate a decorative character into a graph.*
2) *Calculate Laplacian matrix of the graph.*
3) *Simplify the graph using graph spectral decomposition.*

A. Graph Genaration from Decorative Character

A graph is generated from a decorative character by the following way. In the algorithm, each black pixel is assumed to be a node of a graph. We add an edge between one pixel and each black pixel. Each edge has a weight defined by

$$W(i,j)=\frac{1}{d^4} \tag{1}$$

where $\boldsymbol{W}(i, j)$ indicates the weight of edge between node i and node j, and d represents the Euclidean distance in the image plane. If there is no edge, $\boldsymbol{W}(i, j)$ is set to 0. By these operations, we can get a weight matrix $\boldsymbol{W}$ which represents the adjacency condition between nodes.

B. Calculation of Laplacian Matrix

Laplacian matrix $\boldsymbol{L}$ [3] is an $n \times n$ matrix, where n is the number of nodes which are equivalent to the number of black pixels contained in the character. The matrix represents the structure of the given graph. The Laplacian matrix is defined as follows: Non-diagonal components of $\boldsymbol{L}$ is set to the weight as

$$L(i,j)=-W(i,j) \tag{2}$$

and diagonal components of $\boldsymbol{L}$ is set to the summation of all weights connecting to the center pixel as

$$L(i,i) = \sum_j W(i,j) \tag{3}$$

C. Graph Spectral Decomposition

The eigenvalues and eigenvectors are obtained by eigenvalue decomposition of the Laplacian matrix. This decomposition is called as graph spectral decomposition [3].

Assume node i has a coordinate (x_i, y_i) in image plane, vector $\boldsymbol{X}$ and $\boldsymbol{Y}$ are defined as:

$$\mathbf{X} = (x_1, x_2, \cdots, x_n)^T \tag{4}$$

$$\mathbf{Y} = (y_1, y_2, \cdots, y_n)^T \tag{5}$$

where T denotes a transpose.

The coordinate of each node in image plane are decomposed by the spectral decomposition as follows:

$$\mathbf{X} = (\mathbf{X} \cdot e_1)e_1 + \cdots + (\mathbf{X} \cdot e_n)e_n \tag{6}$$

$$\mathbf{Y} = (\mathbf{Y} \cdot e_1)e_1 + \cdots + (\mathbf{Y} \cdot e_n)e_n \tag{7}$$

where e_i denotes eigenvectors obtained by the eigenvalue decomposition of the Laplacian matrix.

In Eqs. (4) and (5), m eigenvectors from the minimal eigenvalue are used to reconstruct the graph. The resultant graph is condensed according to the number of employed eigenvectors. As an example, we show the result of graph spectral decomposition in Fig.2. (a) shows an input image, (b), (c), and (d) show reconstructed results with first to nth, first to fifth, and sixth to nth eigenvectors, respectively. (c) is the image of low-frequency components of the original one, and (d) is the image of high-frequency components of the original one.

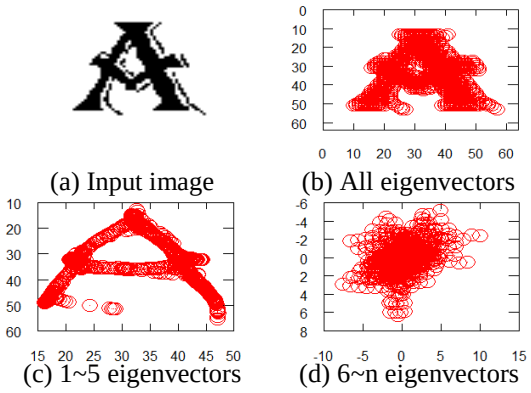


Figure 2. Graph spectral decomposition. (a) Input image, (b) Nodes for1~n eigenvectors, (c) Nodes for 1~5 eigenvectors, (d)Nodes for 6~n eigenvectors

III. Proposed Method

A. Separating Decoration Part

By graph spectral decomposition, decorative characters are decomposed to low frequency component and high frequency component. Now we pay attention to high frequency component. In Fig.3 (a), nodes that are far from origin are colored by blue. This means that these nodes are far from low frequency component. In other words, we can regard these nodes as decorative part. Extracting corresponding nodes from input image, we can separate decoration part like (e). Remaining nodes (d) represent skeleton part.

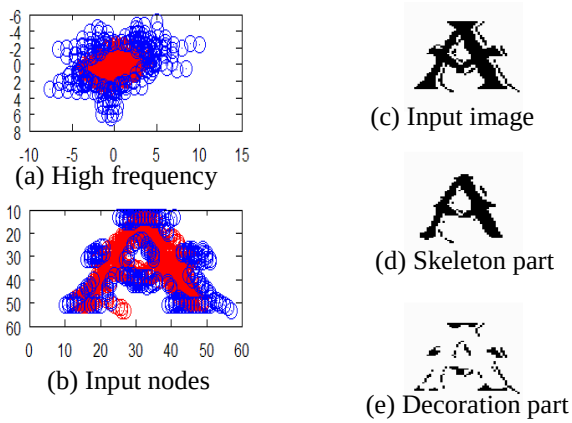


Figure 3. Separating decorative part

B. Generation of Decorative Character

Our purpose is generating decorative character from skeleton of character and other sample decorative character. So we combine low-frequency component of non-decorative character and high-frequency component of decorative character. To find corresponding nodes between characters, we use shape contexts [4]. At first, we apply graph spectral decomposition to both non-decorative character and decorative character. Next we calculate shape context of low-frequency component. Finally, we combine low-frequency component of non-decorative character and corresponding high-frequency component of decorative character. Fig.4 shows an example of finding corresponding nodes. Almost all nodes have good correspondence. However, too high-density nodes and too low-density nodes tend to have wrong correspondence.

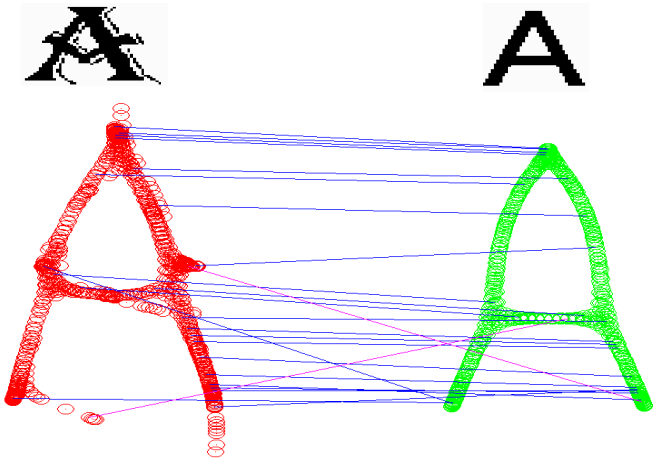


Figure 4. Finding corresponding points of low frequency

IV. EXPERIMENT

We carried out an experiment to separate some decorative characters into skeleton parts and decoration parts. The results are shown in Fig.5. in every sample, skeleton parts and decoration parts are separated successfully.

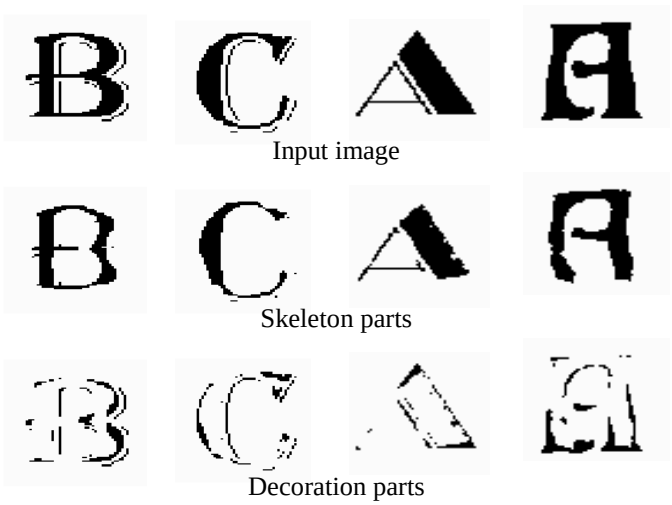


Figure 5. Experimental Result of Separating

Next, we made an experiment to generate a decorative character. We generated decorative character "R" from decorative characters "A", "B" and skeleton of "R". Generated character is shown in Fig.6.

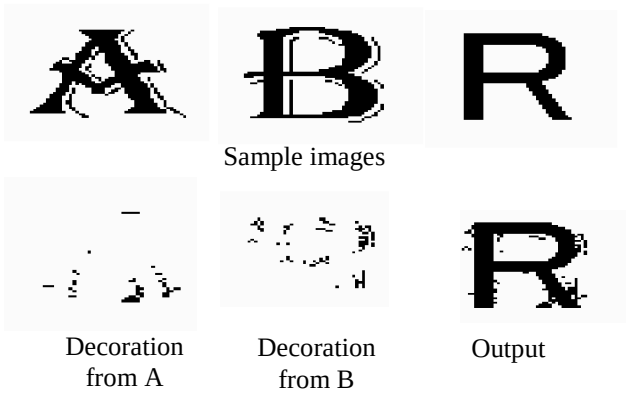


Figure 6. .Experimental result of decorative character generation

Lower part of "R" is similar to that of "A" and upper part of "R" is similar to that of "B". So, bottom decoration was made from "A" and upper decoration was made from "B". As you will know from looking at the result, decoration looks unnatural. This is because process of this method is done pixel by pixel. So smoothing post-processing is necessary.

V. CONCLUSION

We proposed a method of separating characters into skeleton parts and decoration parts. Experimental results show that separation process works well for some kinds of decorative characters. This method can be used for skeletonization of characters. We also proposed a method for generating decorative character from decoration of sample characters. Generated decoration looks a little unnatural yet, so there is room for improvement using smoothing process.

ACKNOWLEDGMENT

This work was partially supported by JST, CREST.

REFERENCES

[1] S. Omachi, M. Inoue, and H. Aso, "Structure Extraction from Decorative Characters Using Multiscale Images," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.23, No.3, March 2001.

[2] Hideaki Kawano, Akito Shimamura, Hideaki Orii, Hiroshi Maeda, and Norikazu Ikoma, "Skeletonization of Decorative Characters by Graph Spectral Decomposition," 2008 IEEE International Conference on Systems, Man and Cybernetics. Singapore, vol. A247, pp. 266–270, Oct. 2008.

[3] F. R. K. Chung, "Spectral Graph Theory," CBMS Lecture Notes, AMS Publication, 1997.

[4] Serge Belongie , Jitendra Malik , Jan Puzicha, "Shape Context: A new descriptor for shape matching and object recognition" Neural Information Processing Systems, pp. 831-837, 2000

A Novel MAC Protocol to Improve Performance in Mobile Ad Hoc Networks

Maksudur Rahman Jonayed and Sung Won Kim
Department of Information and Communication Engineering
Yeungnam University
Gyeongsan-si, Republic of Korea

Abstract—**The collision in wireless networks is a foremost factor of performance degradation. IEEE 802.11 MAC is quite incompetent in mobile ad hoc environment. The collision over the air is much more severe in the mobile ad hoc environment than that in the wireless LANs. Besides the IEEE 802.11 MAC, some multi channel MAC protocols were also proposed to improve the performance of mobile ad hoc networks. Though those improve the overall performance of the networks, there is still collision in the control channel, which causes lower utilization of DATA channel. In this paper we propose a new MAC protocol which uses two communications channels (for control and data frames separately). In our proposed protocol we have considered the collision probability of the control packets in control channel that was not considered in the existing protocols. Analysis results show that our proposed protocol can perform better and reduce the data and control packet collision.**

Keywords- IEEE 802.11, Medium access control, (MAC), Dual-channel, Collision Probability, Mobile Ad Hoc Network, Control channel.

I. INTRODUCTION

Due to low cost and easy implementation, contention based MAC protocol [1] has been deployed widely. IEEE 802.11 MAC, which is the standard for wireless local area network (WLAN), is also integrated in countless wireless simulation packages for mobile ad hoc networks and even real life implementation. The IEEE 802.11 adopts the four way handshaking procedure i.e. Ready to Send (RTS), Clear to Send (CTS), DATA, Acknowledgement (ACK). Packets like RTS and CTS are short in size and used for avoiding collision between long data packets. The network allocation vector (NAV), which is carried by a packet is used to avoid the probable collision and mitigate the hidden terminal problem. The ACK is an affirmation of a successful transmission.

The efficacy of the above mentioned protocol in mobile ad hoc networks has been acknowledged as serious problem. The collision over the air is much more severe in mobile ad hoc network than the wireless LANs [1]-[4]. The packet losses which are results of such MAC layer protocol will affect the performance of the overall networks.

Many schemes were proposed in this issue to reduce the severe collision of DATA packets at the MAC layer. Busy Tone Multiple Access (BTMA) [5] uses a busy tone to address the hidden terminal problem. The base station broadcasts a busy tone signal to keep the hidden terminals for accessing the channel when it senses a transmission. It relies on a centralized network infrastructure which is not applicable in mobile ad hoc networks. FAMA-NCS [6] uses the long dominating CTS packets to act as the busy tone to prevent any competing transmitters in the receiver range from transmitting. This requires any nodes hearing interference keep quiet for the period of one maximum data packet to guarantee no collisions with the ongoing data transmission, which is obviously hot efficient especially when the RTS/CTS negotiation process fails or the DATA packet is very short [7].

Multi-channel protocols were also proposed. The usual approach to avoid collision between control packets and data packets to use separate channel for separate kinds of packets i.e. one channel for control packets and one channel for data packets. Dual busy tone multiple access (DBTMA) [8] uses two separate channel for DATA and control packets. To avoid hidden terminal and exposed terminal problem it uses two busy tone signals. Transmitter-based busy tone is used to prevent the exposed terminal problem and receiver-based busy tone is used to prevent hidden terminal problem. Power Aware Multiple Access (PAMAS) [9] uses a separate control channel to transmit both RTS/CTS packets and busy tone signals. It gives a solution to the hidden terminal problem and mainly focuses on power savings. Dual channel MAC protocol (DUCHA) [7] also uses two channels for data and control packets and a busy tone signal to deal with the hidden terminal problem. It reduces the collision in the data channel significantly but the collision of the control channel was not considered.

Though the performance of data channel depends on the effective use of the control channel, it was not well studied. To the best of our knowledge, there are no extensive studies about the control channel or the probability of control packet collision in control channel. This may lead to a lower utilization of data channel due to the collision of the control packets in control channel. Most of the current schemes concentrate on the data channel and other problems of the MAC protocol.

In this paper, we utilize two channels for control and data packets separately. RTS and CTS are transmitted in control channel to avoid collision with the data packets. An out of band receiver-based busy tone is used to solve the hidden terminal problem. If the data is corrupted due to link error (as there will be no collision in data channel), we use a NACK signal which

is an extended busy tone. Our protocol minimizes the collision in control channel significantly and increases the spatial reuse and network maximum throughput.

The rest of the paper is organized as follows. Section II presents some related works. Section III discusses the proposed MAC protocol. Results and analysis are presented in Section IV. In Section V, concluding remarks are given.

II. RELATED WORKS

A. Overview of Mobile Ad Hoc Networks

For past few decades, the wireless and mobile communication field has experienced an extraordinary growth. The popularity of cellular phone, increasing number of WLAN hot spots allow travelers with portable computers to surf the internet from airports, railways, hotels or other public places. However, these networks are infrastructure with centralized administration. For the devices like laptops, personal digital assistants (PDAs), pocket PCs, smart phones having short range wireless interfaces getting smaller, cheaper, more user friendly and more powerful. And for these devices, a new alternative way for mobile communication is evolving where mobile devices create, organize and administrate the wireless network. This type of network is known as mobile ad hoc network

B. Charecteristics of Ad Hoc Networks

The characteristics and complexities of mobile ad hoc network are summarized in following table. In addition these networks are faced with the traditional problems of wireless network i.e. lower reliability than wired network, limited physical security, time varying channels, infrastructure, etc.

TABLE I. CHARACTERISTICS OF AD HOC NETWORKS

Autonomous and infrastructure less
Multi-hop routing
Dynamic network topology
Device heterogeneity
Energy constrained operation
Bandwidth constrained variable capacity links
Limited physical security
Network scalability
Self-creation, self-organization and self administration

C. Dual Channel MAC protocol

In [7] the authors propose DUCHA to overcome the limitations of the single channel MAC protocols. They utilize dual channel for DATA and control packets, separately. DATA is transmitted over the data channel. RTS and CTS are transmitted over the control channel. Negative CTS (NCTS) is used to solve the receiver blocking problem and is also transmitted on the control channel. In case of any collision in control channel it should differ long enough to avoid possible collision (at least SIFS+CTS + 2 × max-propagation-delay). An out of band receiver-based busy tone [5] [8] is used to solve the hidden terminal problem. ACK is unnecessary here because our protocol guarantees that there is no collision to DATA packets. To deal with wireless channel errors, the authors introduce a NACK signal which is a continuing busy tone signal when the receiver determines that the received DATA packet is corrupted. The sender will not misinterpret this NACK signal since there are no other receivers in its sensing range and hence no interfering NACK signals. It will conclude that the transmission is successful if no NACK signal is sensed.

In brief DUCHA solves a lot of problems of ad hoc networks and improves the spatial reuse of the data channel but did not consider the collision on the control channel. The control channel is still prone to collide and harm the data channel utilization.

III. THE PROPOSED MAC PROTOCOL

In this section, we present the new MAC protocol for ad hoc mobile networks.

A. Protocol Overview

In this protocol dual channel has been used for DATA and control packets individually. The data rate of the DATA channel will be higher than the data rate of the control, channel. DATA packets are transmitted through DATA channel and control packets (RTS/CTS) are transmitted through control channel. An out of band receiver-based busy tone is used to protect the DATA packets from collision and solve the hidden terminal problem. We use an extended busy tone as a NACK signal when the receiver determines that the received DATA packet is corrupted due to channel errors. The control channel time is counted in a slotted manner. Every node will monitor the channel and choose an idle slot for its control packet transmission. It will keep the slot for a while for its packet transmission.

Figure 1. Proposed protocol overview

B. Basic Operations

A node choose $\lfloor m/n \rfloor$ unique slots in the range from 0 to m – 1 in control channel for transmission, where m is the maximum number of nodes supported by the network, and n is the actual number of nodes in the network. The node transmits its control packets only when the current slot number corresponds to its slots. Once a node chooses a slot (assumed to be not occupied by other nodes) and successes to transmit its packet in the slot, the state of the slot is set to 1. For a successful transmission in the slot, the node increases the state of the slot by one. If a transmission attempt fails, it is decreased

by one. If the state becomes zero, the node releases the slot and finds another slot. Initially, all slots are set to state 0. A node is regarded to have a slot when the state of the slot is greater than one. However, if the value of the slot state is too high, slot management may not respond to topology changes. For this reason the value of slot state should not exceed 2.

- Initialization: At the beginning the nodes of the network monitors the link (control channel) and picks $\lfloor m/n \rfloor$ number of (assumed to be empty) candidate slots. Initially, the states of the slots are set to zero. Then, it attempts to transmit its packets in the candidate slots, and increases the state of the slot by one for a successful transmission. If it fails, it chooses another slot until it has $\lfloor m/n \rfloor$ number of slots.
- Detecting a New Node: If a new node joins the network and n increases, each node in the network keeps the $\lfloor m/n \rfloor$ largest state of slots (where n is the current number of nodes) and releases the rest of slots for the new node. The existing nodes do not need to inform which slots they release to the new node. The new node finds its slots via the Initialization Process in the above.
- Detecting Left of a Node: If n decreases, in order to maintain link utilization, each node in a network finds more slots to have $\lfloor m/n \rfloor$ number of slots via the *initialization* process

C. Basic Message Exchange

- RTS: For a new transmission of RTS, a node must send it to any of its own time slots. It can send the RTS to its different time slots to ensure successful transmission
- CTS/NCTS: Any node correctly receiving the RTS should return CTS on its own slots if the data channel is idle. If the data channel is busy it returns a NCTS. **I**t can provide the estimate for the remaining DATA transmission time in the duration field according to the difference between the transmission time of the maximum DATA packet and the length it has sensed a busy medium in the DATA channel.
- DATA: A transmitter, after correctly receiving the CTS, should start DATA transmission if no busy tone signal is detected. If the transmitter receives an NCTS, it defers its transmission according to the duration field of NCTS. Otherwise, it assumes that there is a collision, and starts the *initialization* process.
- Busy Tone: The intended receiver begins to sense the data channel after it transmits CTS. If the receiver does not receive signal in the data channel in the due time (for the first few bits of the DATA packet), it will assume that the sender does not transmit DATA. Otherwise, it transmits the busy tone signal to prevent hidden terminals from possible transmissions.
- NACK: The intended receiver has a timer to indicate when it should finish the reception of the DATA packet according to the duration field in the previously received RTS. If the timer expires and has not received the correct DATA packet, it assumes that the DATA transmission fails and sends NACK by continuing the busy tone signal for an appropriate period. If it correctly receives the DATA packet, it stops the busy tone signal and finishes the receiving procedure. The sender assumes that its DATA transmission is successful if there is no NACK signal sensed during the NACK period. Otherwise, it assumes that its transmission fails because of wireless channel error and then starts the retransmission procedure

In addition, during the NACK and DATA transmission period, any other nodes in the sensing range of the sender are not allowed to become the receiver of DATA packets, and any other nodes in the sensing range of the receiver are not allowed to become the sender of DATA packets. This is to avoid confusion between NACK signals and the normal busy tone signals.

D. Mathematical Model

Collision probability of control packets in control channel and its relation with the data channel performance was not studied in recent past. In this section, we describe the collision probability of control packets in control channel and compare the results with DUCHA.

In DUCHA, the control channel access is contention-based. It senses the control channel before it transmits any control packet. It will wait a certain time in case of busy control channel. It will discard the packet after a certain number of retransmission attempts [7].

The probability of collision in control channel in DUCHA can be derived as

$$P_{CD} = 1 - (1 - \frac{1}{W_{avg}})^{n-1}, \quad (1)$$

where P_{CD} is the collision probability of the control channel, W_{avg} is the average contention window size and n is the number of active nodes.

In our proposed protocol, the control channel is divided in to slotted manner. The probability of choosing the same slot by n nodes is

$$P_{Cn} = \frac{1}{m^n}. \quad (2)$$

So, the total probability of choosing the same slot by more than or equal to two nodes are

$$P_{CD} = \sum_{i=2}^{n} \frac{1}{m^i}, \quad (3)$$

where m is the total number of slots and n is the total number of nodes in the network.

Equation (3) represents the total probability of collision because there will be a collision if more than one nodes selects the same slots at the same time.

The maximum throughput of the control channel is then,

$$Throughput_{cmax} = \frac{(1-P_{CD})L_C}{\frac{L_C}{R_C}}, \qquad (4)$$

where L_C is the total bits successfully transmitted (in case of packet loss due to collision, it will be zero) and R_C is the control channel bandwidth.

IV. PERFORMANCE EVALUATION

To evaluate the performance of our proposed protocol and compare it with DUCHA protocol, we use MatLab. For the simplicity of the calculation, we assume that there is no link error and hence no packet is lost due to connection loss. We also consider a medium size network where the network will not allow more than 50 nodes. We analyze the protocol performance by varying the number of active nodes in the network. The parameters we use for the evaluation are given in Table II.

TABLE II. CHARACTERISTICS OF AD HOC NETWORKS

Data rate of control channel	1Mbps
Data rate of DATA channel	10Mbps
Control packet size	20 bytes
Data packet size	1000 bytes
Size of control channel slot	0.02 msec
Physical preamble	192 bits

Figure 2 shows the comparison of the control packet collision probability in control channel between our proposed protocol and the DUCHA. It has been observed that the collision probability is significantly low in our proposed protocol.

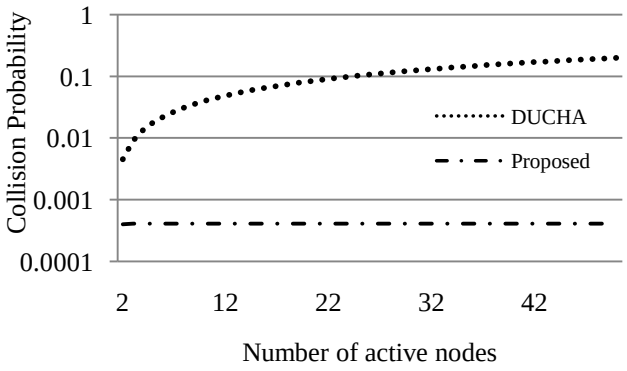


Figure 2. Comparison of control channel collision probability

The effect of less collision leads to higher throughput in data channel. Figure 3 describes the data channel throughput under a very congested traffic. It has been observed that in DUCHA the throughput decreases when the number of nodes increases. This is an effect of increased collision in control channel while our proposed protocol succeeded to keep a threshold point.

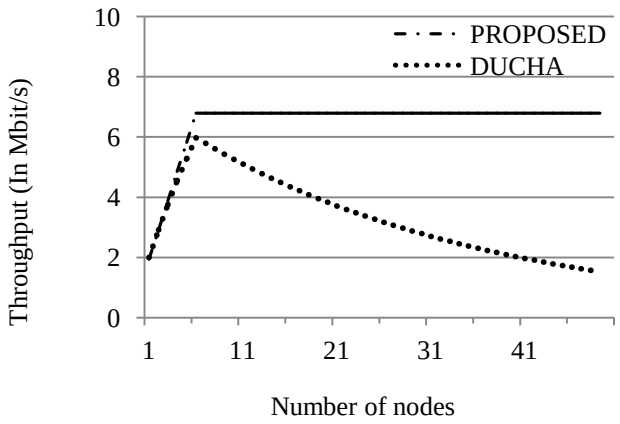


Figure 3. Throughput with high traffic load in each node

Figure 4 explains the performance degradation due to the mobility. In a mobile condition, any node can join and leave any time. That puts pressure to the time slot allocation or the initialization process. As the initialization process takes time, it is normal to experience performance degradation if there is frequent arrival of new nodes.

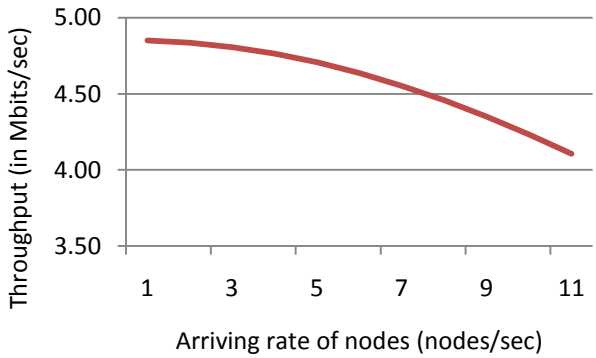


Figure 4. Network throughput analysis in a mobile environment

V. CONCLUSION AND FUTURE WORKS

In this paper we studied the inefficiency of existing MAC protocols in mobile ad hoc environment. The control packet collision probability in control channel and the effects of that on the overall throughput of the network are also observed in this paper. Finally we proposed a new MAC protocol consisting of two different channels for control and data packets. And a slotted approach is taken to manage control channel time. The proposed scheme decreases the control packet collision in control channel and hence improves the data channel throughput. In a future work, delay analysis and fairness could be studied.

REFERENCES

[1] IEEE Standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications, ISO/IEC 8802-11. I999(E)," Aug, 1999.

[2] H.lhai, X. Chen, and Y.Fang, "How well can the IEEE 802.11 wireless LAN support quality of service?" IEEE Trans. Wireless Commun, vol. 4, no. 6, pp.3084-3094, Nov. 2005.

[3] H.Zhai, Y. Kwon, and Y.Fang "Performance analysis of IEEE 802.11 MAC protocols in wireless LANs," *Wiley Wireless Communications and Mobile Computing*, vol. 4, pp. 917-931, Dec. 2004

[4] H. Zhai, J. Wang, X. Chen, and Y.Fang, "Medium access control in mobile ad hoc networks: Challenges and solutions," *Wiley Wireless Communications and Mobile Computing*, vol. 6, no. 2, pp. 151-170, Mar 2006.

[5] F. A. Tobagi and 1. Kleinrock, "Packet switching in radio channels Part II-The hidden terminal problem in carrier sense multiple-access and the busy-tone solution," *IEEE Trans. Commun.*, vol. 23, pp. 1417-1433, Dec.1975.

[6] C. L. Fullmer and J. J. Garcia-Luna-Aceves, "Solutions to hidden terminal problems in wireless networks," in Proc. ACM SIGCOMM, Sep. 1997, pp. 39-49.

[7] H. Zhai, 1. Wang, and Y. Fang, "DUCHA: A New Dual-Channel MAC Protocol for Multihop Ad Hoc Networks," *IEEE Transactions on Wireless Communications*, Vol. 5, no. 11, Nov2006.

[8] Z. 1. Haas and 1. Deng, "Dual busy tone multiple access (DBTMA)-A multiple access control for ad hoc networks," *IEEE Trans. Commun.*, vol. 50, pp. 975-985, June 2002.

[9] S. Singh and C. S. Raghavendra, "PAMAS-Power aware multi-access protocol with signalling for ad hoc networks," *Computer Commun. Review*, July 1998, pp. 5-26.

An Assistant System for Writing in Easy Japanese

Meng Zhang
Graduate School of Engineering
Tohoku University
Sendai, Japan
zhangm@spcom.ecei.tohoku.ac.jp

Akinori Ito
Graduate School of Engineering
Tohoku University
Sendai, Japan
aito@spcom.ecei.tohoku.ac.jp

Abstract— **As the foreigner population increasing for travel and short-term academic exchange in Japan, it is necessary to prompt "easy Japanese language" for them. Especially, they are encountered unexpected disasters, such as earthquake and tsunami, easy Japanese language become even important for the foreigners. Because the Japanese government, TV channels and radio stations broadcast the newest disasters information to them and they are avoid it. However, it is hard to identify the "easy Japanese language" by Japanese newscasters or reporters. In this paper, we demonstrate to what we proposed computer-assisted detection system identify the "Easy Japanese Language" by framework descriptions and experiments. The purpose of this research was to investigate the difficulty level of Japanese using the linguistic features and give introductions to the users in order to make a sentence with easy Japanese.**

Keywords-component; Easy Japanese; assistant system

I. INTRODUCTION

Japan attracts amounts of foreign people who cannot speak Japanese well, for either traveling or studying due to there has beautiful scenic spots and famous universities. Unfortunately, Japan often is encountered unexpected disasters, such as earthquake and tsunami. For both local Japanese people and mastered Japanese foreigners, they can understand Japanese emergence information from broadcast or newspaper either easy Japanese or difficult Japanese. Hence, they can avoid the unexpected disasters or deal with it well even in disaster environment. However, it is not easy to understand the urgent news for foreigners of beginner's level of Japanese because Japanese government, TV channels, broadcasts and volunteers do not know how to say "Easy Japanese" for these foreigners. Hence, it is necessary to identify "Easy Japanese" for Japanese learning beginner's foreigners. Using Easy Japanese for announcement of disaster will be a great benefit for both native Japanese speakers and non-native Japanese speakers [1].

Easy Japanese is similar concept to Basic English [2], but we do not intend create a new language; Easy Japanese is a kind of guideline for writing Japanese document that can be understood by foreigners of various language skill.

In this paper, we demonstrate a system for assisting Japanese native speakers in writing "Easy Japanese Language". The system, called YANSIS (YAsasii NIhongo SIen System), receives sentences written by a Japanese native speaker and points out the parts where foreigners might have difficulty on understanding those sentences.

II. EASY JAPANESE

"Easy Japanese" (*Yasasii Nihongo*) was proposed by Kazuyuki Sato in 1999, considering the fact that many foreigners suffered from language barrier after the Hanshin-Awaji Great Disaster happened in 1995 [1]. Many urgent announcements, as well as notices posted at shelters, were written in ordinary Japanese, and hence many foreigner did not understand them, which caused severe disadvantages for the foreigners. If those messages had been written in easier Japanese, it would be a great benefit for those people.

Easy Japanese (EJ) is basically a set of Japanese expressions that can be understood by those who passed the 3rd grade of the Japanese proficiency test. EJ is intended to convey various information needed within 72 hours after a disaster. Although most Japanese sentences cannot be translated into EJ, it is still useful because of the limited situation.

III. SYSTEM DESIGN

A. *Uses and objective of the system*

We made two assumptions on the user of YANSIS

1. Users of the system are employees of the government, staff men of broadcast stations or volunteers. These people are assumed to be native Japanese speakers, who are familiar with writing various documents.
2. The users of the system are not specialists of Japanese education of foreigners, nor have experience of special education for teaching Japanese.

Ultimately, it is ideal that the system translates normal Japanese (NJ) sentences into EJ automatically. However, considering the current natural language processing technologies, it is almost impossible to develop such system, because most of complicated sentences are impossible to translate into EJ in principle.

Considering that, we decided the objective of the system as helping the user making sentences in EJ. What YANSIS do is to point out the parts difficult for foreigners, and the correction of the sentences are the user's responsibility. Therefore, role of YANSIS is similar to a spell checker or a grammar checker.

It should be noticed that translating NJ sentences into EJ sentences is not a simple word substitution. In addition to use easy vocabulary and grammar, presentation of information in EJ should be clearer and more formal than that in NJ. Comprehension of sentences can be enhanced by employing simple arrangement of sentences. The current version of

YANSIS does not make any advice on the sentence-level structure.

B. *The framework of system*

Figure 1. Basic block diagram of YANSIS.

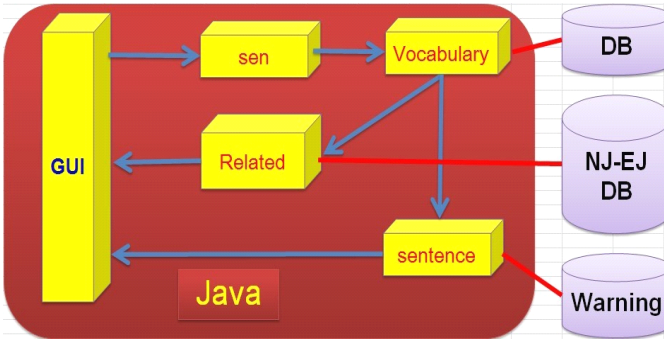


Figure 1 shows the basic framework of YANSIS. The user inputs a sentence using the GUI. Then the sentence is split into words using a morphological analyzer *Sen* [3]. Then the words are matched to the words in the database, and the difficulty of each word is determined. In addition, patterns of the word sequence is matched with the warning pattern database, and warning message is displayed if the sentence matches to one of the patterns of difficult expressions. Finally, the system can present EJ sentences related to a selected word.

IV. DEVELOPMENT OF THE SYSTEM

A. *Development environment*

The system is developed by Java, and the morphemic analyzer Sen 1.2.2.1 is used. As Sen is a library that is fully written in Java, YANSIS can be realized only using Java.

B. *Vocabulary database*

Table I shows the number of vocabulary registered in the database. The database consists of the function words, symbols and vocabulary, which has four levels in the Japanese Language Proficiency Test (JLPT), the level with smaller number is more difficult.

TABLE I. THE VOCABULARY

Vocabulary level	Number of distinct words
Level 1	3025
Level 2	3771
Level 3	718
Level 4	791
Function words	39
Symbol	133
Total	8477

C. *Warning pattern*

Sentences in EJ should have clear meaning, and ambiguity should be avoided. To avoid ambiguous expressions, we prepared pattern database that should be avoided. Table II shows examples of expressions in the database. We prepared 30 patterns in total. When the input sentence is matched to one of the patterns of the database, the system presents the corresponding warning.

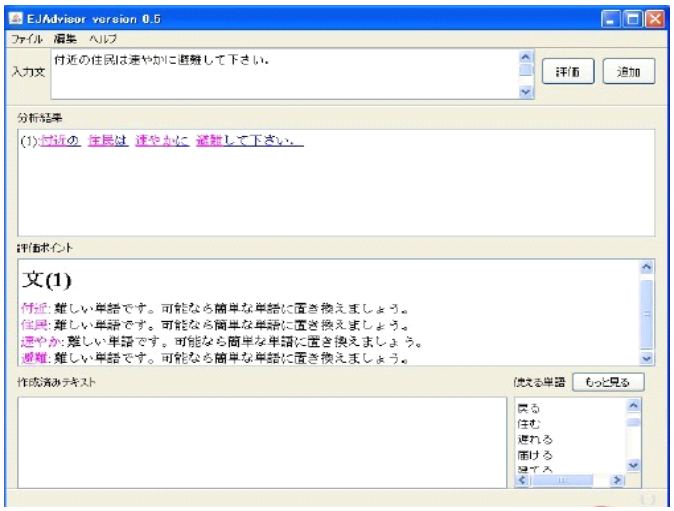
TABLE II. EXAMPLE OF THE WARNING DATABASE

pattern
mashou (Let's ...)
masenka (Why not...)
mashitaka (Did you ...)

D. *User Interface*

The GUI of the system is shown in figure 2. The GUI has one input and four output areas: the analysis area, the evaluation area, created sentence area and the example sentence area.

Figure 2. Graphical User Interface



The user of the system first inputs Japanese sentence in the input area and press the evaluation button. Then the system shows the analysis result in the analysis area, where difficult words are colored. The words in level 1 or 2 are marked as "difficult", while the words that does not appear in the vocabulary are marked as "more difficult." At the same time, description of the difficulties are listed in the evaluation area. Then the user modifies the difficult words in the input area, and repeats the modification and evaluation until the warning messages get fewer.

Besides, a morphological analysis result of a word shown in the analysis area will be displayed by clicking the word. The user can search the relevant EJ sentences to that word using the morphemic analysis dialog.

V. CONCLUSIONS

In this paper, overview of the writing support system of EJ sentences is described. The system points out the difficult words and expressions in the user's input sentence to help the user writing easy Japanese sentences.

YANSIS is now actually used many workshops of EJ writing, and hundreds of officers, volunteers and teachers have been learned EJ writing using YANSIS.

REFERENCES

[1] K. Sato, “Easy expressions save foreigners!”, http://human.cc.hirosaki-u.ac.jp/kokugo/tagengoenglish.html

[2] C. K. Ogden, “Basic English as an international second language,” Harcourt, Brace & World , 1968.

[3] Morphemic analyzer Sen. http://java.net/projects/sen/

Stability Analysis of Discrete-Time Large-Scale Systems with Time-Delays

Qiong Huang
School of Electronic Engineering,
University of Electronic Science and Technology of China,
Chengdu, 611731, P.R.China.
Email: subaru6725@163.com

Long Sheng and Guanjie Yu
School of Electronic Engineering,
University of Electronic Science and Technology of China,
Chengdu, 611731, P.R.China.

***Abstract*—This paper presents the issue of stability analysis of discrete-time fuzzy large-scale systems with time-varying delays based on a set of Takagi-Sugeno(T-S) fuzzy model. The fuzzy large-scale systems consists of N time-delay T-S fuzzy subsystems. Via the delay-dependent piecewise Lyapunov-Krqasovskii functional(DDPLKF), we analyze the stability and H_∞ performance of the prescribed systems. Through DDPLKF, system description , solving a set of linear matrix inequalities (LMIs) and with an improved free weighting matrix approach, we obtain the qualification of the overall fuzzy system when we use MATLAB to work out the LMIs. Consequently, a numerical example is given to demonstrate the feasibilities and efficiencies of the proposed methods. The approach illustrated in this paper has great practical value in the stability judgments of the overall fuzzy systems.**

Keywords: Delay-dependent , time-varying delays, H_∞ performance, stability, nonlinear interconnected system, Takagi-Sugeno(T-S) fuzzy model.

I. Introduction

With the development of science, many practical systems such as the power systems, nuclear reaction systems, space vehicle systems, economic systems, and many other mathematical models of such systems are involved in the increasing size and complexity. The methodologies of large-scale system provide a technique through the manipulation of system structure in some way to overcome this problem. There were considerable interests in the research of large-scale systems in past years [1].

In 1985, Takagi and Sugeno proposed a well-known T-S fuzzy modeling method [2]. The dynamics of the nonlinear plant could be systematically represented by the T-S fuzzy model. Based on the large-scale systems theory, the stability analysis and control synthesis of the complex systems can be solved.

Stability is one of the most important indicator of dynamical systems. Time-delay is widespread in many practical engineering systems, such as mechanical drive, temperature control systems, neural networks and financial management. In recent years, the stability of the fuzzy control systems with time-delay has drawn a deluge of research attention. To investigate this class of nonlinear systems, the Lyapunov-Krasovskii function [3,4] was employed. In general, there are two categories of stability analysis found in literature: delay-dependent and delay-independent. For the delay-dependent stability analysis approach, the time-delay is considered during the stability analysis.

In modeling the actual systems, due to changes in the systems of the external environment and a variety of unpredictable interference, there are certain gaps between the mathematical model and the actual dynamics system. Therefore, we should consider robustness of the system. In this article, we address the issue of stability analysis and the H_∞ performance of discrete-time large-scale systems via DDPLKF approach[5].

The rest of this paper is organized as follows. System description and preliminaries are presented in Section 2. Stability analysis is presented in Section 3. Piecewise H_∞ performance analysis is developed in Section 4. In Section 5, numerical example is provided to demonstrate the effectiveness and feasibility of the method.

Notations: The superscript"T" stands for matrix transposition, I is an identity matrix, $\Re^n$ denotes the n-dimensional Euclidean space, the notation $P > 0 (P \geq 0)$ means that P is symmetric and positive (nonnegative) definite, $diag\{\cdots\}$ stands for a block diagonal matrix, $l_2[0,\infty)$ denotes the space of square-integrable vector functions over $[0,\infty)$, $\|(\cdot)\|_2$ refers to the standard l_2 norm, i.e., $\|(\cdot)\|_2 = (\sum_{k=0}^{\infty} (\cdot)_k^T (\cdot)_k)^{\frac{1}{2}}$. In addition, matrices, if not explicitly stated, are assumed to have compatible dimensions.

II. System Descriptions And Preliminaries

Consider the following discrete-time T-S fuzzy model with time-varying delays which could be used to represent a class of complex nonlinear delay systems. This fuzzy dynamic model, composed of N interconnected subsystems S_i, $i = 1, 2, \ldots, N$, is described by fuzzy IF-THEN rules as follows[6]:

Rule j: IF $\theta_{i1}(k)$ is M_{i1j} and $\ldots$ and $\theta_{ig}(k)$ is M_{igj} THEN

$$\begin{cases} x_i(k+1) = A_{ij}x_i(k) + A_{dij}x_i(k-\tau_i(k)) \\ \qquad\qquad +D_{ij}v_i(k) + \sum\limits_{\substack{n=1 \\ n\neq i}}^{N} C_{in}x_n(k) \\ y_i(k) = H_{ij}x_i(k) + H_{dij}x_i(k-\tau_i(k)) \\ x_i(k) = \phi_i(k),\ k = 0, -1, ..., -\tau_2, \end{cases} \tag{1}$$

where $j \in L_i = \{1, 2, \ldots, r_i\}$ denotes the jth fuzzy inference rule; $\theta_i(k) = (\theta_{i1}(k), \theta_{i2}(k), \ldots, \theta_{ig}(k)) \in R^{n_{ig}}$ are some

measurable premise variables; and $M_{ipj}(p = 1, 2, \ldots, g)$ are the fuzzy sets; $x_i(k) \in R^{n_{ix}}$, $y_i(k) \in R^{n_{iy}}$ are the state vector and the output vector, $u_i(k) \in R^{p_i}$ is the control signal, $v_i(k) \in R^{q_i}$ is the disturbance which belongs to $l_2[0, \infty)$; (A_{ij}, A_{dij}, B_{ij}, D_{ij}, H_{ij}, H_{dij}, G_{ij}) is the jth local model of the ith fuzzy subsystem, and C_{in} is the interconnection between the ith and nth subsystems; r_i is the number of fuzzy implications, $\phi_i(k)$ the given initial condition sequence. Time-varying delays $\tau_i(k)$ are positive integers satisfying the following assumption:

$$\tau_1 \leq \tau_i(k) \leq \tau_2, \tag{2}$$

where τ_1 and τ_2 are two constant positive integers representing the minimum and maximum time-delay.

Through using "fuzzy blending", the final output of the ith fuzzy subsystem could be inferred as follows:

$$\begin{aligned} x_i(k+1) = & \sum_{j=1}^{r_i} h_{ij}(k)[A_{ij}x_i(k) + A_{dij}x_i(k - \tau_i(k)) \\ & + D_{ij}v_i(k) + \sum_{\substack{n=1 \\ n \neq i}}^{N} C_{in}x_n(k)] \\ y_i(k) = & \sum_{j=1}^{r_i} h_{ij}(k)[H_{ij}x_i(k) + A_{dij}x_i(k - \tau_i(k))] \end{aligned} \tag{3}$$

with

$$\omega_{ij}(\theta_i(k)) = \prod_{p=1}^{g} M_{jpi}(\theta_{ip}(k)),\ h_{ij}(\theta_i(k)) = \frac{\omega_{ij}(\theta_i(k))}{\sum_{j=1}^{r_i} \omega_{ij}(\theta_i(k))}, \tag{4}$$

in which $M_{jpi}(\theta_{ip}(k))$ is the grade of membership of $\theta_{ip}(k)$ in M_{ipj}. It is assumed that $w_{ij}(k) \geq 0$ for all k, $j = 1, 2, \cdots, r_i$. Therefore, the normalized membership function $h_{ij}(\theta_i(k))$ satisfies $h_{ij}(\theta_i(k)) \geq 0$, $\sum_{j=1}^{r_i} h_{ij}(\theta_i(k)) = 1$ for all k.

For the purpose of develop less conservative analysis and synthesis, by using DDPLKF candidate, in subsystem S_i, we partition the premise variable space $\Omega_i \subseteq R^{n_{ig}}$ by the boundaries[5]:

$$\begin{aligned} & \partial\Omega_{ij}^{v} = \{\theta_i(k) | h_{ij}(\theta_i(k)) = 1, \\ & 0 \leq \underset{0<|\epsilon|\ll 1}{h_{ij}} (\theta_i(k) + \epsilon) < 1, j \in L_i\}, \end{aligned} \tag{5}$$

where v is the set of the face indexes of the polyhedral hull under the circumstance of $\partial\Omega_{ij} = \cup_v \partial\Omega_{ij}^{v},\ j \in L_i$. Based on the boundaries, the variable space Ω_i can be partitioned into $\hat{m}_i$ independent polyhedral hull $\Omega_{ip},\ p \in \aleph_i = 1, 2, 3, \ldots, \hat{m}_i$ which satisfies

$$\Omega_{ip} \cap \Omega_{iq} = \partial\Omega_{ij}^{v},\ p \neq q, p, q \in \aleph_i, j \in L_i,$$

where $\aleph_i$ denotes the set of subspace indexes.

In each subspace Ω_{ip}, we define the set

$$\Lambda_i(p) := \{\xi\, | h_{i\xi}(\theta_i(k)) > 0,\ \xi \in L_i, \theta_i(k) \in \Omega_{ip},\ p \in \aleph_i\}.$$

The $h_{i\xi}(k)$ symbolizes the membership of the fuzzy sets in each region Ω_{ip} and satisfies

$$\sum_{\xi \in \Lambda_i(p)} h_{i\xi}(\theta_i(k)) = 1.$$

Then the subsystem (3) can be inferred as follows:

$$S_i : \begin{cases} x_i(k+1) = \sum\limits_{\xi \in \Lambda_i(p)} h_{i\xi}(A_{i\xi}x_i(k) + A_{di\xi}x_i(k - \tau_i(k)) \\ \qquad + D_{i\xi}v_i(k) + \sum\limits_{\substack{n=1 \\ n \neq i}}^{N} C_{in}x_n(k)) \\ y_i(k) = \sum\limits_{\xi \in \Lambda_i(p)} h_{i\xi}(H_{i\xi}x_i(k) + H_{di\xi}x_i(k - \tau_i(k)) \\ \theta_i(k) \in \Omega_{ip}, \quad p \in \aleph_i. \end{cases} \tag{6}$$

III. Stability Analysis

In this section, we consider the stability analysis of the large-scale fuzzy systems (6) illustrated in the last section without external disturbance, that is, $v_i(k) \equiv 0, i = 1, 2, \cdots, N$. Omitting the output term, the ith subsystem S_i becomes the following form:

$$\begin{aligned} S_i : x_i(k+1) = & \sum_{\xi \in \Lambda_i(p)} h_{i\xi}(A_{i\xi}x_i(k) + A_{di\xi}x_i(k - \tau_i(k)) \\ & + \sum_{\substack{n=1 \\ n \neq i}}^{N} C_{in}x_n(k)) \end{aligned} \tag{7}$$

The delay-dependent stability condition for the large-scale system without external disturbance can be summarized in the following theorem.

Theorem 1 (Stability Analysis): The fuzzy large-scale system composed of N fuzzy subsystems as (6) with $v_i = 0$, $i = 1, 2, \cdots, N$, is asymptotically stable if there exist symmetric positive definite matrices $P_{ip} > 0$, $Q_i > 0$, $U_i \geq 0$ and X_{ip}, M_{ip}, N_{ip} with appropriate dimension, symmetric matrix R_i and positive constant ε, satisfying

$$\begin{bmatrix} \Pi & \Psi_i \\ * & -\varepsilon * I \end{bmatrix} < 0 \tag{8}$$

$$\begin{bmatrix} X_{ip} & M_{ip} \\ * & U_i \end{bmatrix} \geq 0 \tag{9}$$

$$\begin{bmatrix} X_{ip} & N_{ip} \\ * & U_i \end{bmatrix} \geq 0 \tag{10}$$

for all $i = 1, 2, \ldots, N,\ p \in \aleph_i,\ \xi \in \Lambda_i(p)$, where

$$\Pi = \Pi_1^{'} + \Pi_2 + \Pi_3 + \Pi_3^{T} + \tau_2 X_{ip}$$

$$\Pi_1^{'} = \begin{bmatrix} P_{iq} - P_{iq} + (\tau_2 - \tau_1 + 2)Q_i \\ \quad + (N-1)\varepsilon \sum\limits_{\substack{n=1 \\ n \neq i}}^{N} C_{ni}^{T}C_{ni} & & \\ * & & \\ * & & \\ * & & \end{bmatrix}$$

$$\begin{bmatrix} P_{iq} & 0 & 0 \\ \tau_2 U_i + P_{iq} & 0 & 0 \\ * & -Q_i & 0 \\ * & * & -Q_i \end{bmatrix}$$

$$\Pi_2 = \Psi_i \begin{bmatrix} (A_{i\xi} - I) & -I & A_{di\xi} & 0 \end{bmatrix} + \begin{bmatrix} (A_{i\xi} - I) & -I & A_{di\xi} & 0 \end{bmatrix}^{T} \Psi_i^{T}$$

$$\Pi_3 = \left[\begin{array}{cccc} M_{ip} & 0 & (N_{ip} - M_{ip}) & -N_{ip} \end{array} \right]$$

$$\Psi_i = \left[\begin{array}{cccc} R_i^T & R_i^T & 0 & 0 \end{array} \right]^T$$

$$N_{ip} = \left[\begin{array}{c} N_{i1p} \\ N_{i2p} \\ N_{i3p} \\ N_{i4p} \end{array} \right], \; M_{ip} = \left[\begin{array}{c} M_{i1p} \\ M_{i2p} \\ M_{i3p} \\ M_{i4p} \end{array} \right]$$

$$X_{ip} = \left[\begin{array}{cccc} X_{i11p} & X_{i12p} & X_{i13p} & X_{i14p} \\ & X_{i22p} & X_{i23p} & X_{i24p} \\ & * & X_{i33p} & X_{i34p} \\ & * & * & X_{i44p} \end{array} \right].$$

Proof : The proof is omitted here.

IV. Piecewise H_∞ Performance Analysis

In this section, we will analyze the H_∞ disturbance attenuation performance for the open-loop delay-dependent discrete-time fuzzy large-scale system. With external disturbance, the ith subsystem S_i can be expressed as (6). Now, we give the definition of H_∞ disturbance attenuation performance of a discrete-time fuzzy large-scale system as follows.

Definition 1: Given a constant $\gamma > 0$, the open-loop discrete-time fuzzy large-scale system composed of N subsystems as (6) is said to be stable with γ-disturbance attenuation if it is asymptotically stable and the output satisfies

$$\|y\|_2 < \gamma \|v\|_2 \tag{11}$$

for all nonzero $v(k) \in l_2[0, \infty)$ under the zero initial condition. Here, $y(k) = [y_1^T(k), y_2^T(k), \ldots, y_N^T(k)]^T$, $v(k) = [v_1^T(k), v_2^T(k), \ldots, v_N^T(k)]^T$. Then, we are ready to present the H_∞ performance analysis of (6) in terms of LMIs as follows.

Theorem 2 (H_∞ Performance Analysis): The fuzzy large-scale system composed of N fuzzy subsystems as (6) is stable if there exist symmetric positive definite matrices $P_{ip} > 0$, $Q_i > 0$, $U_i^{'} > 0$, and $X_{ip}^{'}$, $M_{ip}^{'}$, $N_{ip}^{'}$, R_i with appropriate dimensions and positive, and positive constant ε, satisfying

$$\left[\begin{array}{ccc} \Phi & \Psi_i^{'} & \mathrm{H}_i^T \\ * & -\varepsilon & 0 \\ * & * & -I \end{array} \right] < 0 \tag{12}$$

$$\left[\begin{array}{cc} X_{ip}^{'} & N_{ip}^{'} \\ * & U_i^{'} \end{array} \right] \geq 0 \tag{13}$$

$$\left[\begin{array}{cc} X_{ip}^{'} & M_{ip}^{'} \\ * & U_i^{'} \end{array} \right] \geq 0 \tag{14}$$

for $i = 1, 2, \ldots, N$, $p \in \aleph_i$, where

$$\Phi = \Phi_1 + \Phi_2 + \Phi_2^T + \tau_2 X_{ip}^{'}$$

$$\Phi_1 = \left[\begin{array}{cccc} \Phi_{11} & \Phi_{12} & R_i A_{di\xi} & 0 & R_i D_{i\xi} \\ & \Phi_{22} & R_i A_{di\xi} & 0 & R_i D_{i\xi} \\ & * & -Q_i & 0 & 0 \\ & * & * & -Q_i & 0 \\ & * & * & * & -\gamma^2 \end{array} \right]$$

$$\Phi_{11} = P_{iq} - P_{ip} + (\tau_2 - \tau_1 + 2)Q_i + R_i A_{i\xi} + A_{i\xi}^T R_i^T - R_i - R_i^T + (N-1)\varepsilon \sum_{\substack{n=1 \\ n \neq i}}^{N} C_{ni}^T C_{ni}$$

$$\Phi_{12} = P_{iq} - R_i - R_i^T + A_{i\xi}^T R_i^T$$

$$\Phi_{22} = P_{iq} + \tau_2 U_i^{'} - R_i - R_i^T$$

$$\Phi_2 = \left[\begin{array}{ccccc} M_{ip}^{'} & 0 & (N_{ip}^{'} - M_{ip}^{'}) & -N_{ip}^{'} & 0 \end{array} \right]$$

$$\Psi_i^{'} = \left[\begin{array}{ccccc} R_i^T & R_i^T & 0 & 0 & 0 \end{array} \right]^T$$

$$H_i = \left[\begin{array}{ccccc} H_{ij} & 0 & H_{dij} & 0 & 0 \end{array} \right]$$

$$N_{ip}^{'} = \left[\begin{array}{c} N_{i1p} \\ N_{i2p} \\ N_{i3p} \\ N_{i4p} \\ N_{i5p} \end{array} \right], \; M_{ip}^{'} = \left[\begin{array}{c} M_{i1p} \\ M_{i2p} \\ M_{i3p} \\ M_{i4p} \\ M_{i5p} \end{array} \right]$$

$$X_{ip}^{'} = \left[\begin{array}{ccccc} X_{i11p} & X_{i12p} & X_{i13p} & X_{i14p} & X_{i15p} \\ & X_{i22p} & X_{i23p} & X_{i24p} & X_{i25p} \\ & * & X_{i33p} & X_{i34p} & X_{i35p} \\ & * & * & X_{i44p} & X_{i45p} \\ & * & * & * & X_{i55p} \end{array} \right].$$

Proof : The proof is omitted here.

V. Numerical Example

Example 1: In this example, we consider a fuzzy interconnected system which composed of two subsystems S_i represent by the T-S fuzzy model with time-varying delay in the form of (6) without input, output, and disturbance.

Subsystem S_i:

Rule 1) IF $x_{i1}(k)$ is M_{i1l}, THEN

$$x_i(k+1) = A_{i1}x_i(k) + A_{di1}x_{di}(k - \tau_i(k)) + \sum_{\substack{n=1 \\ n \neq i}}^{2} C_{ni1}x_n(k)$$

Rule 2) IF $x_{i1}(k)$ is M_{i2l}, THEN

$$x_i(k+1) = A_{i2}x_i(k) + A_{di2}x_{di}(k - \tau_i(k)) + \sum_{\substack{n=1 \\ n \neq i}}^{2} C_{ni2}x_n(k)$$

with $i = 1, 2$. Where $x_i(k) = [x_{i1}(k), x_{i2}(k)]^T$

$$A_{11} = \left[\begin{array}{cc} 0.7 & 0 \\ 0 & 0.5 \end{array} \right], \; A_{12} = \left[\begin{array}{cc} 0.6 & 0 \\ 0 & 0.3 \end{array} \right]$$

$$A_{21} = \left[\begin{array}{cc} 0.65 & 0 \\ 0 & 0.45 \end{array} \right], \; A_{22} = \left[\begin{array}{cc} 0.55 & 0 \\ 0 & 0.25 \end{array} \right]$$

$$A_{d11} = \left[\begin{array}{cc} 0.1 & 0 \\ 0 & 0.05 \end{array} \right], \; A_{d12} = \left[\begin{array}{cc} 0.1 & 0 \\ 0 & 0.05 \end{array} \right]$$

$$A_{d21} = \left[\begin{array}{cc} 0.08 & 0 \\ 0 & 0.02 \end{array} \right], \; A_{d22} = \left[\begin{array}{cc} 0.08 & 0 \\ 0 & 0.02 \end{array} \right]$$

$$C_{121} = \left[\begin{array}{cc} 0 & 0.01 \\ 0 & 0.01 \end{array} \right], \; C_{122} = \left[\begin{array}{cc} 0 & 0.01 \\ 0 & 0.01 \end{array} \right]$$

$$C_{211} = \left[\begin{array}{cc} 0 & 0.01 \\ 0 & 0.01 \end{array} \right], \; C_{212} = \left[\begin{array}{cc} 0 & 0.01 \\ 0 & 0.01 \end{array} \right].$$

The normalized membership functions and cell partitions of

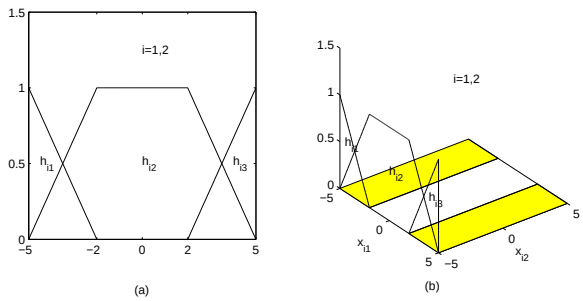


Fig. 1. (a)Normalized membership functions. (b)Cell partitions of the state space.

each subsystem are the same and shown in Fig.1.
Based on the proposed approach, there are three cell partitions for each subsystem.
Subsystem S_i: $x_{i1} = [-5,-2),[-2,2),[2,5)$,
which are denoted by Ω_{i1}, Ω_{i2}, Ω_{i3}, respectively, where $i = 1,2$. We supposed that the $\tau_i(k)$ be a constant, and $\tau_i = 1$, for $i = 1,2,\ldots,N$. It is noted that there is no solution for the stability analysis approach based on the delay-dependent method indicated in [7]. To successfully curb such unfavorable blights, via applying theorem 1, a piecewise quadratic Lyapunov functions $V(k)$ can be found so that the asymptotic stability of the origin delay-dependent fuzzy large-scale system is certified. Furthermore, we can found the following solution to those LMIs when $\varepsilon = 1$:

$$P_{11} = \begin{bmatrix} 0.0024 & -0.0000 \\ -0.0000 & 0.4808 \end{bmatrix},$$

$$P_{12} = \begin{bmatrix} 0.0239 & -0.0004 \\ -0.0004 & 0.4230 \end{bmatrix},$$

$$P_{21} = \begin{bmatrix} 0.1848 & 0 \\ 0 & 0.5571 \end{bmatrix},$$

$$P_{22} = \begin{bmatrix} 0.1772 & 0 \\ 0 & 0.5112 \end{bmatrix},$$

$$Q_1 = \begin{bmatrix} 0.0004 & -0.0000 \\ -0.0000 & 0.0106 \end{bmatrix}, \; Q_2 = \begin{bmatrix} 0.0035 & 0 \\ 0 & 0.0146 \end{bmatrix}$$

$$U_1 = \begin{bmatrix} 0.0015 & -0.0003 \\ -0.0003 & 0.0103 \end{bmatrix}, \; U_2 = \begin{bmatrix} 0.0087 & 0 \\ 0 & 0.0130 \end{bmatrix}.$$

VI. Conclusions

In this paper, the stability analysis based on DDPLKF approach is developed for discrete-time nonlinear interconnected systems with time-varying delays. The H_∞ performance is presented as well. It is shown that the stability analysis can be determined by solving a set of LMIs. A numerical example is also presented to demonstrate the effectiveness of the proposed approaches.

References

[1] D. D. Siljak, "Large-Scale Dynamic Systems: Stability and Structure", Elsevier North-Hollad, New York, 1978.

[2] T. Takagi and M. Sugeno, "Fuzzy identification of systtems and its application to modeling and control," IEEE Trans. Syst. Man. Cybern., vol. 15, pp. 116-132, 1985.

[3] J. K. Hale, S. M. Lunel. "Introduction of functional equations", Springer-VerlagNew York, 1993.

[4] V. B. Kolmanovskii, V. R. Nosov. "Stability of functional differential equations", New York: Academic Press, 1986.

[5] C. L. Chen, G. Feng, X. P Guan. "Delay-dependent stability analysis and controller synthesis for discrete-time T-S fuzzy systems with time delays", IEEE Trans. Fuzzy Syst., 2005,vol.13,no.5,pp630-643.

[6] H. B. Zhang, G. Feng. "Stability analysis and H_∞ controller design of discrete-time fuzzy large-scale systems based on piecewise Lyapunov functions", IEEE Trans. Syst., Man, Cybern., 2008,vol.38,no.5,pp.1390-1401.

[7] M. Chen, G. feng, H. B. Ma, G. Chen. "Delay-dependent H_∞ filter design for discrete-time fuzzy systems with time-varying delays", IEEE Trans. Fuzzy Syst., 2009, vol.17,no.3,pp.604-616.

[8] X. W. Liu and H. B. Zhang, "Delay-dependent robust stability of uncertain fuzzy large-scale systems with time-varying delays," Automatica, 2008, vol.44, no.1, pp.193-198.

[9] H. B. Zhang, C. G. Li and X. F. Liao, "Stability analysis and H_∞ controller design of fuzzy large-scale systems based on piecewise Lyapunov functions," IEEE Trans. Syst., Man, Cybern. B, Cybern., Jun. 2006 vol. 36, no. 3, pp. 685-698.

Design and Implementation of an Emergency Response System using GPS

Junyeol Song
Dept. of Info. and Comm. Eng.,
Yeungnam University
Gyeongsan, South Korea
jrljjang@gmail.com

Yongwan Park
Dept. of Info. and Comm. Eng.,
Yeungnam University
Gyeongsan, South Korea
ywpark@yu.ac.kr

Abstract— **In this paper, the weak persons who felt a risk of a person can inform own position by operation of simple appliances and in lost child, to parents and give them their position in less time can be found by tracing the location of Mia Wi-Fi communications experience in using GPS in an emergency response system was designed and implemented.**

Keywords-component; Embedded System; GPS; Wi-Fi; LBS; SMS;

I. INTRODUCTION

Recently Rising High Oil Prices as two-income family with the people who use public transportation is gaining. This targeting children alone or in a crowded during rush women who use public transport to target an increase in crime and the kidnapping and sex crimes to that kind and there are those types. Because of this crime, learn about ways to cope but actually it is difficult to deal with it when the situation.

The development of wireless Internet technology and mobile terminal device, with the rapid development of technology for mobile computing, one after another is being developed by a variety of applications. In addition, GPS-based location technology with satellite navigation technology at a rapid pace, and they are spread a lot of applications has been put to practical use.

In this paper, according to a recent social phenomenon, which developed rapidly in the mobile computing, wireless infrastructure, using the information infrastructure, without regard to time or place to move freely and to enjoy Wi-Fi service in an environment where children and women with a GPS receiver for a simple device that is attached through an accessory for a rainy day on the server wirelessly transmit information to the user's current location and the management server received a user's location information is registered with the police due to parent's mobile users by giving notice of the to detect anomalies, such as kidnapping or sexual offenses more about your life that prevents the system was designed and implemented.

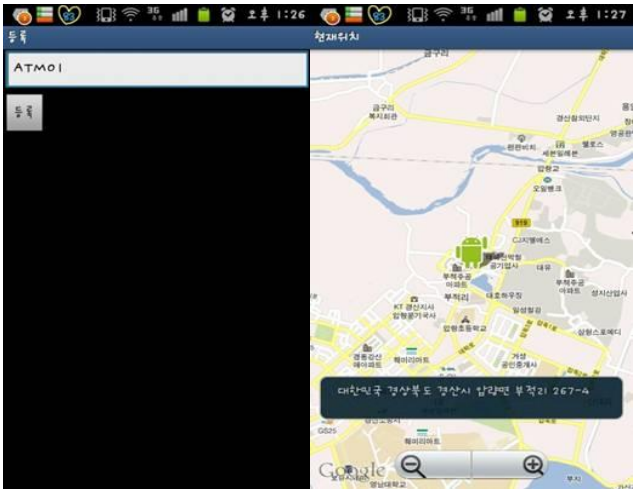
II. SYSTEM MODEL

Processing capabilities of embedded devices, desktop computers, compared to the fall, because it uses batteries for mobile devices, many limitations of the comes to power. Also as transmit and receive by radio data, reliability regarding data drops. I use client - server structure in order to to raise efficiency regarding a module considering this non-efficiency. In other words, It is composed with a form I access it if I put one management server, and a client operates, and to use, and I use Wi-Fi between the desk tower computer and embedded devices which are management server for data delivery, and I communicate through wireless networks.

GPS-based systems for emergency response system in the management server using the JDBC and C2DM and was composed of MS SQL database in this proposed paper. I planned it in the device which a GPS receiver was attached to accessories in order to do naturally it as it was comfortable for a user environmental a client carrying it.But I used actual implementation and the PCB which I composed to GPS receipt modules etc. for convenience of simulation experiment, and I did it so that server can easily do access through Wi-Fi. And if a user ran device, I composed a system to a mobile of Android OS so that thing used C2DM to push to a mobile of a guardian to SMS of position information of a user.

III. THE DESIGN OF GPS BASED ON EMERGENCY RESPONSE SYSTEM

Main component of management server is composed of user management, a socket, database, agent. User management component manages an account regarding mobile information of device of a user and a guardian. And name of each user and information regarding device are stored to actual database, but management does it here. If there is a request of a user, socket component takes charge of the communication section which sends by radio communication in position information to guardians. Database component stores device information of device and a guardian, and managing. Finally, Agent component is a piece managing other component. If there is a request of a user, it alert to guardians.



(a) **(b)**

Figure 1. Implementation results. (a) a screen that register on the server in guardian's mobile device. (b) location information of devices is sent to registered Guardian's mobile

A client of a user is composed to GPS module, ATmega128, Wi-Fi module, a buzzer, a switch. So a user has send own information of position through Wi-Fi to server by simple operation like pressing a switch only. Also the buzzer is at you can inform around own risk.

IV. EXPERIMENTAL RESULTS

In this section the proposed GPS-based design and implement an emergency response system displays the results. Fig. 1(a) is implemented in this paper, GPS-based emergency response system is the guardian's mobile screen as C2DM information and to register the user device. It is user's location through C2DM transmitted to guardian's mobile to Fig. 1.(b) .

V. CONCLUSIONS

Recently, we mainly around children and crimes that target the same weak persons is a increased trend. Also the rapid development of wireless technologies and mobile devices, with a rapidly developing technology and applications are being deployed.

In this paper, according to a recent social phenomenon, this developed rapidly uses radio-based information infrastructure. GPS receiver attached via a simple device by radio transmits the current position. The management server is received information about the location of the user's device and transfer it to the safety of users indicates anomalies in the system designed and implemented.

REFERENCES

[1] Anthony Jones, Jim Ohlund, "Network Programming For Microsoft Windows", Microsoft Press, 2003.

[2] Charles Petzold, Programming Windows, 5th ed, Microsoft Press, 2005, pp.1536

[3] Kaplan, Elliott D, "GPS: Principles and Applications", Bostion: Artech House Publishers, 1996

[4] Asharf A. Tahat, "Mobile Messaging Services-Based Personal Electrocardiogram Monitoring System", in Int. Journal of Telemedicine and Applications, Vol. 2009.

[5] Cleonilson Protasio de Souza, Tiago Pontes Pereira, Raimundo C. Silverio Freire, "Electrocardiogram by Mobile Phone: A Compression Method for SMS," In Proc. of XIX IMEKO World Congress: Fundamental and Applied Metrology, pp.1707-1710, Sep, 2009, Lisbon, Portugal.

[6] G. Bodic, Mobile Messaging Technologies and Services, John Wiley & Sons, West Sussex, UK, 2005.

Narrow spectral linewidth wavelength tunable laser diodes consisting of SOI ring resonators

Keita NEMOTO[1], Tomohiro KITA[2] and Hirohito YAMADA[2]
[1]Department of Information and Intelligent Systems, Faculty of Engineering, Tohoku University
[2]Department of Electrical & Communication Engineering, Graduate School of Engineering, Tohoku University
6-6-05 Aramaki-Aza-Aoba Aoba-ku Sendai 980-8579, Japan
{k.nemoto, tkita, yamada}@ecei.tohoku.ac.jp

Abstract—**We fabricated wavelength tunable laser diodes with Si-wire waveguide ring resonators as an external optical cavity. Less than 100 kHz narrow spectral linewidth was obtained by optimizing design of the ring resonators. The wavelength tunable laser diodes with narrow spectral linewidth are suitable as light sources for digital coherent optical communication systems.**

Keywords-component; Si-wire waveguide; wavelengh tunable laser diode; ring resonator; lasing spectral linewidth

I. INTRODUCTION

Compact and low-power consumption wavelength tunable laser diodes are required for use in high capacity optical communication systems. Recent coherent digital optical transmission systems require light sources with extremely low phase noise. It means less than 100 kHz narrow spectral linewidth is required for higher-order multilevel modulation schemes such as 16 quadrature amplitude modulation (QAM) [1]. Wavelength tunable lasers constructed with ring resonator wavelength filters and a semiconductor optical amplifier (SOA) have high mechanical stability, durability, and higher productivity. Less than 100 kHz narrow spectral linewidth has been already demonstrated by tunable laser diodes consisting of triple ring resonators with silica (SiON) waveguides [2]. However, no one has reported narrow spectral linewidth for wavelength tunable laser diodes consisted with Si-wire waveguide ring resonators. Si-wire waveguides that can make sharp bend and having high thermo-optic (TO) effect are attractive for constructing ring resonators for tunable lasers because it can drastically reduce size of tunable lasers and also reduce power consumption by tuning operation [3]-[5].

In this paper, we describe narrow spectral linewidth operation for the tunable lasers with Si-wire waveguide ring resonators.

II. STRUCTURE AND OPERATION OF THE TUNABLE LASER

Schematic structure of the wavelength tunable laser is shown in Fig. 1. The optical cavity is composed by two ring resonators with different free-spectral ranges (FSRs) made of Si-wire waveguides, bus waveguides that is optically coupled with the ring resonators, and a loop mirror with a Y-junction. The waveguides have a core with cross-sectional size of 400 nm × 220 nm, and the FSRs of the double ring resonators are 472 GHz and 501 GHz, respectively. The lasing wavelength can be determined by the vernier effect of resonance wavelength difference between the two ring resonators. Therefore, the lasing wavelength can be tuned by changing refractive index of one ring resonator controlled by a micro heater through TO effect. At the butt-joint section of the SOA and the ring resonator wavelength filter, low loss and low reflective optical coupling is essential. In order to obtain low loss and low reflection at the end facets, we formed a spot size converter (SSC). The SSC has a tapered silicon waveguide core structure that gradually becomes thinner toward the end facet of the filter which adiabatically converts the optical field profile in the Si-wire waveguide to the waveguide in the SOA. The output waveguide of the tunable filter is tilted by 16° to couple with the waveguide of the SOA which is tilted by about 7° satisfying Snell's law.

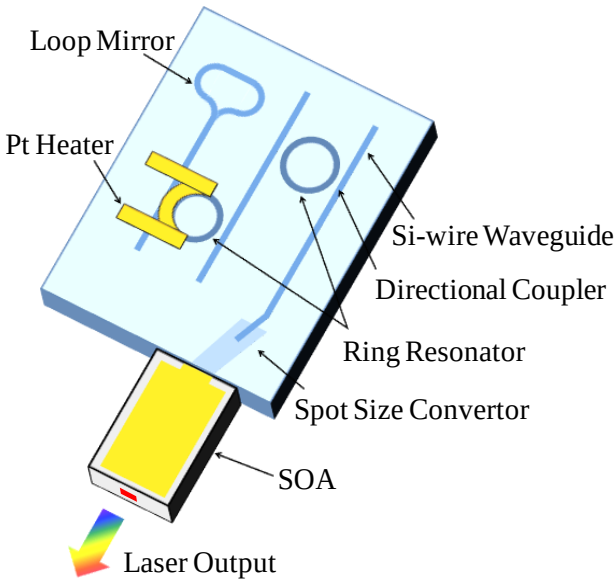


Figure 1. Schematic structure of the wavelength tunable laser

III. SPECTRAL LINEWIDTH OF LASING

Less than 100 kHz of narrow spectral linewidth is required for use as light sources for digital coherent systems. The spectral linewidth of semiconductor laser can be expressed by the Henry's formula [6];

$$\Delta\nu = \frac{\upsilon_g^2 h\nu n_{sp}(1+\alpha^2)}{8\pi P_0}\cdot(\alpha_i+\alpha_m)\cdot\alpha_m . \quad (1)$$

Here υ_g is the group velocity in the laser cavity, $h\nu$ is the photon energy, n_{sp} is the spontaneous emission coefficient, α is the linewidth enhancement factor, P_0 is the lasing output power, α_m is the mirror loss, and α_i is the internal loss. From the model of laser cavity as shown in Fig. 2, we expressed υ_g, α_m and α_i as

$$\upsilon_g = \frac{c}{n_g} = \frac{c}{\left(\frac{L_{SOA}n_{SOA}+L_{filter}n_{filter}}{L_{SOA}+L_{filter}}\right)} . \quad (2)$$

$$\alpha_m = \frac{1}{L_{SOA}+L_{filter}}\ln\left(\frac{1}{\sqrt{R_1}\sqrt{R_2}}\right) . \quad (3)$$

$$\alpha_i = \frac{L_{SOA}\alpha_{SOA}+L_{filter}\alpha_{filter}+\alpha_{coupling}}{L_{SOA}+L_{filter}} . \quad (4)$$

Here, n_g is the average group index of the laser cavity. n_{SOA} and n_{filter} is the group index of the SOA and the tunable filter, respectively. L_{SOA} and L_{filter} is the effective cavity length of the SOA and the tunable filter, respectively. R_1 and R_2 is the end-facet reflectivity of the SOA and the tunable filter, respectively. α_{SOA} and α_{filter} is the internal loss of the SOA and the tunable filter, respectively. $\alpha_{coupling}$ is the coupling loss between the SOA and the tunable filter. Those parameters used in this calculation were listed in Table I.

In order to obtain narrower spectral linewidth, longer L_{filter} is preferable according to (1) ~ (4). By decreasing optical coupling efficiency between bus waveguides and ring resonators, the number of turning around a ring resonator increase. This makes the effective cavity length of tunable filter drastically increases, and enables to narrow spectral linewidth. This is a big difference from conventional DBR lasers where the effective length is always smaller than the physical length of the optical cavities. The effective length of ring resonator is given by

$$L_{eff} = \left(\frac{1}{2}+\frac{1-K}{K}\right)\cdot L_r . \quad (5)$$

Here, L_{eff} is the effective length of ring resonator, K is coupling efficiency between the ring resonator and the bus waveguide and L_r is an orbiting length of the ring resonator. Total external cavity length is sum of the effective length of ring resonators and the total length of bus waveguides. We fabricated two types of tunable filters with external cavity length of 4.75 mm and 6.3 mm. The coupling efficiency for both was 0.086. We calculated spectral linewidth of the tunable lasers using (1) ~ (4). Less than 100 kHz spectral linewidth is estimated when laser output is 9 mW.

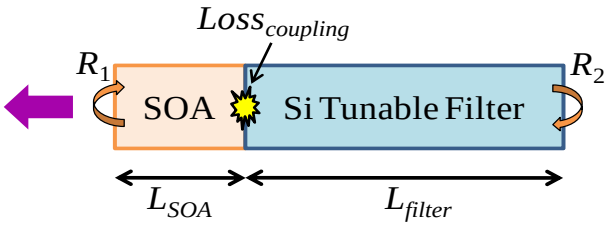


Figure 2. Calculated model of the laser cavity

TABLE I. PARAMETERS USED FOR CALUCULATION

L_{SOA}	1.5 mm	α	5.0
R_1	0.02	R_2	0.95
α_{SOA}	20 cm^{-1}	α_{filter}	0.69 cm^{-1}
$\alpha_{coupling}$	1.39 Neper	n_{SOA}	3.4
n_{filter}	4.3	n_{sp}	2
$h\nu$	0.78 eV		

IV. MEASUREMENT RESULT

Fig. 3 shows measured lasing spectra of the tunable laser superimposed for various lasing wavelengths. The temperature of the SOA chip was controlled to around 25°C by using a TEC, and injected current to the SOA was set to 80 mA. The lasing wavelength was controlled by heating one ring resonator which makes it possible to tune discrete wavelength interval. If both of the ring resonators were heated, the lasing wavelength could be continuously tuned. Fig. 4 shows the heating power and the side-mode suppression ratio (SMSR) as a function of lasing wavelength. The lasing wavelength was linearly changed for the heating power, and single-mode oscillation with more than 40 dB SMSR was obtained for the wavelength range. 45.1 nm of tuning wavelength range which covers entire L-band of optical communication wavelength range was obtained by 115.7 mW heating power. In our fabricated structure, the heating power might not be used effectively to heat ring resonator because heat from the micro heater diffuses into the large area of the substrate. By introducing a thermal isolating groove around the core of the waveguide, the core may be efficiently heated and power consumption in the heater may be drastically reduced [4].

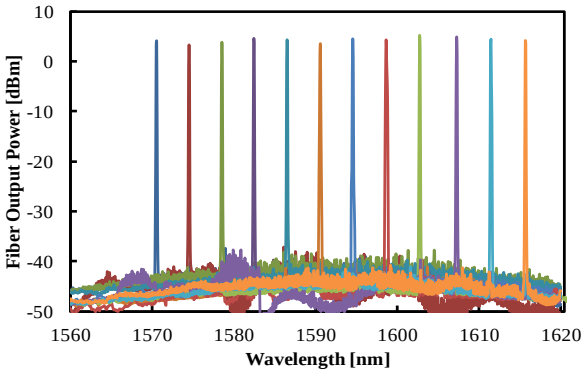


Figure 3. Measured lasing spectra of the tunable laser

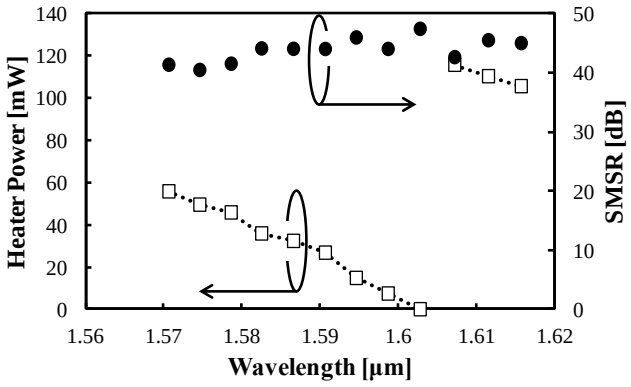


Figure 4. Heating power and SMSR as a function of lasing wavelength

Fig. 5 shows a relationship between lasing output power and spectral linewidth measured by the delayed self-homodyne method using a 30 km long single-mode optical fiber as a delay line. Measured values of spectral linewidth were in partially agreement with those calculated. The possible reason that measured linewidth is little wider than that of calculated might that we calculated the complicated optical cavity of tunable lasers using a simple Fabry-Perot cavity model. 87.5 kHz of narrowest spectral linewidth was obtained when the external cavity length was 6.3 mm and the lasing output power was 9 mW. The spectral linewidth measured at wavelength range across L-band were shown in Fig. 6. Less than 100 kHz spectral linewidth was obtained for entire L-band.

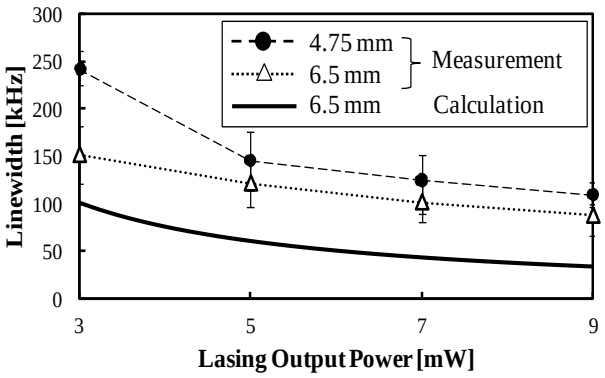


Figure 5. Relationship between spectral linewidth and lasing power

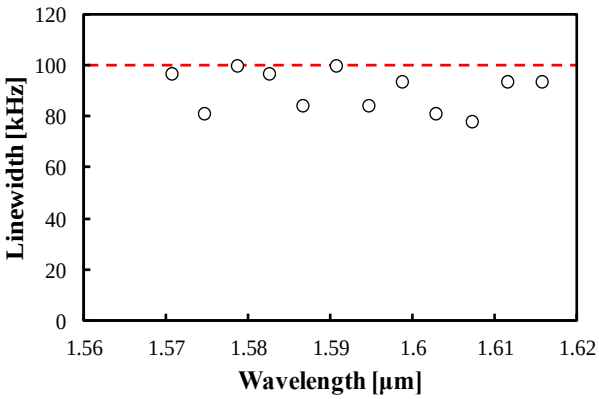


Figure 6. Measured spectral linewidth over wavelength range across L-band

V. CONCLUSION

We fabricated wavelength tunable laser diodes with Si-wire waveguide ring resonators as an external cavity. More than 45 nm wavelength tuning operation was obtained by about 115 mW heating power. Less than 100 kHz spectral linewidth was obtained for entire L-band by optimizing design of the ring resonators. The wavelength tunable laser diodes with narrow spectral linewidth are suitable as light sources for digital coherent optical communication systems.

REFERENCES

[1] M. Seimetz, "Laser linewidth limitations for optical systems with high-order modulation employing feed forward digital carrier phase estimation," OFC2008, OTuM2 (2008).

[2] T. Matsumoto, A. Suzuki, M. Takahashi, S. Watanabe, S. Ishii, K. Suzuki, T. Kaneko, H. Yamazaki and N. Sakuma, "Narrow spectral linewidth full band tunable laser based on waveguide ring resonators with low power consumption," OFC2010, OThQ5 (2010).

[3] T. Chu, N. Fujioka and M. Ishizaka, "Compact, lower-power-consumption wavelength tunable laser fabricated with silicon photonic-wire waveguide micro-ring resonators," Optics Express **17**(16), 14063-14068 (2009).

[4] N. Fujioka, T. Chu and M. Ishizaka, "Compact and low power consumption hybrid integrated wavelength tunable laser module using silicon waveguide resonators," Journal of Lightwave Technology **28**(21), 3115-3120 (2010).

[5] K. Suzuki, T. Kita and H. Yamada, "Wavelength tunable laser diodes with Si-wire waveguide ring resonator wavelength filters," Proc. SPIE 7943, 79431G (2011).

[6] C. H. Henry, "Theory of linewidth of semiconductor lasers," IEEE J. Quantum Electron. **18**(2), 259-264 (1982).

Design of Wireless Monitoring System Based on ZigBee and GPRS Technology

Wang Xiaoqiang
School of Electronic Engineering
University of Electronic Science and Technology of China
Chengdu, Sichuan, China
xiaoqiang08wang@163.com

Zhang Hongyu
School of Electronic Engineering
University of Electronic Science and Technology of China
Chengdu, Sichuan, China

Abstract—**With the development of wireless sensor and RFID technology, wireless monitoring system has attracted much attention. A good system can collect the sensor data periodically and transfer the data to the gateway automatically. Here the author proposed a general design reference based on ZigBee and GPRS to solve the problem of such systems where user can get the sensor data from sever and control the system by sending a message with the cell phone.**

Keywords-ZigBee; GPRS; wireless network;

I. INTRODUCTION

Wireless monitoring system is of great importance in many applications, due to the complexity and uncertainty of the environment as well as other factors. Since the monitoring system usually collects the data from sensors, it is reasonable for developing a common data transferring system.

ZigBee is an intelligent, easy-to-deploy and low-cost infrastructure technology, which provides a solid technological solution for the data acquisition system [1]. GPRS module can send the data to the Internet automatically and we can get the data from the server easily, therefore, combining ZigBee with GPRS is a promising way to establish a new wireless monitoring system.

As mentioned above, in this paper, the author presents a wireless monitoring system. By using this system, users can exchange various kinds of sensors according to the application and get the data from the sever easily.

II. SYSTEM DESIGN

The wireless monitoring system includes Sensor Node and ZigBee Gateway, as shown in Figure 1. The Sensor Node，which can join the ZigBee Network automatically when the network was established，collects data from the sensor and sends it to the Coordinator in the ZigBee Network. The ZigBee Gateway contains the ZigBee Coordinator and GPRS Module. The Coordinator can setup the ZigBee Network automatically when power is on by using z-stack. The ZigBee Coordinator and GPRS Module can exchange data and commands through UART interface.

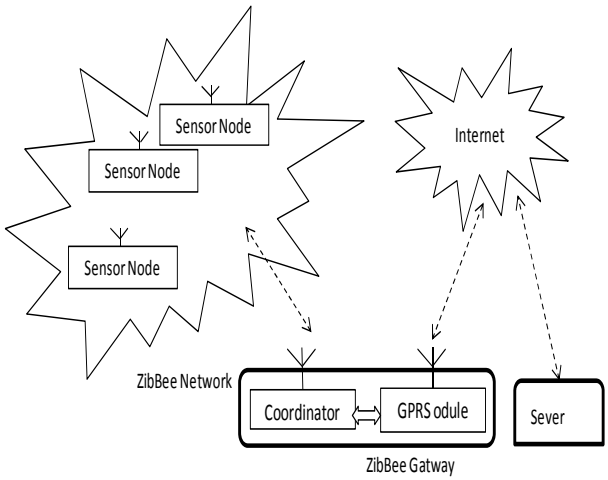


Figure 1 System block Diagram

In the ZigBee Network, the ZigBee Coordinator establishes the mesh network according to the compile options [2]. The Sensor Nodes join the Network automatically, and then data from a certain node can reach the Coordinator in the ZigBee network, regardless of the distance as long as there are enough routers in between to pass the message along.

III. HARDWARE DESIGN

The system hardware is composed of the Sensor Node's hardware and the Gateway's hardware. The Sensor Node is responsible for the data acquisition of various kinds of sensors, such as obliquity sensor, temperature sensor and accelerometer. The Gateway is responsible for collecting the data from the Sensor Nodes in the ZigBee Network and forwarding it to the Internet.

A. The Sensor Node's Hardware Desi*gn*

The Sensor Node's hardware is shown in Figure 2, including CC2530 RF chip and sensors. The CC2530 RF chip, which contains an IEEE 802.15.4-compliant radio transceiver, provides a System-on-Chip Solution for 2.4 GHz IEEE 802.15.4 and ZigBee Applications. The RF core controls the analog radio modules by using TI's Z-Stack, which provides a software support for the ZigBee protocol, through which users can build a ZigBee wireless network easily.

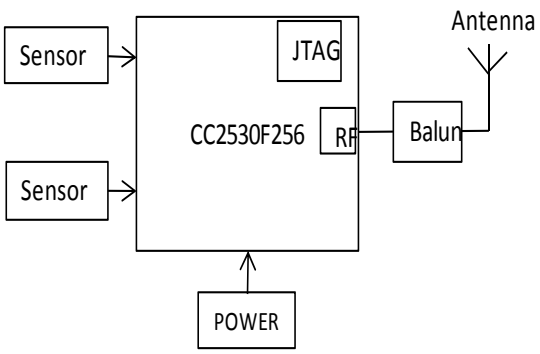


Figure 2 Sensor Node Hardware

The sensors can be connected to the CC2530 I/O pins, and the CC2530 can get the sensor data easily with the help of its onchip ADC. In some applications, a signal conditioning circuit, which would be used to provide a proper voltage level for the ADC pins, is needed.

B. The Gateway's Hardware Design

The Gateway's hardware is shown in Figure 3, including ZigBee Coordinator and GPRS Module [3]. The ZigBee Coordinator's hardware is similar to that of Sensor Node, however, the difference is that the ZigBee Coordinator's logical type is different from that of the Sensor Node, which is determined by the compile options in the Z-Stack.

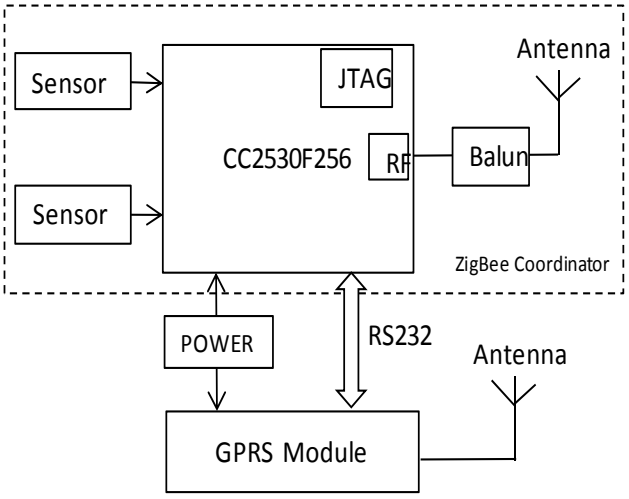


Figure 3 Gateway's hardware

The GPRS Module can send data to the sever automatically according to the configurations and communicate with the ZigBee Coordinator using RS232 interface.

IV. SOFTWARE DESIGN

The System software is composed of Sensor Node's software and ZigBee Gateway's software. The Sensor Node's software is running on the Sensor Node. The ZigBee Gateway's software is running on the Coordinator in the Gateway.

A. Sensor Node's software

The Sensor Node's software program flow chart is shown in Figure 4. The Sensor Node will try to join the ZigBee Network automatically when power is on and send Self-Info to the ZigBee Coordinator when joining the network sucessfully [4]. Then it will detect if there is ACK signal received which indicates that it has joined the ZigBee Netork and its Self-Info has received by the ZigBee Coordinator sucessfully. After that it will read the sensors data periodically and send it to the ZigBee Coordinator.

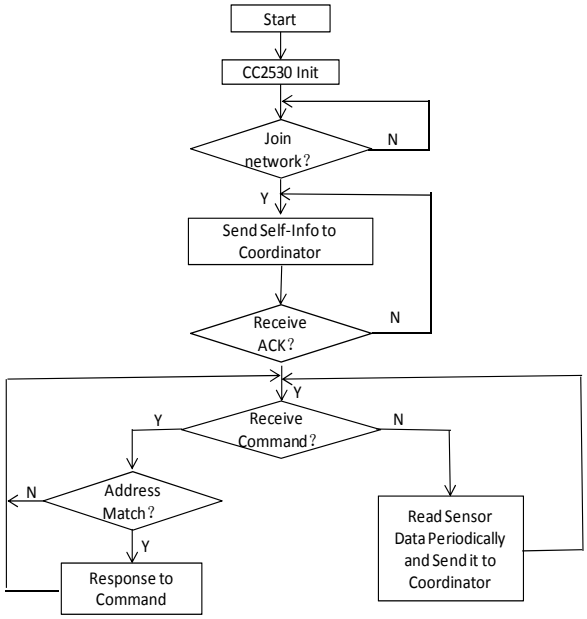


Figure 4 Sendor Node program flow chart

What's more, the Sensor Node's Self-Info can be implemented by using the data struct SELFINFO as fallows：

```
typedef union SELFINFO
{
        unsigned char TxBuf[16] ;
        struct RFTXBUF
        {
                unsigned char Head ;
                unsigned char Devicetype[3] ;
                unsigned char macaddr[8] ;
                unsigned char networkaddr[2] ;
                unsigned char CRC ;
                unsigned char Tail;
        } INFODATA ;
};
```

The SELFIFO containes the packet head and tail, 3-Bytes devicetype code, 8-Bytes Node's MAC address, 2-Bytes Node's network address and one byte checksum.

When the ZigBee Network is established the Node can send data to the ZigBee Coordinator, using the DATAFORMAT as follows:

```
typedef union DATAFORMAT
{
    uint8 DataBuf[39] ;
    struct DATABUF
```

```
{
        unsigned char Head ;
        unsigned char macaddr [8] ;
        unsigned char networkaddr [2] ;
        unsigned char parentmacaddr [8] ;
        unsigned char parentnetworkaddr [2] ;
        unsigned char DataBuf[16] ;
        unsigned char CRC ;
        unsigned char Tail ;
    }RXDATA ;
};
```

The DATAFORMAT containes the packet head and tail, 8-Bytes Node's MAC address, 2-Bytes Node's network address, 8-Bytes Node's parent's MAC address, 2-the Node's parent's network address, 16-Bytes sensor data and one byte checksum. Notice that the 8-Bytes Node's parent's MAC address, 2-the Node's parent's network address are used to form the network topology.

B. ZigBee Gateway's software

The ZigBee Gateway's software program flow chart is shown in Figure 5 and Figure 6.

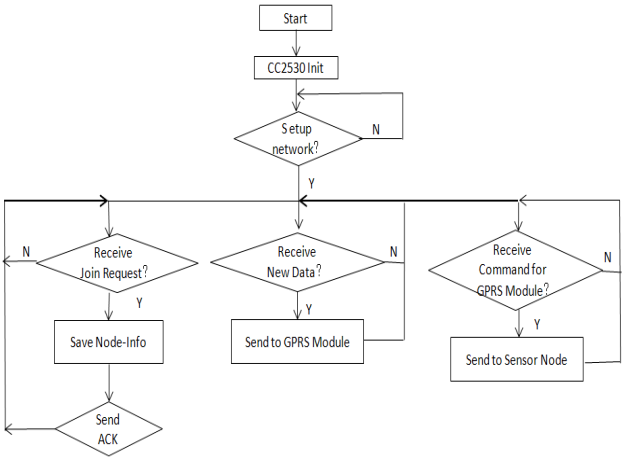


Figure 5 ZigBee Gateway program flow chart

The ZigBee Coordinator will try to setup the ZigBee Network automatically when power is on. Then it will test if there is a Node joining request or having received any new data coming from the Node, and if either is possible, it will process the event immediately.

OSAL is a multi-Non-preemptive task, polling operation System by means of using polling mechanism for task scheduling [5] ~[6]. Z-Stack running on the basis of OSAL, which provides services and managements such as task management, task synchronization, message management and memory management. Each task running independently in the Z-Stack can be seen as an OSAL task. Therefore, the multi-task programming becomes easy to implement.

The task initialization process is:

- Initialize the system task event handler list array tasksArr[].
- Assign a unique task ID and allocate the stack memory for each task.
- Register a application object in the AF layer using the afRegister () function
- Registered the corresponding OSAL or HAL system services.

The data format used between the Coordinator and the GPRS Module can be implemented by the data struct GATEWAYDATA as follows:

```
typedef union GATEWAYDATA
{
    uint8 DataBuf[29] ;
    struct UARTBUF
    {
        unsigned char Head ;
        unsigned char macaddr [8] ;
        unsigned char networkaddr [2] ;
        unsigned char DataBuf[16] ;
        unsigned char CRC ;
        unsigned char Tail ;
    }DATA ;
};
```

The GATEWAYDATA containes the packet head and tail, 8-Bytes Node's MAC address, 2-Bytes network address, the 16-Bytes sensor data and one byte checksum [6].

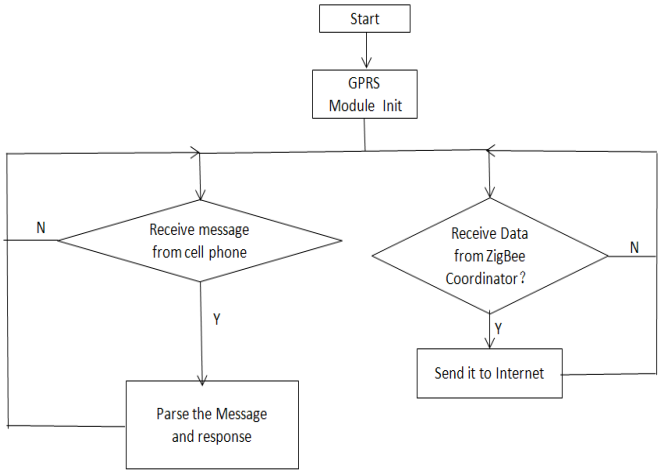


Figure 6 GPRS Module program flow chart

The GPRS Module will load the conFigure uration options when power is on. Then it will detect whether there is a message coming from user's cell phone or any data transferred from the ZigBee Coordinator [7]~[8]. If a message has been received, the GPRS Modele will parse and response to it according to the content. What's more, when the data transferred from the ZigBee Coordinator has been received, it will send it to Internet automatically. Therefore, users can get the data from the sever.

V. CONCLUSION

Based on ZigBee and GPRS Module, we not only put forward a wireless monitoring system, which combines the

ZigBee’s intelligent network building with GPRS’s advantage in long distance data transfering, but demonstrate this design in both hardware and software. The system can be used in remote control and wireless monitoring applications. The further research is to pay specific attention to ZigBee Gateway functional expansion References

[1] ZigBee Standards Organization. ZigBee Specification.2007.

[2] Texas Instruments Inc, Z-Stack Sample Application for CC2530-DB[Z].2010

[3] Drew Gislason, ZigBee Wireless Networking.2008

[4] Shahin Farahani, ZigBee Wireless Networks and Transceivers.2008

[5] Texas Instruments Inc, Z-Stack Developer’s Guid for CC2530.2010

[6] Li Jun, Huang Lan and Wang Zhong-yi, Energy Management Strategy for WSN Based on Z-Stack, Computer Engineering, vol. 37, no.7, 2011, pp.121–124.

[7] 孙维新.基于 GPRS 的制冷设备远程监控系统的设计与实现[D].成都：电子科技大学, 2007.

[8] 刘文萍,杨学超,王景中.基于 GPRS 的远程报警视频监控系统[J].计算机工程, 2007(24):253257.

Parking Assist System Using Fisheye lens

MinSung Kang
Dept. of Info. and Comm. Eng.,
Yeungnam University
Gyeongsan, Korea
zxc2100@ynu.ac.kr

Yongwan Park
Dept. of Info. and Comm. Eng.,
Yeungnam University
Gyeongsan, Korea
ywpark@yu.ac.kr

Abstract—**in this paper, we offer Parking Assist System. The current system using cameras or sensors on the driver by an acoustic parking tells about the environment. Our systems use wide angle of fisheye lens camera with existing systems. So we can provide wide visibility to the driver.**

Keywords-component; Parking Assist Systes, Fisheye lens

I. INTRODUCTION (HEADING 1)

Recently, many consumers have focused on Parking Assist System. J. D Power of the '2001 Emerging Technologies Study', according to, 66 percent of consumers willing to purchase a Parking Assist System that was answered, as consumers interest has led the actual purchase. To meet this consumer demand of Parking Assist System is being developed in various forms. Parking Assist System can be divided into the two broad.

First method is automatically parking. This systems using sensor in mounted front and back of car and System detect of parking space. This systems is very helpful because automatically detection of space and parking. Recently, these technical is mounted on the vehicle, but the cost is very high to consumers.

Figure 1. Parking Assist System using sensor

Figure 2. Parking Assist System using camera

Second method provides environment information of parking. This system use sensor or camera. The system in Fig. 1 uses sensor that provide distance to the object. The system in Fig. 2 uses a camera that provide estimated steering angle. In addition, method is a bird's-eye view around the vehicle

The advantage of the Sensor is precise measurement of the distance. But the sensor cannot use object of wires, ropes, chains, cotton, and sponge. In addition, higher or lower than of the bumper object is not available.

Camera is less precise measurement of the distance. But consumers prefer camera than sensor because camera image is very similar to person's eye

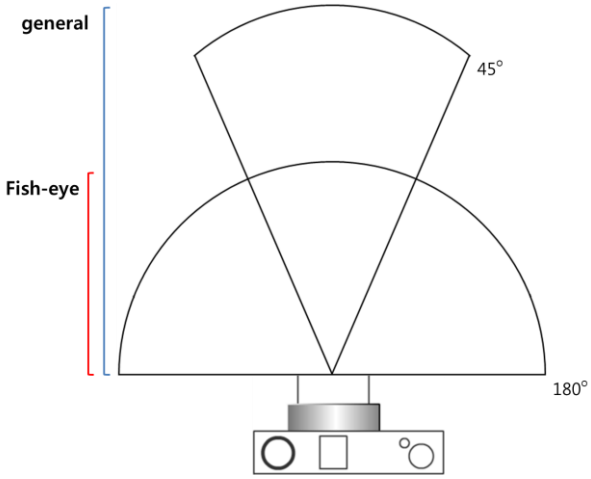


Figure 3. range of Fisheye lens

In this paper, we offer Parking Assist System using fisheye lens camera with existing systems. Angle of fisheye lens in Fig. 3 is 180°. So we can take maximum viewing angle by using minimal camera. And we calculate the distance to the output by a still image of fisheye lens.

II. MEASUREMENT OF DISTANCE USING FISHEYE IMAGE

A. Distortion correction of fisheye lens

Unlike a general camera lens, angle of fisheye lens is 180°. We can take wide still image. But fisheye lens make distortion. So we must be distortion correction.

Figure 4. Distortion still image

Figure 5. Correction still image

Fig. 4 is distortion still image of fisheye lens. In order to eliminate the distortion of the image reconcile the major axis of image with the minor axis of image. And we calculated camera parameters through a calibration process. Finally we make corrected view in Fig. 5 using camera parameter.

B. Measurement of the distance using correction image

Figure 6. remove noise

In order to remove noise of still image in Fig. 6, we uses sharpening, smooth, dilate, and erode.

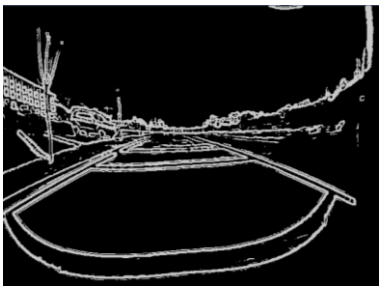


Figure 7. binarization

Second, binarization in Fig. 7 use sobel mask.

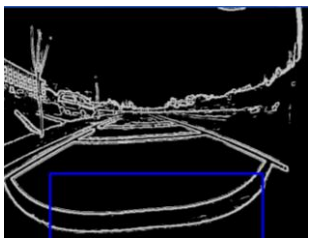


Figure 8 Set ROI

Third, ROI in Fig. 8 is set where location is high probability of lane

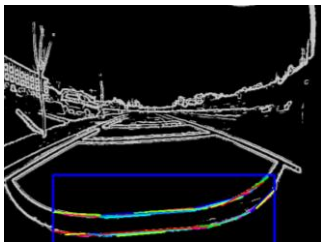


Figure 9. detection all straight line

Fourth, Detection all straight line in Fig. 9 use hough transform.

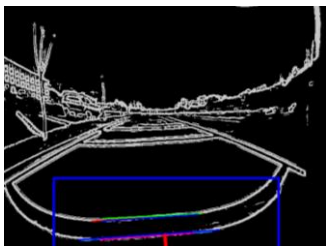


Figure 10. detection lane

Fifth, Detection lane in Fig. 10 use slope, length, location of line and difference slope, difference length, difference location between line and other line.

And we calculate pixel distance between fisheye and lane. Finally, Pixel distance changed the actual distance using the ratio between pixels and the actual distance, fisheye characteristic and perspective.

III. CONCLUSION

We offer Parking Assist System using fisheye lens. So we can take maximum viewing angle by using minimal camera. Simulation error is between 1cm and 5cm in actual parking environment. Measurement of the distance for Parking Assist System is similar compare camera with sensor. So using camera system performance is good than using sensor system.

REFERENCES

[1] Young Ju Cho, Sung Hee Kim, Ji Young Park, Jin Woo Son, Joong Ryoul Lee, Myoung Hee Kim, “Image Data Loss Minimized Geometric Correction for Asymmetric Distortion Fish-eye Lens” Vol. 19, No. 1, pp. 22-31, March 2010. (references)

[2] Ho Gi Jung, Hyoung Jin Kang, Jaihie Kim, “Target Parking Position Designation Method – Review and Perspective”

Evaluation of V_{th}-Variation Effect on Multiple-Valued Current-Mode Circuits

Kiyohiro Kashiuchi, Masanori Natsui, and Takahiro Hanyu
Research Institute of Electrical Communication, Tohoku University, JAPAN

1. Introduction

Multiple-valued current-mode (MVCM) circuit has an attractive feature to reduce the number of active devices and wiring in computation module because the frequently used linear sum operation can be performed simply by wiring. It is effective in designing high-speed/low-power VLSI, however, its realization requires measures against process variability due to further miniaturization of MOS transistors. In this paper, DC characteristics and variability of two types of differential-pair circuits, which are important components in the MVCM circuit, are evaluated for designing high-performance and highly-dependable MVCM circuits.

2. Differential-Pair Circuit

Fig.1(a) shows a circuit diagram of the simple differential-pair circuit. It uses a gate-grounded PMOSFET as the linear resistance. When an input voltage (V_{in}) is greater than a voltage for comparison (V_t), the drain current on M1 becomes larger than that on M2. At this time, the voltage of V_{gt1} becomes less than that of V_{gt2} due to the voltage drop. In the same way, when V_{in} is less than V_t, the voltage of V_{gt1} becomes greater than that of V_{gt2}.

Fig.1(b) shows a circuit diagram of the different-pair circuit with active load[1]. The active load consists of cross-coupled PMOSFETs and diode-connected PMOSFETs. The use of the diode-connected structure can improve the switching delay, and as a result, both large output voltage amplitude (the difference between V_{gt1} and V_{gt2}) at small input voltage amplitude (the difference between V_{in} and V_t) and the high-speed switching can be obtained.

3. Evaluation

Fig.2 and Fig.3 show the DC characteristics of the two types of differential-pair circuit where input voltages are changed from 0 to 1.2V in the condition of voltage source (V_{dd}) = 1.2V and V_t = 0.6V. Note that the threshold voltage (V_{th}) variation with ±0.05V are applied to each NMOSFET. We can confirm that the circuit with active load has larger output voltage amplitude than that with linear load, while the V_{th} variation seems to cause larger performance degradation in the output of the former one.

4. Conclusion

For designing variation-aware MVCM circuit in future nano-scale process technology, we should devise a new schematic of MVCM circuit components that are less subject to V_{th} variation with small performance overhead. As a future prospect, we will consider new technologies for variation-aware VLSI in various design level, such as post-process-oriented variation compensation technique and new-device-based logic architecture.

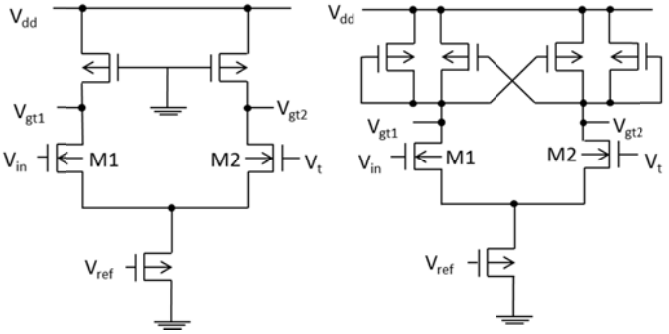


Fig.1. Two types of schematics differential-pair circuit: (a) linear load type, (b) active load type.

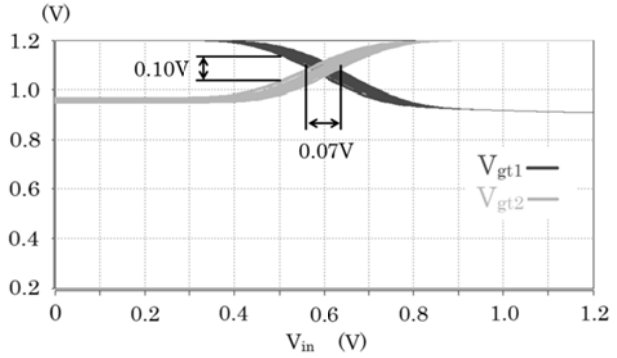


Fig.2. DC characteristics of differential-pair circuit with linear load.

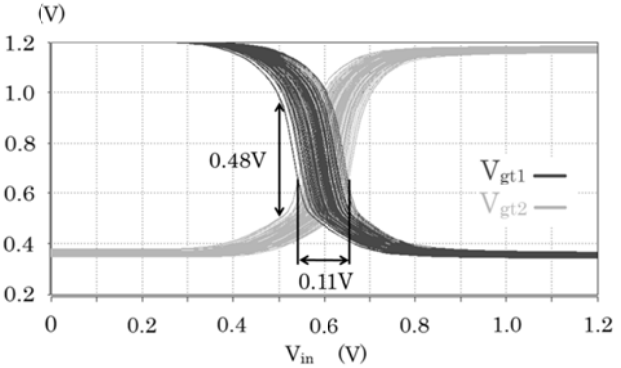


Fig.3. DC characteristics of differential-pair circuit with active load.

Reference

[1] A.Mochizuki, et al., "High-Performance Multiple-Valued Comparator Based on Active-load Dual-Rail Differential Logic for Crosstalk-Noise Reduction," ISMVL2007.

A Novel MASH Structure for Fractional-N Synthesizer

Bohai Zhang
Department of Integrated Circuit and System
UESTC
Chengdu, China
tmac.bohai@gmail.com

Songbai He
Department of Integrated Circuit and System
UESTC
Chengdu, China
sbhe@uestc.edu.cn

Abstract—**The Fractional-N synthesizers can achieve low phase noise, high frequency resolution and fast frequency switching at the same time. A Sigma-Delta Fractional-N synthesizer can achieve better phase noise performance by the Sigma-Delta modulators' spur suppression method. A novel structure MASH 1-2-2 with dithering is presented in this paper to attenuate the noise power in-band. Then a new circuit structure of PLL is presented and it adopts the novel structure MASH 1-2-2 with dithering. The phase noise value is -92.13dBc/Hz at 1 KHz offset the carrier frequency and -108.47dBc/Hz at 10 KHz offset the carrier frequency, 55.001MHz.**

Keywords-Fractional-N synthesizers; MASH; phase noise

I. INTRODUCTION

High demanding on frequency synthesizers is put forward with the rapid change of the wireless systems. The classical PLL synthesizers with integer-N dividers suffer from limitations in meeting various requirements such as low phase noise and fast frequency switching. The Fractional-N synthesizers overcome the limitation by changing the value of a programmable divider, and then high frequency resolution can be achieved by the average division ratio of the divider's changing values without using low reference frequency. Compared to the integer-N PLL, the Fractional-N synthesizers have higher phase comparison frequency and then the Fractional-N synthesizers can achieve low phase noise, high frequency resolution and fast frequency switching at the same time. Sigma-Delta modulation is regarded as a better technique to achieve the Fractional-N synthesizers [1]. Fig.1 shows the block form of Fractional-N synthesizers using the technique of Sigma-Delta modulator. F is the fractional part of frequency dividing ratio, and the clock of the Sigma-Delta modulator is the reference frequency. Because the PLL work as low-pass filter to the quantization noise, the noise out of band will be eliminated. That's why we need less noise power in the band. In the Fig.1, PD is phase discrimination; LPF is low-pass filter; VCO is voltage controlled oscillator; ÷N is programmable frequency divider controlled by Sigma-Delta modulator. Then the f_r is the reference frequency and the f_o is the output frequency. As a deterministic finite state machine, when the input of the MASH DDSM is constant and the quantization noise exists in the spectrum periodically. There are many factors affecting the period of the quantization noise. For example, the initial condition of the register, the input value, the modulus of the quantizer. The shorter length produced by the MASH, the higher the noise power shared by per tone. Increasing the length of the MASH's output can spread the noise power over more tones and the noise power of the PLL's bandwidth can be smaller.

The objective of this work is to develop a novel MASH structure whose length of the output is longer than the classical MASH, and then it will have much more smaller noise power in the bandwidth of the PLL. In the following we will present the new structure of the MASH and the spectrum simulation the MASH, and then we compare the performance of the new structure and the classical MASH in a specific PLL hardware.

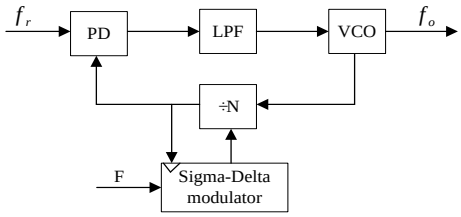


Figure 1. Block form of Fractional-N synthesizers

II. A NOVEL STRUCTURE MASH 1-2-2 WITH DITHERING

Increasing the order of the MASH DDSM can increase the length of the MASH's output. Similarly, adding random dithering can break down the period of the MASH's output and then the noise power can spread to more tones. However, adding random dithering can bring in extra noise to the system, and it should be carefully considered [2].

In Fig.2 we show a first order error feedback modulator with dithering. The dithering comes from $y_1[n]$, which is one time delay of the output y[n] multiply by the constant a. M is the step size of the quantizer and $M=2^{n0}$ and n0 is the word length of the input.

The first order error feedback modulator with dithering introduces a feedback line from the output to the input, so the signal transfer function STF (z) and noise transfer function NTF (z) become:

$$STF(z)=\frac{1}{1-az^{-1}} \quad (1)$$

$$NTF(z)=\frac{1-z^{-1}}{1-az^{-1}} \quad (2)$$

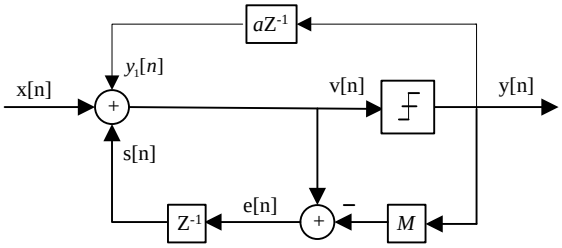


Figure 2. The first order error feedback modulator with dithering (EFM1)

The advantage of this first order error feedback modulator is that the length of the output can achieve the maximum with a specific a, and the specific a must meet the requirement that M-a is the largest prime number less than M. Table Ⅰ gives some example of specific a according to the word length of the input.

TABLE I. THE SPECIFIC A WITH RESPECT TO THE WORD LENGTH N0

N0	a
5,7,13,17,19,31	1
9,10,12,14,20,22,24,29	3
8,18,25,26,32	5
11,12	9

In the Fig.3 we show a second order error feedback modulator with dithering. Compared with first order error feedback modulator with dithering, the second order error feedback modulator suppresses the quantization noise much better. We can easily see it from the signal transfer function STF (z) and noise transfer function NTF (z):

$$STF(z)=\frac{1}{1-az^{-1}} \quad (3)$$

$$NTF(z)=\frac{(1-z^{-1})^2}{1-az^{-1}} \quad (4)$$

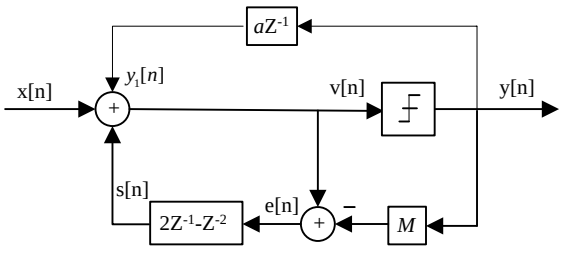


Figure 3. The second order error feedback modulator with dithering (EFM2)

MASH DDSM 1-2-2 is the combination of a cascade of EFM1, EFM2, EFM2 and some delay of signal. As shown in Fig.4, the quantization error e[n] of each stage is fed to the next stage and the output y[n] is fed to some delay network. Compared to the classical MASH structure, this new structure introduces a feedback line from the quantization error $e_3[n]$ to the input x[n]. $e_3[n]$ is a long sequences and it can be used as dithering added to the input without extra hardware. The transfers function from $e_3[n]$ to x[n] is V[Z] and V[Z] is used to eliminate the effects of the noise. The V[Z] and the relationship between X[Z] and Y[Z] are given as followed:

$$V(z)=(1-z^{-1})^l \quad (5)$$

$$Y(z)=X(z)+(1-z^{-1})^5 E(z) \quad (6)$$

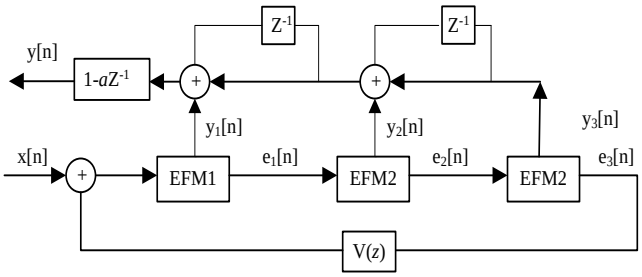


Figure 4. MASH 1-2-2 with dithering

III. DESIGN AND SIMULATION

The Sigma-Delta modulator adopts the new structure we describe in the last chapter, MASH 1-2-2 with dithering. In order to testify the performance of the MASH 1-2-2 with dithering, we compare the power spectra of the MASH 1-2-2 with dithering with the classical MASH 1-2-2 without dithering [3][4]. Fig.5 shows the output spectra of 13 bits MASH 1-2-2 with dithering and 13bits MASH 1-2-2 without dithering. We can easily see MASH 1-2-2 with dithering have lower noise power in the bandwidth of the PLL.

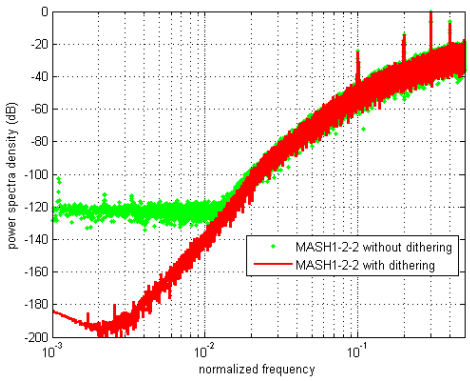


Figure 5. Output spectra of a 13 bits MASH 1-2-2 with dithering, a 13 bits MASH 1-2-2 without dithering with normalized input 0.3

The low-pass filter consists of two first order RC filters and an active filter. The use of active filter is to provide enough controlled voltage and the use of the two RC filters are to make enough attenuation of reference spurs. Besides, In order to achieve low noise, the value of resistance is chose as small as possible.

In this work, we adopt the AD9901-type PFD (phase frequency discrimination). As is shown in the Fig.6, the main components include four "D" flip-flops, an exclusive-Or gate (XOR) and some combinational output logic. The circuits operate in two distinct modes: as a linear phase detector and a frequency discriminator. The benefits of these circuits are that when the phase difference is π, then the PLL is locked, and it avoids the "dead zone" of the phase frequency discrimination.

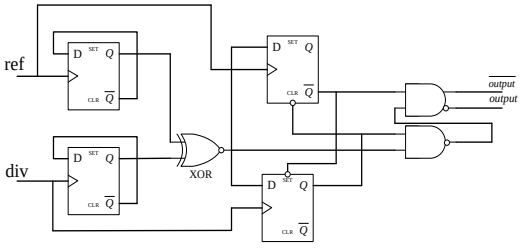


Figure 6. The block form of PD

IV. HARDWARE IMPLEMENTATION

In order to testify the performance of the MASH 1-2-2 in the hardware, we apply it to a project of 38-76 MHz PLL. A 32 bits MASH 1-2-2 with dithering was modeled using VHDL hardware description language and implemented on a Xilinx Spartan3 XC3S400. The block form of the PLL is shown in Fig.7 and the circuit diagram is shown in Fig.8.

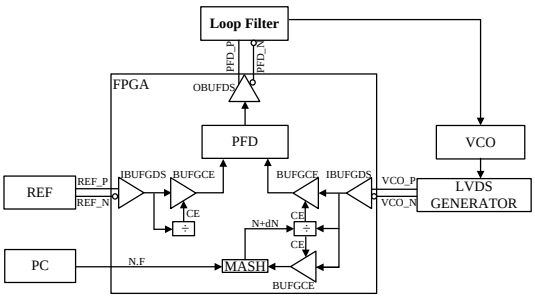


Figure 7. block form of the PLL

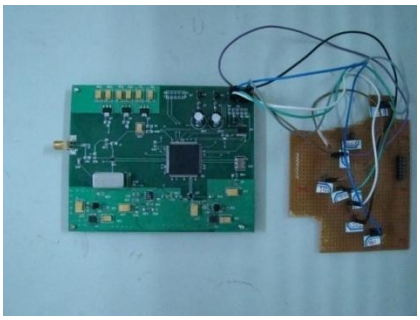


Figure 8. circuit diagram of the PLL

As is shown in the Fig.7, The PFD output PFD_P and PED_N, the signal REF_P and REF_N which come from REF (reference clock), and the signal VCO_P and VCO_N which come from VCO all adopt the signal type "LVDS", which will make sure their amplitude so small that they cannot modulate each other via ground bounce, supply dip and so on. Besides, in order to prevent the modulation between VCO signal and REF signal, The VCO signal and REF signal must enter on opposite sides of the FPGA and PFD output exists on a third part.

This Fractional-N synthesizer covers the frequency from 38 to 76MHz, and the reference frequency and VCO dividers [5], MASH 1-2-2 with dithering and AD9901-type PFD are all implemented in a Xilinx Spartan3 XCS400. The PFD rate is 1MHz, and the VCO and reference oscillator are low phase noise oscillators.

In order to attenuate the reference spurs, the bandwidth of the PLL is set to 2 KHz. Phase margin is 59.8° and magnitude margin is 28.7 dB. The two indexes meet the requirements of stabilities of the PLL.

Fig.9 shows the output phase noise of the Fractional-N synthesizer at 50.001MHz with the new structure of MASH 1-2-2 with dithering. The phase noise value is -92.13dBc/Hz at 1 KHz offset the carrier frequency. The phase noise is -108.47dBc/Hz at 10 KHz offset the carrier frequency [6]. And the spurs are attenuated under the noise floor. Table II shows the phase noise comparisons between MASH 1-2-2 with dithering and MASH 1-2-2 without dithering. The phase noise can improve 2-4 dB at the offset frequency we adopt.

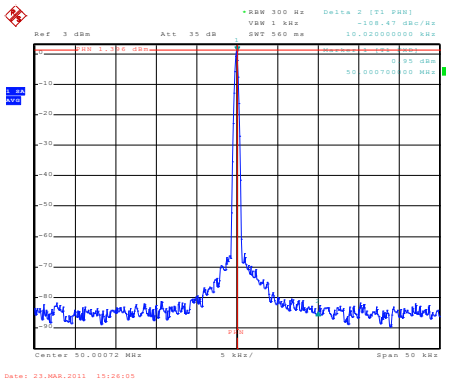


Figure 9. output phase noise of the Fractional-N synthesizer (Span=5MHz, RBW=30 KHz)

TABLE II. PHASE NOISE COMPARISONS BETWEEN MASH 1-2-2 WITH DITHERING AND MASH 1-2-2 WITHOUT DITHERING

Frequency offset	1KHz	10KHz	100KHz
MASH 1-2-2 with dithering(dBc/Hz)	-92.13	-108.47	-115.61
MASH 1-2-2 without dithering(dBc/Hz)	-89.59	-104.80	-111.32

V. CONCLUSION

In order to achieve better noise performance, a novel structure MASH 1-2-2 with dithering was presented. The simulation shows that MASH 1-2-2 with dithering has lower

noise power than MASH 1-2-2 without dithering in-band. Then the MASH 1-2-2 with dithering was achieved in a FPGA to testify its performance. And then a new circuit structure of PLL was presented and it adopted the novel structure of MASH 1-2-2 with dithering. The output spectrum shows that the phase noise can improve 2-4 dB at the offset frequency when we use the novel structure MASH 1-2-2 with dithering.

REFERENCES

[1] Tom A. D. Riley. Delta-Sigma Modulation in Fractional-N Frequency Synthesis. IEEE Journal Of Solid-state Circuit, Vol. 28, No. 5. May 1993.

[2] K. Hosseini and M. P. Kennedy. Maximum sequence length MASH digital delta sigma modulators. IEEE Trans. Circuits and Systems I: fundamental theory and applications, vol. 54, no. 54, pp. 2628–2638, Dec. 2007.

[3] M. Kozak and I. Kale. Rigorous analysis of delta–sigma modulator for fractional-N PLL frequency synthesis. IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 51, no. 6, pp. 1148–1162, Jun. 2004.

[4] M. J. Borkowski, T. A. D. Riley, J. Hakkinen, and J. Kostamovaara. A practical modulator design method based on periodical behavior analysis. IEEE Trans. Circuits Syst. II, Exp. Briefs, vol. 52, no. 10, pp. 626–630, Oct. 2005.

[5] P. Larsson. High-speed architecture for a programmable frequency divider and a dual-modulus prescaler. IEEE Journal of Solid-State Circuits, 31(5):744-748, May 1996.

[6] S. Pamarti and I. Galton. LSB dithering in MASH delta-sigma D/A converters. IEEE Trans. Circuits and Systems I, vol. 54, no. 4, pp. 779–790, Apr. 2007.

Enhancing Radar Angular Resolution Algorithm with Maximum a Posteriori

Guan Jinchen, Huang Yulin, Yang Jianyu
School of electronic engineering, University of Electronic Science and Technology of China
Chengdu, Sichuan, 611731, P.R.China
E-mail:jchguan@yahoo.com.cn

***Abstract*—With pulse compression technique providing a means of achieving good detection range and fine range resolution, angular resolution is the key parameter of radar. Conventional means of increasing angular resolution are dependent on the size of the antenna aperture. In framework of maximum a posteriori, this paper presents an iterative deconvolution algorithm to restore the target location information and then to enhance angular resolution. Simulation results show that the algorithm can efficiently attenuate the effect of antenna pattern convolution and highly enhance angular resolution; real radar experiment results show an order of improvement in angular resolution over real aperture image, and confirm that the algorithm could obtain angular superresolution.**

Keywords-angular resolution; maximum a posteriori; iterative deconvolution

I. INTRODUCTION

With pulse compression technique providing a means of achieving fine range resolution, the largest remaining constraint in imaging radars is azimuth resolution. In an unprocessed scanning real beam radar system, the angular resolution θ is limited by the antenna aperture size D to the value $\theta \propto \lambda / D$, where λ is the wavelength[1]. While beam width control is adequate for target detection, prohibitively large antennas result for the narrow beams required to determine detailed characteristics such as the number, size or separation of targets.

The need for information of this sort led to the development of synthetic aperture radar. Unfortunately SAR systems also have limitations. Synthetic aperture technique (SAR, ISAR) obtains a larger synthetic aperture through strict correlation processing, SAR processing require radar platform motion as well as a considerable amount of dwell time; in ISAR processing, target rotation is outside the control of the radar, and thus, not an optimal solution[2]. Furthermore, SAR systems are unable to image at forward-looking, are expensive, and are hardware dependent, most importantly, they are comparatively slow in imaging due to the extensive amount of signal processing required. These limitations have led to renewed efforts to improve the angular resolution of real beam systems. Real aperture radar improves the azimuth resolution mainly by increasing the physical aperture, but constrained by the antenna weight, size and other physical factors, it's difficult to obtain high azimuth resolution.

Radar response is the convolution of the antenna pattern and the surface scatter, thus deconvolution methods could recover the target location information in theory, past attempts of deconvolution to enhance radar angular resolution fail because the presence of the noise and pattern zero make the deconvolution to an inherent ill-posed problem. In order to solve this problem, many deconvolution methods have presented. In [3], a fast CID algorithm is proposed to avoid division in frequency domain, but azimuth resolution improvement is approximately 2 times when the SNR is 10dB; in [4], authors develop a generalized inverse filtering method and deconvolution is computed in time domain, however, computation is too complex to fit for real-time implement; in [5],[6], a scheme of multi-channel deconvolution for monopulse radar is presented, a order improvement in angular resolution when the SNR is higher than 30dB. These exiting methods can't enhance resolution high enough or require higher SNR.

This paper presents an iterative method to achieve radar angular super-resolution. In framework of maximum a posteriori, the iterative algorithm is derived to restore the target location information, in theory the algorithm converges quickly and has low calculation complexity. Both simulation and real radar data experiment results prove that algorithm could achieve higher angular resolution by a few iterations at low SNR condition.

II. SCANNING RADAR SIGNAL MODEL

Imagine a mapping radar scanning a region containing only a few point targets. As shown in Fig. 1, the baseband video output signal will trace out the antenna voltage pattern as it scans past an isolated single reflector; if two point reflectors are close together, the output voltage will be proportional to the superposition of two replicas of the antenna pattern, appropriately separated.

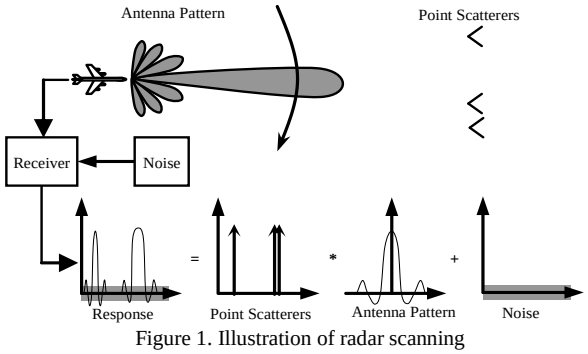


Figure 1. Illustration of radar scanning

This project was supported by the Basic Research Program
（No.B1420110182）

Clearly, if two targets are too close together the composite response will have a single peak and they will not be resolved. Furthermore, the spacing limit is determined by the beam width of the antenna pattern.

Consider a radar scanning a region, antenna pattern is equivalent to the angular impulse response, and return in angular is the convolution of the antenna pattern and the surface scatterers. Consider return in only one range bin: $\sigma(\theta)$ denotes the surface scatterers of M point, $P(\theta)$, $n(\theta)$ denote antenna pattern and noise respectively, then signal model for the return in one range bin is given as:

$$r(\theta) = \sigma(\theta) * P(\theta) + n(\theta) \tag{1}$$

where $r(\theta)$ is return, " $*$ " is convolution operator. In frequency domain, signal model (1) is expressed as:

$$R(\omega) = \sigma(\omega)P(\omega) + N(\omega) \tag{2}$$

Where ω is spatial frequency.

Typical antenna patterns result in strong lowpass filtering, and low resolution in the output data results from filtering out high spatial frequencies from the true scene. In time domain, the azimuth resolution is proportional to the beam width, pattern modulation lead to the poor azimuth resolution; in frequency domain, the band of pattern is finite, after the convolution the spectrum outside antenna pass-band is equal to zero, azimuth resolution narrows down.

This filtering interpretation of the sensor physics clearly suggests that we can restore the lost resolution by deconvolution, or inverse filtering. In theory, using knowledge of the antenna pattern to recover surface scatterers:

$$\tilde{\sigma}(\omega) = r(\omega) / P(\omega) \tag{3}$$

In practice, conventional inverse filtering does not work. The reason is the band limited nature of $P(\omega)$, and (3) requires to form $1/P(\omega)$; since $P(\omega)$ is zero for spatial frequencies above some cutoff value, $1/P(\omega) \to \infty$ at those frequencies, resulting in tremendous noise amplification and useless results.

III. ALGORITHM WITH MAXIMUM A POSTERIORI

For mathematical simplicity, signal model in (1) can be expressed in a matrix-vector form [7]:

$$\bar{r} = \bar{P}\bar{\sigma} + \bar{n} \tag{4}$$

where $\bar{P}$ is the convolution operator, defined on size $M \times M$, $\bar{r}, \bar{\sigma}, \bar{n}$ are vectors of size $M \times 1$ containing return, surface scatterers and noise respectively.

Refer to MAP (maximum a posteriori), the actual $\bar{\sigma}$ should be $\tilde{\bar{\sigma}}$.

$$\tilde{\bar{\sigma}} = \arg \max_{\bar{r}} [\, p(\bar{r} \,|\, \bar{\sigma})\, p(\bar{r})] \tag{5}$$

The iterative method is derived in framework of maximum a posteriori, considering that the return $\bar{r}$ is the convolution of pattern and surface scatterers, then the expected value at the ith element in the return is $E(r_i) = \sum_j p_{ij}\sigma_j$,where h_{ij} is the (i ,j)th element of matrix $\bar{P}$, σ_j is the jth element of vector $\bar{\sigma}$ .In the case of Poisson noise, the actual ith pixel value r_i in $\bar{r}$ is the one realization of Poisson distribution with mean $\sum_j p_{ij}\sigma_j$. Thus, we have the following relation:

$$p(r_i / \bar{\sigma}) = \frac{(\sum_j p_{ij}\sigma_j)^{r_i} e^{-\sum_j p_{ij}\sigma_j}}{r_i!} \tag{6}$$

Each element in return $\bar{r}$ is realized by an independent Poisson process, then:

$$p(\bar{r} / \bar{\sigma}) = \prod_i \frac{(\sum_j p_{ij}\sigma_j)^{r_i} e^{-\sum_j p_{ij}\sigma_j}}{r_i!} \tag{7}$$

Assuming that $p(\bar{\sigma})$ is Poisson's distribution, then:

$$p(\bar{\sigma}) = \prod_j \frac{(\bar{\sigma}_j)^{\sigma_j} e^{-\bar{\sigma}_j}}{\sigma_j!} \tag{8}$$

where $\bar{\sigma}_j$ is the probability of the j position.

An approximate solution of (7) and (8), for given observed return $\bar{r}$ is obtained by maximizing MAP $p(\bar{r} / \bar{\sigma})p(\bar{r})$, or equivalently $\ln(p(\bar{r} / \bar{\sigma}))p(\bar{r})$.

$$J(\bar{\sigma}) = \ln p(\bar{r} / \bar{\sigma}) + \ln p(\bar{\sigma}) \tag{9}$$

Differentiating $\ln(p(\bar{r} / \bar{\sigma}))$ with respect to $\bar{\sigma}$, we get following relation:

$$\frac{\partial J(\bar{\sigma})}{\partial \bar{\sigma}} = \frac{\partial \ln p(\bar{r} / \bar{\sigma})}{\partial \bar{\sigma}} + \frac{\partial \ln p(\bar{\sigma})}{\partial \bar{\sigma}} = 0 \tag{10}$$

Using the (11),

$$\ln m! \approx m \ln m - m \tag{11}$$

Then

$$\begin{aligned} \frac{\partial J(\bar{\sigma})}{\partial \bar{\sigma}}\Big|_i &= \frac{\partial}{\partial \sigma_i}[\sum_j r_j \ln(\bar{P}\bar{\sigma})_j - \sum_j (\bar{P}\bar{\sigma})_j] \\ &\quad + \frac{\partial}{\partial \sigma_i}[\sum_j \sigma_j \ln(\bar{\sigma}_j) - \sigma_j \ln \sigma_j + \sigma_j] \qquad (12) \\ &= \ln(\bar{\sigma}_j) - \ln \sigma_j + \sum_j p_{ji}[\frac{r_j}{\sum_l p_{jl}\sigma_l} - 1] \end{aligned}$$

Setting the result to zero,

$$\ln(\bar{\sigma}_j) - \ln \sigma_j + \sum_j p_{ji}[\frac{r_j}{\sum_l p_{jl}\sigma_l} - 1] = 0 \quad (13)$$

Because of the complexity of solving the (13), using the iterative method to resolve.

$$\bar{\sigma}^{k+1} = \phi(\bar{\sigma}^k) \quad (14)$$

where $\Phi(\bullet)$ is the renewal function. Rearranging (14), then

$$\sigma_i = \bar{\sigma}_i \exp[\sum_j P_{ji}[\frac{r_j}{\sum_l P_{jl}\sigma_l}] - 1] \quad (15)$$

Suppose that $\bar{\sigma}_i$ is equal to the current σ_i^k ,then

$$\sigma_i^{k+1} = \sigma_i^k \exp[\sum_j P_{ji}[\frac{r_j}{\sum_l P_{jl}\sigma_l}] - 1] \quad (16)$$

and matrix and vector expression is:

$$\bar{\sigma}^{k+1} = \bar{\sigma}^k \exp[\bar{P}^T[\frac{\bar{r}}{\bar{P}\bar{\sigma}^k}] - 1] \quad (17)$$

where superscript T denotes transpose of matrix, and $\bar{r}/(\bar{P}\bar{\sigma})$ denotes the vector obtained by componentwise division of $\bar{r}$ by $\bar{P}\bar{\sigma}$.

Deconvolution processing overcomes poor resolution limit of antenna, achieve higher resolution in time domain or spectrum extrapolation in frequency domain. Antenna pattern passband is limited, the spectrum of return is zero outside passband leading to lower resolution; superresolution or spectrum extrapolation could recover the spectrum components outside the passband, frequency domain broadening means high angular resolution in time domain. To determine the superresolution performance of iterative algorithm, we transform iterative formula (15) into the frequency domain[8].

After the convolution, band frequency components of $\bar{\sigma}^k(f)$ are spread out of the band. In single iteration, the convolution operation can extrapolate the spectrum, so the spectrum of next iteration $\bar{\sigma}^{k+1}(f)$ is greater than $\bar{\sigma}^k(f)$. After a few iterations, Algorithm gradually restores spectrum outside the pattern passband and achieve the azimuth super-resolution [9][10].

IV. SIMULATION AND REAL-DATA EXPERIMENT

A. Angular Resolution Enhancement Simulation

Iterative algorithm for radar angular super-resolution is verified in the simulation. Figure 1 top chart shows the distribution of surface scatterers, three point targets, two extended target, respectively, In this distribution, the narrowest spacing between points scatterers is approximately 0.8°, and for the extended targets is about 0.5°, The chart down shows the normalized antenna pattern, the 3dB bandwidth of main lobe is about 2.65 °.

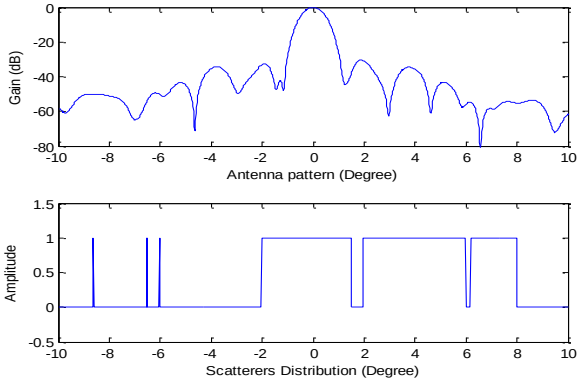


Figure 1. Scatterers distribution and idealized antenna pattern

Fig.2 show return signal and super-resolution processing result. The top chart is the return calculated with formula (1) and SNR is 15dB, resolution is very poor, point targets and extended targets are difficult to distinguish. The bottom chart shows the iterative algorithm output, the number of iterations is 50 times, compared with return, the three points distinguish was very fully distinguished, while two expanded targets separate effectively.

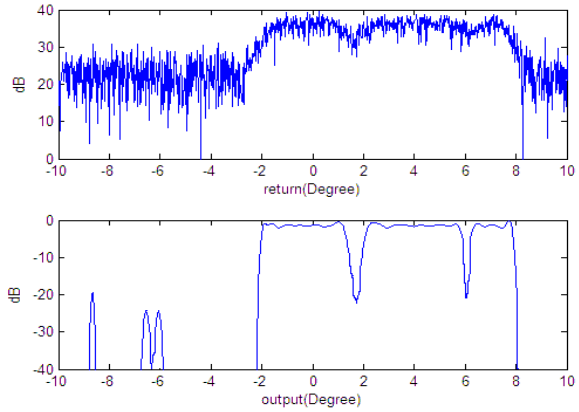


Figure 2. Return signal and superresolution processing result

B. Real-data Experiment

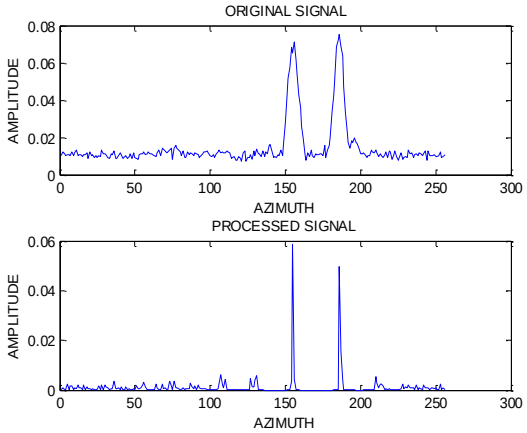


Figure 3. Range bin response in the scene

Now we turn to examine the performance of the angular resolution enhancement algorithm using radar real-data, the number of iterations is 100. Fig.3 show comparative result of signal in a range bin response in the scene. Fig.4 is the original image, azimuth resolution is very low, and it's difficult to distinguish targets.

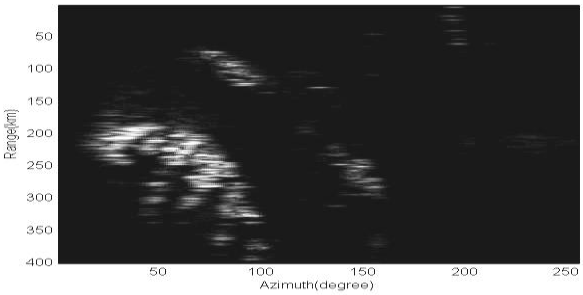


Figure 4. Real beam image

Figure 5 shows the algorithm output, the results show that the proposed approach produces sharper point target image than the original real-beam image, improvement in angular resolution is about 10 times. Also, the approach has effectively suppressed the clutter background, enhanced image features for targets.

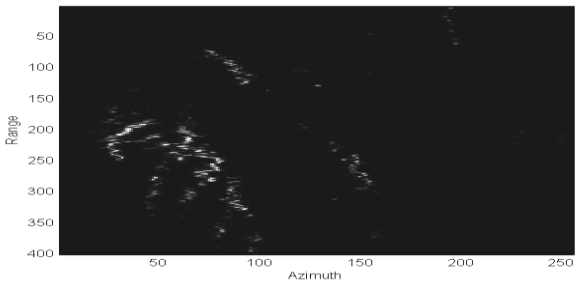


Figure 5. Angular resolution enhanced image by iterative algorithm

V. CONCLUSION

An iterative deconvolution algorithm with MAP for radar angular super-resolution has been developed and tested. Simulation results show that the algorithm effectively attenuates the effect of antenna pattern convolution, largely enhance radar azimuth resolution; real radar data experiment results demonstrate one order improvement of angular resolution compared with original image.

REFERENCES

[1] M.A Richards. Fundamentals of Radar Signal Processing .Publishing House of Electronic Industry,Beijng,2008

[2] I. G. Cumming and F. H. Wong, Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation. Norwood, MA: ArtechHouse, 2005.

[3] M.A Richards.. "Iterative Noncoherent Angular Superresolution," IEEE National Radar Conference, 3rd, Ann Arbor, MI, Apr. 20, 21, pp.100–105. 1988

[4] Yiyuan Ding, Jianyu Yang, Weihua Zhang. "A generalized inverse filtering method to improve the angular resolution of real aperture radar," Journal of Electronics, Vol 21, No 9, pp.15–19, Sep.1993

[5] D Iverson. "Beam sharpening via multikernel deconvolution," Radar, 2001 CIE International Conference on, Proceedings, pp: 693-697, 2001

[6] Yueli Li, Diannong Liang, Xiaotao Huang. "A Multi-Channel Deconvolution based on Forward-look Imaging Method in Monopulse Radar," Signal Processing, Vol 23, No 5, pp.699–703, Oct,2007

[7] Mouyan Zou. Deconvolution and Signal Recovery. National Defence Industry Press, Beijing. 2001

[8] E.S Meinel, "Origins of Linear and Nonlinear Recursive Restoration Algorithms," Journal of the Optical Society of America A, Vol.3, No.6, pp.787–799, 1986

[9] W.H Richardson. "Byesian-based Iterative Method of Image Restoration.,"Journal of the Optical Society of America, 62(1),pp:55-59, 1972

[10] E.S Meinel, "Origins of Linear and Nonlinear Recursive Restoration Algorithms," Journal of the Optical Society of America A, Vol.3, No.6, pp.787–799, 1986

Extraction of 3D Shape of Tooth with Three-Dimensional Contours of CT Images

Toshiki NAGAOKA
Graduate School of Engineering
Tohoku University
Japan
toshiki@iic.ecei.tohoku.ac.jp

Ryuichi YANAGISAWA, Yoshihiro SUGAYA, Shinichiro OMACHI
Graduate School of Engineering
Tohoku University
Japan
machi@ecei.tohoku.ac.jp

Abstract—**This paper proposes a method to extract 3D shapes of teeth automatically from dental CT images obtained from a micro-CT system. The information of 3D shapes is very helpful for dental analysis. Therefore, automatic and accurate method for extracting the shape of tooth is required. In this area, there is a problem of miss-extraction by extracting adjacent tooth or alveolar bone together with the target tooth. An existing method called Active Contour Model (ACM) which uses the information of contour of tooth tried to extract the shape of tooth. But accurate result couldn't be achieved. We propose a method using the characteristics of dental structures to achieve the accurate extraction.**

Keywords-region extraction; dental CT image; 3D shape; region growing;

I. INTRODUCTION

In recent years, benefits of images play a vital role in medical field. X-ray, MRI (Magnetic Resonance Imaging) and CT (Computer Tomography), in particular, are used to aid diagnoses extensively. The benefits derived from using images are reducing burden of medical doctor, externalizing and accurizing diagnoses, making a database and so on.

The diagnosis using images is also made in dental field, and information of the shapes of teeth achieved from the images shown in Fig.1 has great advantages.

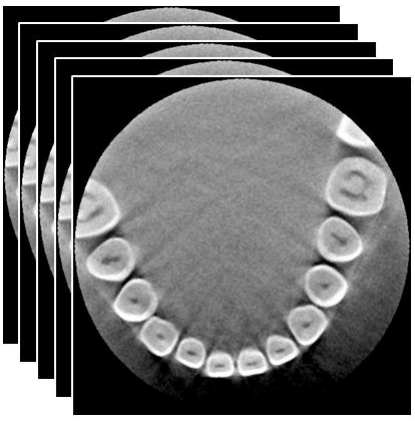


Figure 1. Example of the dental CT images.

One of the ways of extracting the shapes of teeth is tracing the contours of teeth manually in every CT slice. However, this method causes a great deal of trouble because of enhancing resolution of imaging devices recently. Thus the automatic system to extract is required. The method called Active Contour Method (ACM) using the information of teeth contours was introduced in [1]. But it couldn't achieve the accurate extraction due to over-extraction which is extracting the unnecessary structures such as adjacent teeth or alveolar bone together with the target tooth.

We introduce region growing method that considers the characteristics of dental structures for suppressing over-extraction that is the main issue of this field, and for achieving high accuracy extraction.

II. PROPOSED METHOD

The proposed method consists of some steps, which is regularization, enhancing the contrast, smoothing, making the contours and region growing.

A. Regularization

At first, we introduce regularization which is using Slice Energy (SE) [2]. It is necessary that each slice has similar brightness value in this method. SE is the sum of brightness value of all pixels in each slice, and it is calculated by (1).

$$SE = \sum_{x,y} I_z(x,y) \tag{1}$$

$I_z(x, y)$ in (1) is the brightness value in the voxel of coordinates (x, y, z). The brightness value in every slices can be regularized due to all slices have constant values of SE.

B. Enhancing the contrast

Next, contrast is enhanced. In the dental CT images, a tooth and the structures around it such as alveolar bone have a similar brightness, so miss-extraction can be occurred. In order to prevent from extracting the structure except the target tooth, we enhance the contrast by converting the brightness value as (2).

$$\hat{I}_i = \frac{I_i^2}{255} \qquad (0 \le I_i \le 255) \tag{2}$$

I_i is the brightness value of the certain voxel, and $\hat{I}_i$ is its converted value. The result of this process is shown in Fig.2.

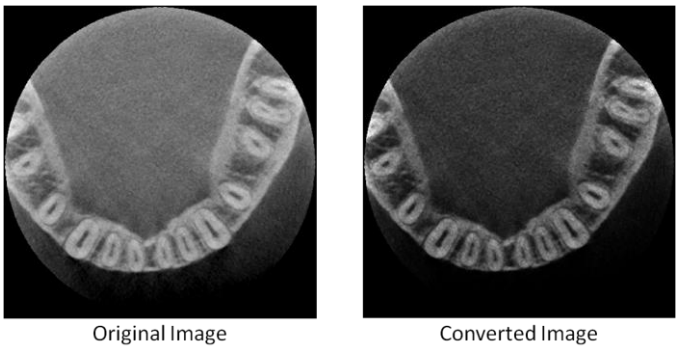


Figure 2. The result of enhancing the contrast.

C. *Smoothing*

The CT images have particular noises, and these noises often cause miss-extraction. In order to solve this problem, we should smooth the CT images before the extracting process. But on the other hand, smoothing tends to decrease the sharpness of the edge of the image, which might also cause accuracy deterioration of extraction. So we introduce bilateral filter [3] that is one of the edge-preserving smoothing filter. The brightness value output by bilateral filter is calculated by (3)

$$f_i = \frac{1}{C} \sum_{j \in N(i)} e^{\frac{-\|j-i\|^2}{2\sigma_d^2}} e^{\frac{-|I(j)-I(i)|^2}{2\sigma_r^2}} f_j \tag{3}$$

$$C = \sum_{j \in N(i)} e^{\frac{-\|j-i\|^2}{2\sigma_d^2}} e^{\frac{-|I(j)-I(i)|^2}{2\sigma_r^2}} .$$

In (3), f is the intensity of certain voxels. i is the voxel which is to be smoothed, and j is a referred voxel. σ_d is the standard deviation of space domain, and σ_r is the standard deviation of the intensity domain. $N(i)$ is a spatial neighborhood of voxel i.

D. *Making the contours*

In the process of region growing, some conditions are needed to stop growing at proper area. The information of a contour of a tooth can be used for a good condition. We make the contours by drawing the lines of certain value between the pixels stepping over certain brightness value, and this step is iterated at every brightness value. In the case of the contour of the height of 5 is shown in Fig.3.

However, we will get unnecessary contours such as one of the noise or structures around the tooth. In consideraion of it, we limit the length of contours so that we can sort out the contours of only teeth and reject unnecessary contours. The image of contour made in above way is showed in Fig.4.

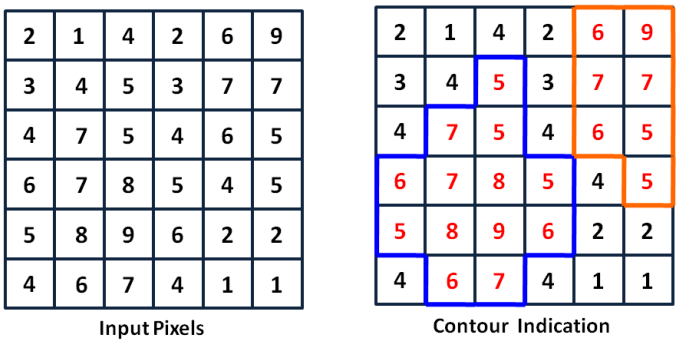


Figure 3. The case of making the contour of the height of 5.

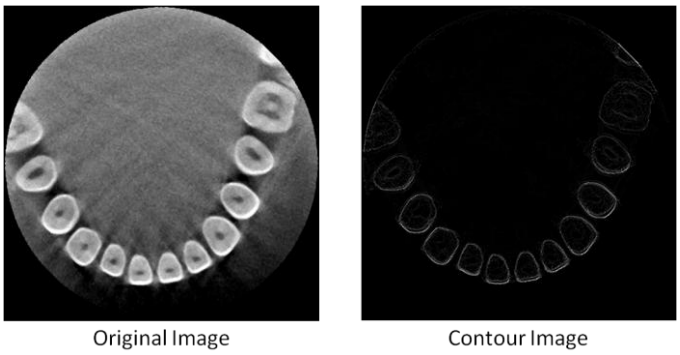


Figure 4. The result of making the contour.

However, the conoturs of terminals of the tooth might be also rejected due to the limitation that they are too short. Thus our method introduces the contours of X and Y axis direction not only Z direction. Fig.5 shows the images of X and Y contours.

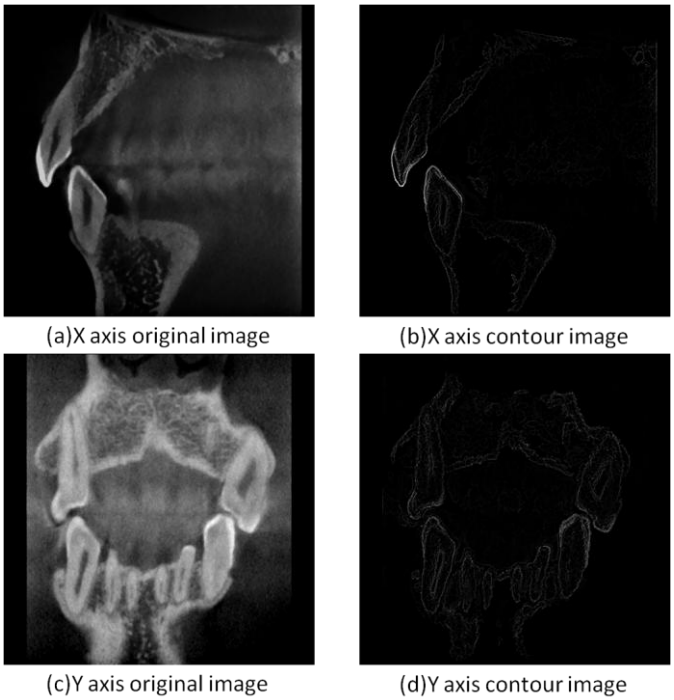


Figure 5. The contour images of X and Y axis

The contour of root is too short to be remained in Z direction contour. However, the contour of root in Z is contained in long contours in other direction contours. That means X and Y contours enable to restore the short contour at the terminal rejected in Z direction by length limitaion (Fig.6).

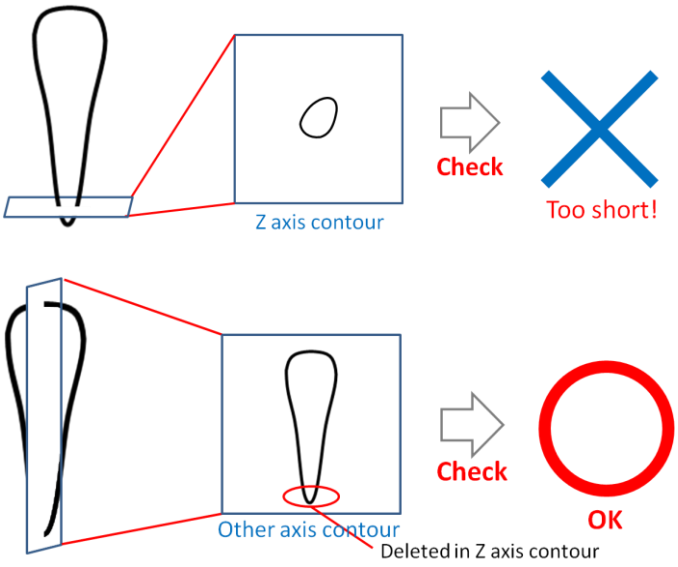


Figure 6. The diagram of other axis contour.

The way of using X and Y contour is to adopt the different contour according to the region growing direction. In region growing step, this method takes in three-dimentional adjacent voxels meeting certain conditions in sequence. We introduce the information of the contours as one of the conditions when the extracted region expands. Then, this method introduce the Y and Z axis contours in the case of expanding along X direction, Z and X axis contours in the case of along Y direction and X and Y contours in the case of along Z direction. This enables to take advantages of contours of each direction.

E. Region growing

The region of a target tooth is extracted by region growing method [4]. In this method, we have to give a start point called "seed" at first. After seed is set, then the seed takes the neighboring voxels within certain conditions one after another as an area to be extracted.

The structures we want to extract are dental pulp, dentine, and enamel. As the characteristics of dental structure, we can see the tendency that the structures near the surface of a tooth tend to have higher brightness. Considering this, we can extract the shape of the target tooth as a cluster of structures by setting the seed on the dental pulp, that has the lowest brightness value, and taking in sequence the neighboring voxels with higher brightness value until reaching the surface as shown in Fig.7.

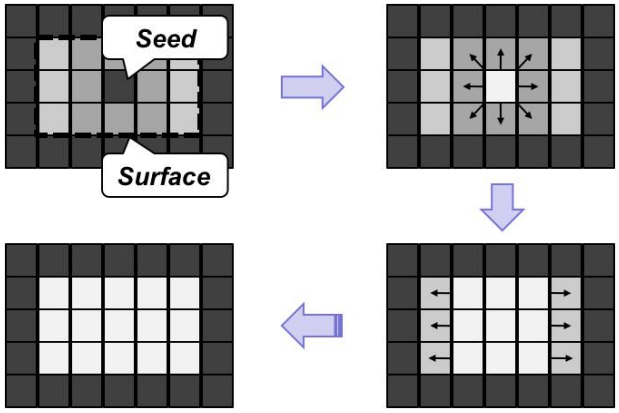


Figure 7. Flow of region growing.

In addition, the region growing method described above is applied similarly to adjacent teeth, and two or more areas which are extracted by the seeds scramble for certain voxels. When more than two regions are going to take in a same certain voxel, the region whose center of gravity is nearest to the voxel will take in it as Fig.8. This process is introduced because extracting the plural teeth simultaneously causes scramble that is mentioned above at the contact zone of teeth, and prevents the leakage to the adjacent teeth. So applying the region growing method to plural teeth achieves reducing miss-extraction at the crown of the teeth.

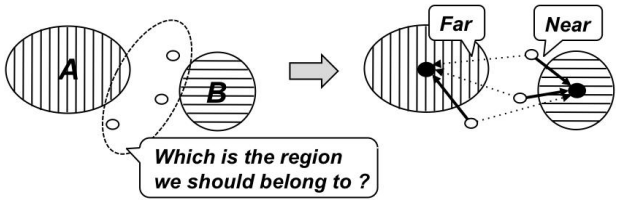


Figure 8. The region whose center of gravity is nearest

III. RESULT

The left side in Fig.9 is the result of the existing method [1], and the right side is the one of the proposed method. The target tooth is the lower right-sided lateral incisor. In the case of the existing method, over-extraction to alveolar bone was caused at the root of the tooth. On the other hand, good extraction could be achieved by the proposed method, and over-extraction was hardly caused.

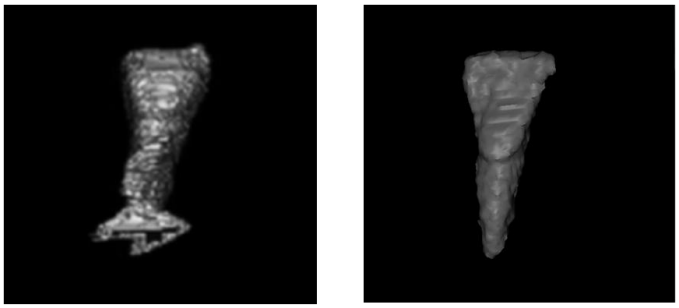


Figure 9. Comparison with existing method.

Fig.10 shows the comparison with the manually extracted teeth. The images of the left side are the results of manual extraction, and that of the right side are the results of the proposed method. In the figure, lower right-sided central incisor as Fig.10(b) and lower left-sided lateral incisor as Fig.10(d) seem to be good extractions. On the other hand, miss-extraction occurred in lower left-sided canine as Fig.10(f).

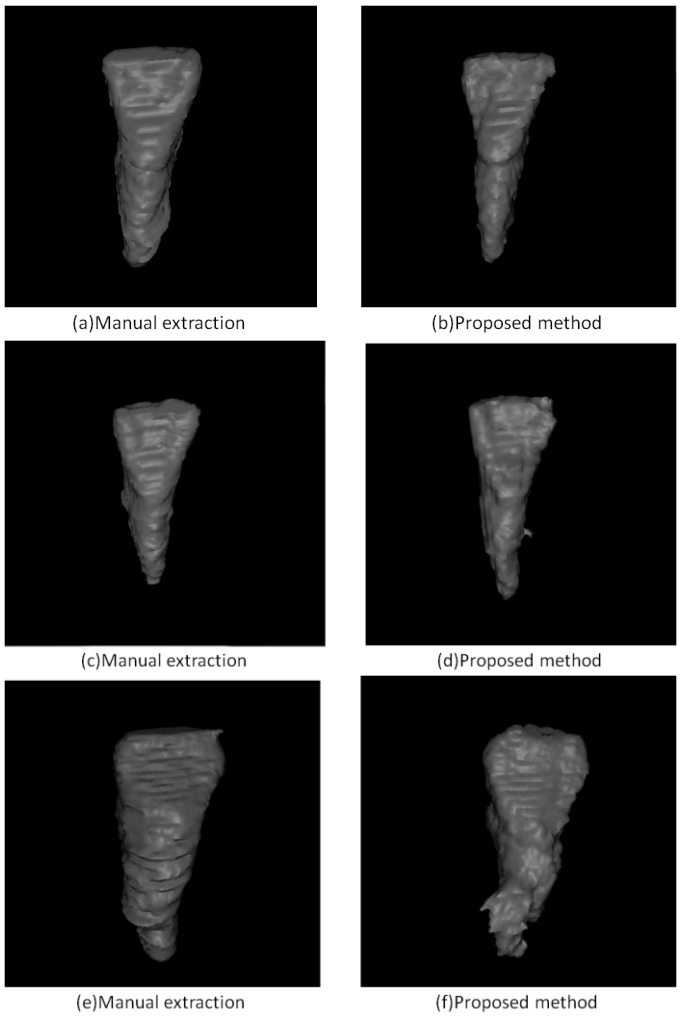


Figure 10. The results of (a)(b) lower right-sided central incisor, (c)(d) lower lefr-sided lateral incisor, (e)(f) lower left-sided canine.

Next we evaluate the accuracy of extraction by the proposed method quantitatively. In this paper, we introduce the value of precision and recall. Precision evaluates the suppression of un-extraction, and recall evaluates the suppression of over-extraction. Each value is calculated as shown in Fig.11.

We applied our proposed method to six teeth, which are lower central incisors, lower lateral incisors and lower canines. The result of quantitative evaluation is shown in Table 1. We achieved high value of precision, but we have to improve the value of recall.

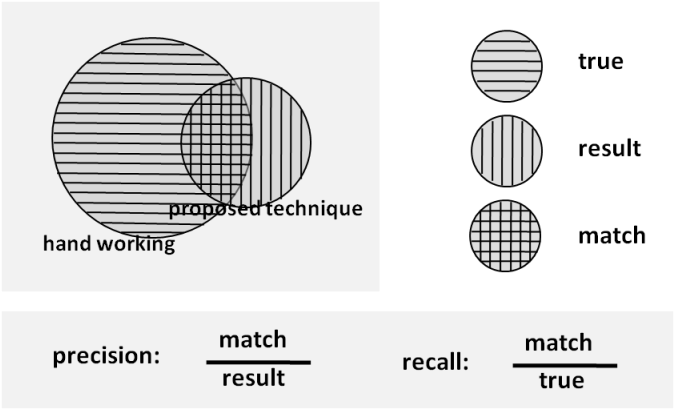


Figure 11. Illustlation of the evaluation values.

TABLE I. THE RESULT OF QUANTITATIVE EVALUATION

Candidate		Precision	Recall
Right-side	canine	0.962	0.737
	lateral incisor	0.973	0.842
	central incisor	0.979	0.774
Left-side	central incisor	0.969	0.780
	lateral incisor	0.968	0.806
	canine	0.985	0.769
	Average	***0.973***	***0.785***

The result shows that we could suppress over-extraction in our method, but un-extracted voxels were still remained. That is to say we couldn't extract the region that have to be extracted. So we should consider the problem of using X and Y axis contour. It is sure that the contour of a terminal of the tooth is achieved well by using not only Z direction contour but also X and Y contour, but the contours of not teeth such as noises also increase. The unnecessary contours disturb the region growing. Then we have to consider the way of making the contour of X and Y direction.

IV. CONCLUSION

In this paper, we proposed a method of extracting 3D shapes of teeth from dental CT images. This method consists of some stages. In contour making step, the X and Y direction contours are used to achieve the contours of a terminal of the tooth which is rejected in the contour of Z direction by limitation of contour length to reject noises.

The shapes of six teeth were extracted by our proposed method. Over-extraction to adjacent teeth or alveolar bone was suppressed. However, un-extracted region, that has to be extracted, was remained. So we must make accurate contours of X and Y direction to achieve better extraction. Applying the proposed method to other teeth is an important future work.

REFERENCES

[1] S. Omachi, K. Saito, H. Aso, S. Kasahara, S. Yamada, and K. Kimura, "Semi-Automatic Reconstruction of Tooth Shape from CT Images by Contour Propagation," IEICE Transaction on Information and Systems, vol. J91-D, no.9, pp2426–2429, 2008.

[2] T. Grenier, C. R. Muller, N. Costes, M. Janier. G. Gimenez, "Automated Seeds Location for Whole Body NaF PET Segmentation," IEEE Transaction On Nuclear Science, vol. 52, no. 8, pp1401–1405, Oct. 2005.

[3] C. Tomasi, R. Manduchi, "Bilateral filtering for gray and color images, "in Proc. Int. Conf. Computer Vision, pp839–846, 1998.

[4] R. Adams, L. Bischof, "Seeded region growing," IEEE Transaction on Pattern Analysis And Machine Intelligence, vol. 16, pp. 641-647, 1994.

Pattern-Search-based Optimized PM signal design

Liang Li, Lingjiang Kong*, and Xiaobo Yang
School of Electronic Engineering
University of Electronic Science and Technology of China
Chengdu, China
*Corresponding author E-mail: lingjiang.kong@gmail.com

***Abstract*—In this paper, a radar system capable of adaptively adjust its transmitted waveform is considered. In which a mechanism that dynamically mitigate the effect of the environment is taken so as to improve the detection performance. Such method is particularly suitable for low speed target detection in compound-Gaussian (CG) sea clutter environment. As a key feature of above adaptable structure under the framework of general likelihood ration test (GLRT), phase-modulate (PM) signal design problem is exploited, analyzed and transferred analytically into a simple mathematic representation. Further more, a pattern search based algorism addressed on this optimization is designed and evaluated.**

Keywords- waveform diversity; PM signal; detection; pattern search;

I. INTRODUCTION

The problem of ocean surface small target detection is challenging due to the heavy clutter background. Such a scenario is typical when radar operates with low grazing angles over the sea. Improvement in detection performance is demanding and can be achieved by taking pulse-to-pulse waveform agility that allows radar to perform a certain level of scenario cognition and adaptation [1] [2].

In this work, we follow the dynamic waveform design method [1] which aims to design the waveform in pursuit of mitigating the side lobe interference from clutter at a specific range bin, so as to increase SCR associated with weak target and improve its detection performance. In this method, the location of potential target range bin and clutter power distribution over range bins near by is estimated and an optimized phase-modulated (PM) waveform is designed accordingly. The generalized likelihood ratio test (GLRT) for the target [2] is then expressed as a function of phase modulate parameters of the waveform and estimated clutter statistics. Maximization of the GLRT criterion has been taken into optimize waveform synthesis process.

Conventional waveform synthesis method in [3] provides a basic framework that reveals a complex relation between waveform parameter and target function. Simplified and altered in current synthesis problem, a series of expressions that encompass a standard form of multi-dimension optimization problem are acquired.

Algorisms such as exhausted search [9], linear programming (LP) [10], gradient descent [11], Newton-Raphson method [2][8] and pattern search [4] etc. are generally applicable to multi-dimension optimization problems. In this case, the LP, gradient descent and Newton-Raphson methods are all computationally costly than patter search, and they require complex measurements to ensure convergence.

Pattern search method, which is also known as direct-search or derivative-free method, operates in a quite simpler way. General pattern search varies one parameter at a time by steps of the same magnitude, and when no such increase or decrease in any one parameter further improved the solution, they halved the step size and repeated the process until the steps were deemed sufficiently small.

A pattern search based waveform parameter optimize algorism has been designed and evaluated.

II. SYSTEM AND SIGNAL MODELS

Consider a two-stage procedure of transmitting waveform to realize an improved detection performance. The first stage transmits a linear frequency modulated (LFM) chirp pulse train to estimate clutter statistics and possible target position, while in the second stage, PM waveform is designed based those estimation and such a pulse train is transmitted to yield optimized detection.

The procedure assumes a low speed Swerling I target that has reflectivity $\boldsymbol{b}=[b^0,b^1,\ldots b^{K-1}]$ with respect to K pulses, which follows a complex Gaussian distribution $\boldsymbol{b}\sim CN(0,\sigma^2\mathrm{I}_K)$. In each stage, series of pulses are transmitted in very short duration so that Doppler effects and processing can be ignored. Further more, it is also assumed clutter power is significant higher than that of addictive noise, there for the effect of addictive noise is ignored. After sampling and matched-filtering of the received signal, the vector from matched-filter according to jth range bin is given by

$$r_j = \boldsymbol{b} z_s[j-n_0] + \sum_{n=-(N_s-1)}^{N_s-1} z_s[n] y_{j+n} \qquad (1)$$

Where n_0 is the rage bin index that target is located, $z_s[n]$, $|n|<N_s$ is the autocorrelation function of transmitted signal with length of N_s.

Compound-Gaussian model [5] is used in the procedure. Clutter statistics are estimated in stage 1 of the procedure. Suppose a target is predicted in range bin j, then it has to estimate

This work is supported by the Fundamental Research Funds for the Central Universities No.E022050205.

$$\Theta = \{T_{j-2(N_s-1)}, ..., T_j, ..., T_{j+2(N_s-1)}, \Sigma\} \quad (2)$$

Where Σ is speckle covariance matrix, and T_x are textures over range bins. From (2), the covariance matrix of the matched-filter output is

$$R_j = E\{r_j r_j^H\} = \sigma^2 \mathrm{I}_k \mid z_s[j-n_0]\mid^2 + \Sigma\beta_j \quad (3)$$

Where $\beta_j = \sum_{n=-(N_s-1)}^{N_s-1} T_{j+n} \mid z_s[n]\mid^2$ represents total clutter power of the matched-filter output at range bin j, is contributed by both clutter reflection at range bin j and reflections of near ranges that caused by range side-lobe.

The maximum likelihood estimate can be found using EM algorism [6]. After the estimating of (2), a subspace-based clutter suppression method [7] has been taken. As assumed that clutter concentrates most of its energy in a lower-rank subspace. By means of projection on a lower dimension subspace matrix $Q_c^{\perp}$ with rank of K', whose columns corresponds to the small eigenvalues of estimated speckle covariance matrix $\hat{\Sigma}$, thus there has

$$r_j^{\perp} = Q_c^{\perp H} r_j \quad (4)$$

Then the clutter-suppressed version of received signal is obtained.

III. GLRT Detector

With respect to a given range bin, hypotheses H_0 which indicate the presence of only clutter while H_1 indicate the presence of both clutter and target are defined. The GLRT is formed [15]

$$\Lambda_j^{GLRT} = \ln \frac{p(r_j^{\perp} \mid H_1, \hat{T}, \hat{\Sigma}, \hat{\sigma}^2)}{p(r_j^{\perp} \mid H_0, \hat{T}, \hat{\Sigma})} \gtrless \gamma' \quad (5)$$

Under hypotheses H_0 and H_1, the covariance matrixes of $r_j^{\perp}$ are

$$\begin{cases} C_{\omega} = \hat{\beta}_j \hat{Q}_c^{\perp H} \hat{\Sigma} \hat{Q}_c^{\perp} \\ C_{s\omega} = \hat{\sigma}^2 \mathrm{I}_{K'} + \hat{\beta}_j \hat{Q}_c^{\perp H} \hat{\Sigma} \hat{Q}_c^{\perp} \end{cases} \quad (6)$$

Considering of both probability density functions in (6) are complex Gaussian with zero means, the GLRT can be reduced to

$$T(r_j^{\perp}) = r_j^{\perp H} C_{\omega}^{-1} \hat{\sigma}^2 C_{s\omega}^{-1} r_j^{\perp} \gtrless \gamma \quad (7)$$

Apply Eigen-decomposition on $C_{\omega}=V_{\omega}\Lambda_{\omega}V_{\omega}^H$, where V_{ω} is unitary matrix and Λ_{ω} is diagonal matrix of Eigen values $\lambda_{\omega}^{(n)}$, n=0,…,K'–1, thus (7) can be written as

$$T(\tilde{r}_j^{\perp}) = \tilde{r}_j^{\perp H} \hat{\sigma}^2 \Lambda_{\omega}^{-1} (\mathrm{I}_{K'} + \hat{\sigma}^2 \Lambda_{\omega}^{-1}) \tilde{r}_j^{\perp} \gtrless \gamma$$

Where $\tilde{r}_j^{\perp} = \Lambda_{\omega}^{-1/2} V_{\omega}^H r_j^{\perp}$, then

$$T(\tilde{r}_j^{\perp}) = \sum_{n=0}^{K'-1} \frac{\hat{\sigma}^2}{\hat{\sigma}^2 \hat{\beta}_j + \lambda_{\omega}^{(n)} \hat{\beta}_j^2} \mid (\lambda_{\omega}^{(n)})^{-1/2} r_j^{\perp}[n] \mid^2 \quad (8)$$

IV. Optimal Signal Design

By the end of Sub-dwell 1, environment parameter Θ is estimated, and a potential target position for further detection in sub-dwell 2 is provided. According to (8), an optimized transmit signal can be designed to maximizing Λ_j^{GLRT}, which is equivalent to minimize $\hat{\beta}_j$.

The waveform of transmitted signal in sub-dwell 2 is a phase-modulated waveform

$$s_2(t) = \exp(j\psi(t)), 0 \le t \le T_s \quad (9)$$

Where the phase modulation is applied in equal length of segments as

$$\psi(t) = \mu_i, (i-1)\Delta T \le t \le i\Delta T \quad (10)$$

Where T_s and ΔT is the pulse duration and modulation interval, and $\boldsymbol{\mu}=[\mu_1, \mu_2, \ldots, \mu_{Ns}]^T$ is phase vector of phase modulation. The autocorrelation function of transmitted signal is then represented as

$$z_s[n] = \frac{1}{N_s} \sum_{i=1}^{N_s-n} \exp(j(\mu_{n+i} - \mu_i)) \quad (11)$$

Which is clear that both $z_s[0]=1$ and $z_s[-n]= z_s[n]^*$ holds for all n. There for the goal of optimal signal design is

$$\arg\min_{\boldsymbol{\mu}} \hat{\beta}_j \quad (12)$$

$$\hat{\beta}_j = T_0 + \sum_{n=1}^{N_s-1}(T_n + T_{-n})\,|\,c_0(n)\,|^2 + \frac{1}{N_s^2}\sum_{n=1}^{\min(j-1,N_s-j)}(T_n + T_{-n})2\cos(2\mu_j - \mu_{j-n} - \mu_{j+n})$$
$$+\frac{1}{N_s}\sum_{n=1}^{j-1}(T_n + T_{-n})[\frac{1}{N_s} + 2\,\mathrm{Re}(c_0(n))\cos(\mu_j - \mu_{j-n}) + 2\,\mathrm{Im}(c_0(n))\sin(\mu_j - \mu_{j-n})] \quad (13)$$
$$+\frac{1}{N_s}\sum_{n=1}^{N-j}(T_n + T_{-n})[\frac{1}{N_s} + 2\,\mathrm{Re}(c_0(n))\cos(\mu_{j+n} - \mu_j) + 2\,\mathrm{Im}(c_0(n))\sin(\mu_{j+n} - \mu_j)]$$

Define $r(i,j) = \frac{1}{N_s}e^{j(\mu_i - \mu_j)}$, and let $c_0(n) = \sum_{\substack{i=1 \\ i\neq j-n,\,j}}^{N_s-n} r(n+i,i)$.

Expand (11), it can be expressed as (13).

It is clear that as a function of μ_j, $\hat{\beta}_j$ has only 5 non-zero Fourier coefficients in its Fourier series, it is periodic and continues and has derivations of any order. Simple as it is however, no analytic solution for its minima over 0~2π can be found, nor dose it fit well with quadric function which is used for Newton method of root finding.

Recall for (12), it is required for an algorism to find a vector $\boldsymbol{\mu}=[\mu_1, \mu_2,\ldots,\mu_{Ns}]^T$, in which each of μ_j should be given exact value within a specific precision (exp. $2\pi/2^8$). An ideal algorism or optimize method should be able to find the global optimized point of N_s dimension hyper space, while in practice ,there are so many local optimum with near-optimal solutions. There for, a local search approach is suitable for the problem (12).

We develop a pattern search based optimize method as follows:

1) Start iteration form $\boldsymbol{\mu}$ directly inherit from LFM waveform, set search step-size to $2\pi/16$, that is 16 equally spaced points range from 0 to 2π.

2) Iteratively take rounds of calculations that update μ_i for every *i*, and progressively shrink the step-size until it reaches the predefined precision.

3) Continue iteration with step-size equal to the predefined precision, shrink the search range according to the maximal changes of μ_i of a round, until it reaches 0.

In process 2 and 3, on μ_i for every *i*, (13) is used in calculation, and value of μ_i that correspond to the smallest $\hat{\beta}_j$ is chosen to be the new value of μ_i, and continue with next *i* until it reaches N_s.

Such an algorism guaranties convergence in this problem [12], though it does not necessarily yields the global optimal.

V. SIMULATION

The simulation is set in order to evaluate the solving of (12) with proposed algorism. Look into (13) where texture estimates $\hat{T}$ among so many parameters in the model are of only concern. For the simulation, the estimation result of texture follows a gamma distribution

$$p(u) = \frac{2b^{2v}}{\Gamma(v)}u^{2v-1}\exp(-b^2u^2)$$

Where *b* is scale parameter and *v* is shape parameter. In this simulation it use *b*=1 and *v*=0.4, which corresponds to a highly non-Gaussian clutter. A typical sample of clutter power over 256 range cells are shown in figure 1.

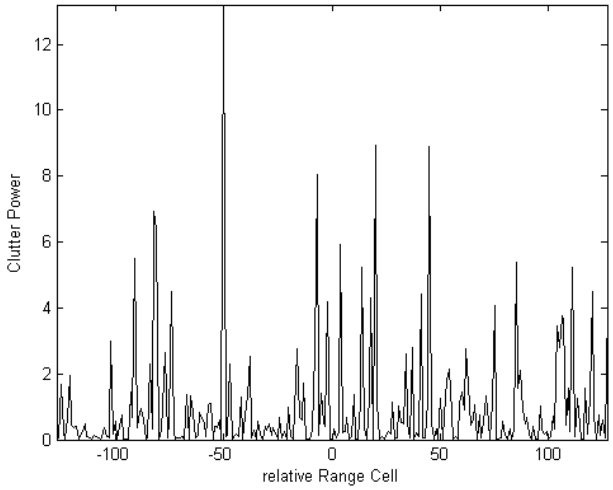


Figure 1. Clutter distribution over range cells

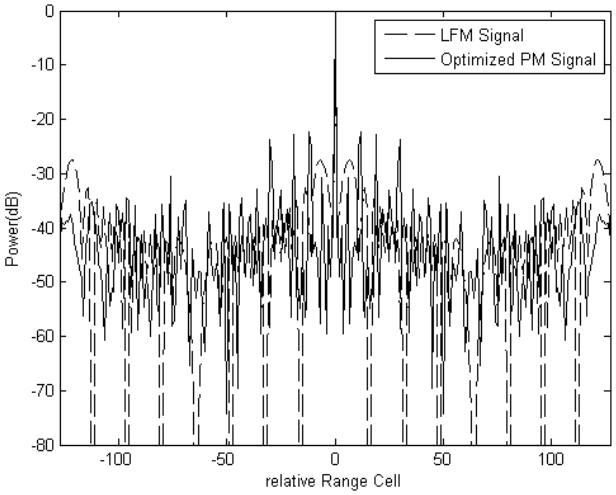


Figure 2. Comparison of the magnitude of the autocorrelation of the designed PM (solid line) with that of the LFM (dashed line)

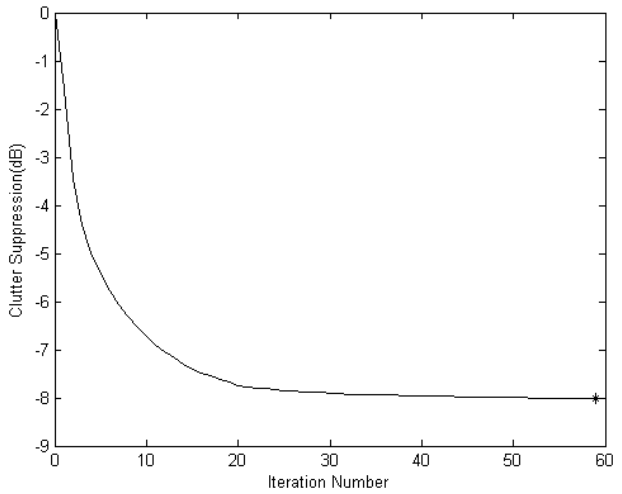


Figure 3. Clutter suppression compared with LFM signal vs. iteration number

Suppose a target is located at 0th of the range cell on figure 1, and use N_s=128 in (9).The 128 dimension $\boldsymbol{\mu}=[\mu_1, \mu_2,\ldots,\mu_{Ns}]^T$ is meant to be optimized to satisfy (12) with accuracy of $2\pi/2^8$. In another word, an alternate autocorrelation function of a signal is obtained that makes $\hat{\beta}_j$ smaller. This is done by the method described in section IV, and the autocorrelation functions of both LFM signal and optimized PM signal ($|z_{s_2}[n]|^2$) are shown in figure 2.

It can be observed that the value of $|z_{s_2}[n]|^2$ has many concaves in comparison with that of LFM's, and those concave are generally aligned with range bins where large clutter power present. Also there are many side-lobes of $|z_{s_2}[n]|^2$ have higher magnitude than that of LFM's. These side-lobes correspond to range bins that have relatively lower power of clutter.

The iteration progressively optimizes the PM parameter $\boldsymbol{\mu}$, and this progress can be seen from figure 3, which shows the value of $\hat{\beta}_j$ with current PM signal compares with that corresponds to LFM signal. It also represent just how much of the clutter power that comes from range side-lobes has been reduced with optimized signal, compared with that LFM signal can produce.

VI. CONCLUSION

The problem of detecting low speed small target in sea clutter has been stated; a clutter mitigation mechanism has been further observed and analyzed. On the design of the optimized phase-modulated waveform that maximized the GLRT, a pattern search based algorism is proposed. Such an algorism is demonstrated and evaluated via a simulation, and the results show effectiveness of the algorism. The results also suggest significant performance improvement by clutter mitigation mechanism with optimized PM signal in the ascribed detection scenario can be attained.

REFERENCES

[1] Sira, S.P.; Cochran, D.; Papandreou-Suppappola, A.; Morrell, D.; Moran, W.; Howard, S.; 'A SUBSPACE-BASED APPROACH TO SEA CLUTTER SUPPRESSION FOR IMPROVED TARGET DETECTION' ACSSC '06. Fortieth Asilomar Conference on Signals, Systems and Computers, 2006. Oct. 29 2006-Nov. 1 2006 pp.752 - 756

[2] Sira, S.P.; Cochran, D.; Papandreou-Suppappola, A.; Morrell, D.; Moran, W.; Howard, S.D.; Calderbank, R.; "Adaptive Waveform Design for Improved Detection of Low-RCS Targets in Heavy Sea Clutter" IEEE Journal of Selected Topics in Signal Processing, Volume 1, Issue 1, June 2007 pp.56 – 66

[3] J. D. Wolf, G. M. Lee, and C. E. Suyo, "Radar waveform synthesis by mean-square optimization techniques," IEEE J. Oceanic Eng., vol. AES-5, no. 4, pp. 611–619, Jul. 1969.

[4] Moharir.P.S, Singh. R. and Maru .,"S-K-H algorithm for signal design", Electronic letter, vol 32, no 18, pp.1642-1649, Aug 1996.

[5] K. Ward, C. Baker, and S. Watts, "Maritime surveillance radar Part I: Radar scattering from the ocean surface," IEE Proceedings F: Communications, Radar and Signal Processing, vol. 137, pp. 51–62, April 1990.

[6] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood estimation from incomplete data via the EM algorithm," Journalof the Royal Statistical Society, Series B, pp. 1–38, 1977.

[7] Li, Y.; Sira, S. P.; Papandreou-Suppappola, A.; Cochran, D.; Scharf, L. L.; 'MAXIMIZING DETECTION PERFORMANCE WITH WAVEFORM DESIGN FOR SENSING IN HEAVY SEA CLUTTER' IEEE/SP 14th Workshop on Statistical Signal Processing, 2007. SSP '07. 2007 , pp. 249 - 253

[8] J. D. Wolf, G. M. Lee, and C. E. Suyo, "Radar waveform synthesis by mean-square optimization techniques," IEEE J. Oceanic Eng., vol. AES-5, no. 4, pp. 611–619, Jul. 1969.

[9] Davis, R.M.; Fante, R.L.; Perry, R.P. 'Phase-coded waveforms for radar', Aerospace and Electronic Systems, IEEE Transactions on Volume: 43 , Issue: 1, 2007 , pp. 401 – 408

[10] Hankang Wang; Hong Gu; Weimin Su; Jianfei Wang, 'Sidelobe Suppression of Phase-Coded Radar Based on the Accurate Signal Estimate', Innovative Computing Information and Control, 2008. ICICIC '08. 3rd International Conference on, 2008, pp. 535 – 535

[11] Jan A. Snyman. Practical Mathematical Optimization: An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms. Springer Publishing. 2005.

[12] Torczon, V.J. "On the convergence of pattern search algorithms". SIAM Journal on Optimization 7 (1): pp. 1-25. 1997.

A Time-domain Sparse Signal Representation Perspective for Speech Denoising

Li Jinyang, Qun Wan
Department of Electronic Engineering
University of Electronic Science and Technology of China
Chengdu, China
iLocalisation@gmail.com

Abstract— **Our work presents a fact that by taking advantage of the sparseness of speech signals in the time domain, sparse representation methodology shows promise in coping with the difficult problem of denoising in time domain directly. We enforce sparsity by imposing penalties based on the l_1 -norm. Explicitly enforcing the sparsity of the representation is motivated by a desire to obtain a noise-reduced estimate of the time sequence that only utterances are reserved. Frame-by-frame basis operation is induced to improve its performance. Experimental results from real data examples show that the proposed algorithms achieve some improvement on quality assessments.**

Keywords- Signal Representation; Speech Denoising; Sparse; Time-domain; Frame-by-frame .

I. INTRODUCTION

Estimating a signal from a noise-corrupted observation of the signal, known as denoising, is one of the fundamental and extensive problems in signal problem. Over past several decades much research has been focused on this area.

There have been many techniques proposed for denoising of speech signal corrupted by an additive noise, which is independent of speech signal. The general algorithms of speech denoising are designed to remove additive background noise for improving speech quality. In our discussion, background noise will refer to any additive broadband noise component include white Gaussian noise, aircraft cockpit noise, or machine noise in a factory environment.

Some techniques based on stochastic process models of speech rely on a given mathematical criterion, while others based on perceptual aspects of speech attempt to improve aspects important in human perception. Based on speech modeling using iterative methods, Wiener filtering [1] performs estimating model parameters that characterize the speech signal, followed by filtering of noisy speech signal by using a filter derived based on the minimum mean-square error criterion and resynthesis of the noise-free signal. It requires a priori knowledge of both noise and speech statistics. Spectral subtraction (SS) [2] ,which has still been popular due to its computational efficiency [3], suppresses noise by subtracting an estimated noise bias in the power spectral, Fourier transform, or autocorrelation domain simply. Applying Independent Component Analysis (ICA) to a large ensemble of clean speech frames, sparse code shrinkage (SCS) [4], as well as wavelet shrinkage [5-6], projects the frames composing a noisy speech signal on some kind of basis, facilitates the application of Bayesian denoising to each of the resulting independent components individually.

Basically, all these methods are to transform signals into some domain to obtain the main components related to the clean original signals subtract the slight components related to the noises and reconstruct the noise-reduced signals. Signal representations corresponding to invertible transforms such as the DFT, the discrete cosine transform (DCT), or the discrete wavelets transform (DWT) [7] are convenient and easy to calculate . However, it is difficult to determine many alternative representations for specific signals.

Over recent years there has been growing interest in finding ways to transform signals into sparse representations [8], i.e., representations where most coefficients are zero. These sparse representations are proving to be a particularly interesting and powerful tool for analysis and processing of audio signals. Finding a sparse representation for a signal has many advantages for applications such as coding, enhancement, or source separation[9]. The large parts of human speech signal are periods of silence among the small parts of utterances, which attributes the sparsity in time-domain. By taking advantage of the sparseness of speech signals in the time-domain, Time-domain Sparse Representations Denoising (TdSRD) show promise in coping with the difficult problem of denoising, but, to my knowledge, its potential for speech enhancement has not yet been exploited.

The goal of this paper is to explore how to utilize the time-domain sparse signal representation methodology for practical speech signal denoising. The main contributions of our paper include explore the possible of apply sparse signal representation in time-domain process and introduce the frame-by-frame basis operation, which make the algorithm tractable and stable. Finally, the selection of regularize parameter involved in our approach, which balances data-fidelity with sparsity, is discussed. In our experiments, the proposed approach exhibits some advantages on quality assessment over other denoising techniques. Another advantage of the approach is its flexibility, since few assumptions are made in the formulation, e.g., a *priori* knowledge of both noise and speech

statistics is not required. Similarly, extensions to many scenarios, such as other signals and non-Gaussian noise, can be readily made.

The paper is organized as follows. The general Time-domain Sparse Representations Denoising and its frame-based version are presented in Section II. The selection of regularize parameter are discussed in Section III. Section IV presents experimental evaluation and results, and Section V gives conclusions.

II. TIME-DOMAIN SPARSE REPRESENTATION DENOISING

A. SPARSE SIGNAL REPRESENTATION

The simplest version of the sparse representation problem without noise is to find a sparse $\mathbf{x} \in C^N$, given $\mathbf{y} \in C^M$, which are related by $\mathbf{y} = \mathbf{A}\mathbf{x}$,with $M < N$ .Assume the matrix $\mathbf{A}$ be known. The assumption of sparsity of is crucial since the problem is ill-posed without it .The ideal measure of sparsity is the number of nonzero entries $\mathbf{x}$,which donated by $\|\mathbf{x}\|_0^0$, so-called l_0 -norm. Looking for argmin $\|\mathbf{x}\|_0^0$, however, is a difficult combinatorial optimization problem and is intractable for even moderately sized problems. It has been shown that in some case[10-11], the l_1 relaxation $\|\mathbf{x}\|_1^1$ can replace $\|\mathbf{x}\|_0^0$ [12]. Min $\|\mathbf{x}\|_1^1$ subject to $\mathbf{y} = \mathbf{A}\mathbf{x}$ is a convex optimization problem, and the global optimum can be found for real-valued data by linear programming.

In practice, a noiseless measurement model is rarely appropriate; therefore, a sparse represen-tation problem with additive Gaussian noise takes the following form:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \tag{1}$$

To extend l_1 -penalization to the noisy case, an appropriate choice of an optimization criterion is

$$\begin{aligned} &\min \quad \|\hat{\mathbf{x}}\|_1 \\ &\textit{subject to } \|\mathbf{y} - \mathbf{A}\hat{\mathbf{x}}\|_2^2 \le \beta^2 \end{aligned} \tag{2}$$

where is β a parameter specifying how much noise we wish to allow.

B. Time-Domain Sparse Representation Denosing

Let us consider a typical inverse problem. Denote by *x* the original clear signal, and by *v* some kind of additive noise. Assume that the random variable $\mathbf{y}$ is the noisy version of signal $\mathbf{x}$

$$\mathbf{y} = \mathbf{x} + \mathbf{v} \tag{3}$$

We are interested in estimating $\hat{\mathbf{x}}$ from the observed noisy signal $\mathbf{y}$ by means of $\hat{\mathbf{x}} = g(\mathbf{y})$. Compare to equation(1), equation(3) has a difference that $M = N$ and $\mathbf{A} = \mathbf{I}_N$ .Here $\mathbf{I}_N$ is a $N \times N$ identity matrix. Correspondingly, the estimation $\hat{\mathbf{x}}$ can be obtained by l_1 -norm of the time-domain sparsity

$$\begin{aligned} &\min \quad \|\hat{\mathbf{x}}\|_1 \\ &\textit{subject to } \|\mathbf{y} - \hat{\mathbf{x}}\|_2^2 \le \beta^2 \end{aligned} \tag{4}$$

where the regularization parameter β^2 is a critical parameter related to the fit of the solution to the data. The selection of the regularization parameter will be discuss in Section III.

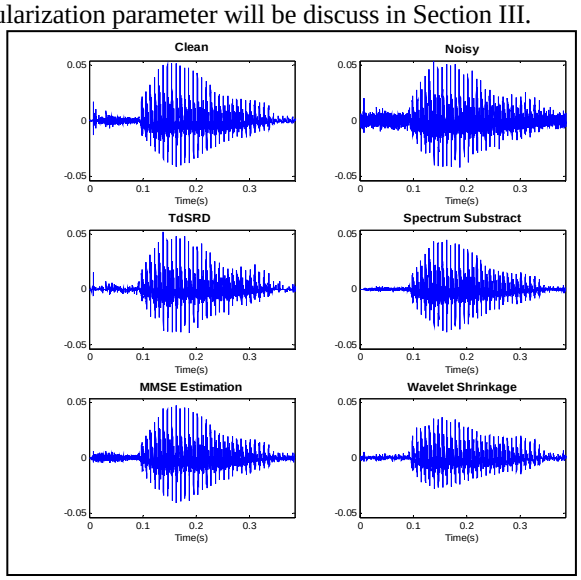


Figure 1. Waveforms of.orignal and denoised signal at SNR = 20

C. TdSRD on Frame-by-frame bais

Given the short-term, usually 10ms-30ms,stationary of speech signal and the fact that the regularization parameter is related to the energy of the noise, the frame-by-frame basis operation is considered, which split the entire *single-channel* noisy speech signal $\mathbf{y}$ into short-term segments or *frame*, each consisting of N samples, and deonising on frame basis. This gives a set of N-dimensional observation vectors $\mathbf{y} = \{\mathbf{y}(t_1),...,\mathbf{y}(t_n)\}$ where $\mathbf{y}(t_n) = [y(t_n), y(t_n+1),..., y(t_n+N-1)]$, t_n donates the start time of each frame. Corresponding notation is used for clean speech signal. According to(4), for large number of frames, the estimation $\hat{\mathbf{x}} = \{\hat{\mathbf{x}}(t_1),...,\hat{\mathbf{x}}(t_n))\}$ can be obtained by

$$\begin{aligned} &\min \quad \|\hat{\mathbf{x}}\|_1 \\ &\textit{subject to} \\ &\quad \|\mathbf{y}(t_1) - \hat{\mathbf{x}}(t_1)\|_2^2 \le \beta_1^2 \\ &\quad \vdots \\ &\quad \|\mathbf{y}(t_n) - \hat{\mathbf{x}}(t_n)\|_2^2 \le \beta_n^2 \end{aligned} \tag{5}$$

where β_1^2,…, β_n^2 are the regularization parameters for relative frames. In the following derivations, for convenience, the time index t_n will be omitted and without causing any confusion.

III. REGULARIZATION PARAMETER SELECTION

An important part of this deonising framework is the choice of the regularization parameter in equation(4), which balances the fit of the solution to the data versus the sparsity prior. The same question arises in many practical inverse problems and is difficult to answer in many cases, especially if the objective function is not quadratic. Let us denote the estimate obtained using as β^2 the regularization parameter by $\hat{\mathbf{s}}(\beta^2)$. A well-known idea under the name of discrepancy principle [13] is to select β to match $\hat{\mathbf{s}}(\beta^2)$ the residuals of the solution to some known statistics of the noise when such are available. For example, if the distribution of the noise is known or can be modeled, then one can select β^2 such that $\|\mathbf{y}(t_n)-\hat{\mathbf{s}}(t_n)\|_2^2 \approx E\left[\|n(t_n)\|_2^2\right]$. With the knowledge of the distribution, we can find a confidence interval $\|n\|_2^2$ for and use its upper value as a choice for β^2 .In simulations we present in Section IV, we find that this procedure generates appropriate regularization parameter choices for our problem when noise is reasonably small.

When noise statistics are not known, and no knowledge of the number of sources is available, the choice of the regularization parameter is a difficult question. Directly searching for a optimal value of β^2 is a difficult but effective way that requires solving equation(4) multiple times for different β^2 s. Another way has been approached in the inverse problem community by methods such as L-curve [14].

IV. EXPERIMENTAL RESULTS

In this section, we present several experimental results for our time-domain sparse representation deonising (TdSSD) scheme. First, we compare the quality assessments of our method to those of classical algorithms under various conditions. Next, we discuss and present results on regularization parameter selection. Finally, in Section V, we present an extension of our framework to general signal scenario and demonstrate its effectiveness on a number of examples.

A. Quality Assessment

Generally speaking, quality assessment techniques fall into two classes: subjective quality measures and objective quality measures. Subjective measures are based on the opinion of a listener or a group of listeners of the quality of an utterance. It is the preferable means of quality assessment, but may be variations or biases among listeners. Here we take the objective quality measures, which are not equivalent to Subjective measures. Experimental evaluation of the proposed algorithms is performed in terms of segmental SNR and Itakura-Satio Distortion Measure (IS) measure.

Donating the original clean signal and estimated speech signal by s and $\hat{s}$,respectively. The segmental SNR (SNR seg) is a much-improved frame-based quality measure that SNR is measured over short frames and the results averaged, and is formulated as

$$SNR_{seg} = \frac{1}{M}\sum_{j=0}^{M-1} 10\log_{10}\left[\frac{\sum_{n=m_{j-N+1}}^{m_j}\hat{s}(n)^2}{\sum_{n=m_{j-N+1}}^{m_j}[s(n)-\hat{s}(n)]^2}\right]$$

where m_0, m_1, ... , m_{M-1} are the end-times for the M frames, each of which is length N. The segmentation of the SNR permits the objective measure to assign equal weight to loud and soft portions of the speech.

As the SNR itself may not provide enough information about the usefulness of an enhancement algorithm, we also provide evaluation by calculating Itakura-Satio (IS) Distortion Measure. Based on linear predictive coding (LPC), IS measure provides a meaningful measure of performance when the two waveforms differ in their phase spectra, as the human auditory system is relatively insensitive to phase distortion, and is defined by

$$d_{IS}\left(\vec{a}_p,\vec{a}_c\right)=\left[\frac{\sigma_c^2}{\sigma_p^2}\right]\left[\frac{\vec{a}_p\mathbf{R}_c\vec{a}_p^T}{\vec{a}_c\mathbf{R}_c\vec{a}_c^T}\right]+\log\left(\frac{\sigma_c^2}{\sigma_p^2}\right)-1$$

where σ_c^2 and σ_p^2 are the LPC gains of the clean and denoised signals, respectively, $\vec{a}_c$ is the LPC vector of the original speech signal frame, $\vec{a}_p$ is the LPC vector of the denoised speech frame, and $\mathbf{R}_c$ is the autocorrelation matrix of the original speech signal. .The IS values were limited in the range of [0, 100]. This was necessary in order to minimize the number of outliers.

The experimental evaluation was performed with speech signals selected from the TIMIT database, separately for each gender. The training set contained 30 sentences for each gender, which were randomly selected part of SA2 from the MSJS1 subset. Noisy speech signals were created by adding a noise to the clean speech signal at various SNRs. We used Gaussian noise from the Noisex92 database. Here moderate SNR = 10dB is considered.

The performance of the proposed algorithms (TdSSD) was compared with the spectral subtraction, Wiener filter, discrete wavelet transform (DWT) enhancement method proposed in [7] employing Symlets wavelet.

An example of clean, noisy and denoised speech signals are depicted in Figure 1. , It can be observed in the time-domain representation that the denoised speech signals obtained by all the algorithms have a low residual noise level as well as some distorts. TdSRD almost sets the silence period to zeros.

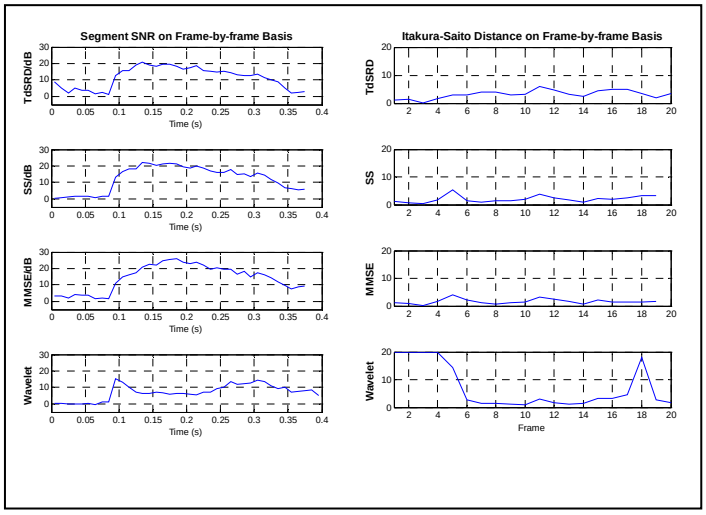


Figure 2. Comparision of Sengment SNR and IS of Denoised Signals at SNR = 0dB

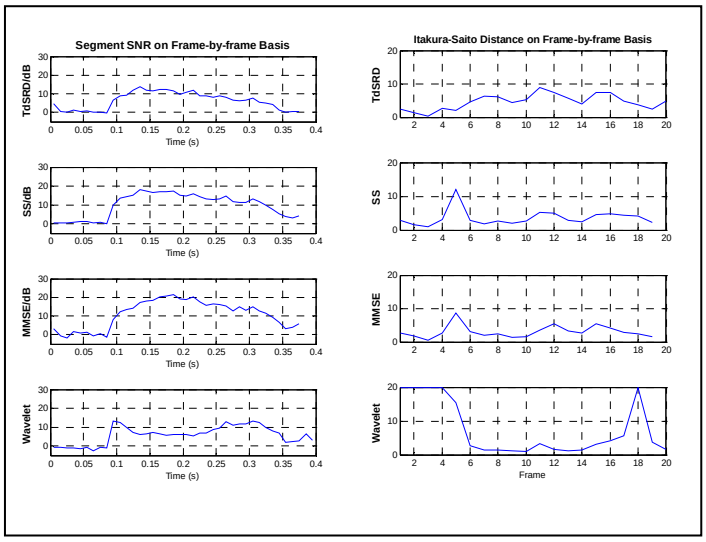


Figure 3. Comparision of Sengment SNR and IS of Denoised Signals at SNR = 5dB

The obtained results in terms of segment SNR and IS are presented in the Figure 2. Figure 3. at SNR = 0dB,5dB, respectively. It shows that It can be seen that TdSRD algorithms in most cases improve SNR with moderate distorts. Compared with other algorithms, TdSRD performs better in the case of SNR = 5dB than in the case of SNR = 0dB.

B. Regularization Parameter Choice

We illustrate the importance of a good choice of the regularization parameter in TABLE I. For "bad choice," the regularization parameter was set lower, that would lead to the increase of IS and deterioration of SNR. The choice relates to many factors such as the SNR, the length of frame and so on. In normal cases, $\beta^2 = 0.1$ usually is a good choice. Another fact we notice is that we can not obtain the same good choice based on different assessments.

TABLE I. THE SELECTION BASED ON SEGMENTSNR AND IS

β^2	**Quality Assessments**	
	Min IS	***Max segment SNR/dB***
1	11.2933	0
10^{-1}	9.0672	11.9
10^{-2}	11.4911	16.12
10^{-3}	11.5472	15.89

V. CONCLUSION

In this paper, we explored a noval formulation of single channel speech denoising problem in a sparse signal representation framework. We started with a scheme for speech deonising with the whole signal and developed a tractable and stable frame-by-frame-based method. We proposed an automatic method for choosing the regularization parameter using the constrained form of the discrepancy principle at high SNR. Finally, Experiments were performed on speech data corrupted by Gaussian noises and the evaluation was performed in terms of segment SNR and IS measure. The experimental results, confirming our theoretical analyses, showed some performance improvements, in some aspects, over conventional speech denoising techniques.

REFERENCES

[1] J. S. Lim and A. V. Oppenheim, "Enhancement and bandwidth compression of noisy speech," *Proceedings of the IEEE*, vol. 67, pp. 1586-1604, 1979.

[2] S. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 27, pp. 113-120, 1979.

[3] C. He and G. Zweig, "Adaptive two-band spectral subtraction with multi-window spectral estimation," 1999, pp. 793-796.

[4] A. Hyvärinen, *et al.*, "Sparse code shrinkage: Denoising by nonlinear maximum likelihood estimation," 1999, p. 473.

[5] D. L. Donoho, *et al.*, "Wavelet shrinkage: asymptopia?," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 301-369, 1995.

[6] V. Balakrishnan, *et al.*, "Wavelet denoising and speech enhancement," *Department of Electrical and Computer Engineering, The Johns Hopkins University, Baltimore*.

[7] D. L. Donoho, "De-noising by soft-thresholding," *Information Theory*, IEEE *Transactions on*, vol. 41, pp. 613-627, 1995.

[8] S. Sardy, *et al.*, "Robust wavelet denoising," *Signal Processing*, IEEE Transactions *on*, vol. 49, pp. 1146-1152, 2001.

[9] M. D. Plumbley, *et al.*, "Sparse representations in audio and music: from coding to source separation," *Proceedings of the IEEE*, vol. 98, pp. 995-1005, 2010.

[10] D. L. Donoho and M. Elad, "Maximal sparsity representation via l1 minimization," *Proc. Nat. Aca. Sci*, vol. 100, pp. 2197-2202, 2003.

[11] R. Gribonval and M. Nielsen, "Sparse representations in unions of bases," Information *Theory, IEEE Transactions on*, vol. 49, pp. 3320-3325, 2003.

[12] S. S. Chen, *et al.*, "Atomic decomposition by basis pursuit," *SIAM journal on scientific computing*, vol. 20, pp. 33-61, 1999.

[13] V.A.Morozov, "On the solution of functional equations by the method of regularization," *Soviet Math. Dokl.*, vol. 7, pp. 414-417, 1966.

[14] P. C. Hansen, "Analysis of discrete ill-posed problems by means of the L-curve," SIAM *review*, pp. 561-580, 1992.

Delay-Aware Model-Based Accurate Simulator for Asynchronous NoC Design

Yuma Watanabe, Atsushi Matsumoto, and Takahiro Hanyu
Research Institute of Electrical Communication, Tohoku University
2-1-1, Katahira, Aoba-ku, Sendai, Miyagi, 980-8577, Japan
E-mail: {yuma_w, matumoto, hanyu}@ngc.riec.tohoku.ac.jp

Abstract— **An accurate circuit-performance simulator using precise delay tables is proposed for asynchronous Network-on-Chip (NoC) design. In the proposed simulation model, LUT-based delay elements are used for estimation of operation delay, while functional elements emulate handshake-based control used in an asynchronous circuit. As a result, fast and accurate simulator for asynchronous NoC can be realized.**

I. Circuit Model used in Proposed simulator

Asynchronous Network-on-Chip (NoC) systems have been recently focused to advance VLSI performance. However, asynchronous NoC simulator with both high speed and accuracy has not been proposed yet. Fig. 1 (a) illustrates a timing chart of an asynchronous circuit simulated by the conventional gate-level simulator. Its timing evaluation is accurate because delay time of each logic gate is actually calculated. Although, caused by the precise delay calculation in many points, it takes too much time to simulate the large-scale NoC. Fig. 1 (b) shows a timing chart by cycle-accurate simulator. In this simulator, delay time in unit circuit component is estimated as a fixed period. Therefore, computation time is greatly reduced compared with it by gate-level simulator because of the simple delay estimation at a few points in cycle-accurate simulator. However, the operations of an asynchronous circuit are not performed at a constant period. Therefore, estimation error is inevitable in a simulation of an asynchronous circuit and it resulted in degradation of accuracy of delay in the simulator.

For fast and accurate simulation of asynchronous NoC like Fig. 1(c), a simulator which can handle variable-period estimation at a few evaluation points is desirable. Fig. 2 shows an example of an asynchronous circuit model used in the proposed simulator. This model consists of variable delay elements and a functional element. Gate-level circuit operation is replaced by the LUT-based delay estimation which makes its computation cost small. Furthermore, handshaking used in the asynchronous control is emulated by the functional elements. As a result, fast and accurate evaluation of asynchronous NoC can be realized by using proposed simulator.

II. Discussion and Conclusion

Table 1 shows the estimated error in latency and normalized computation time of a simple asynchronous circuit evaluated by the conventional and proposed simulators. Latency computed by proposed simulator is close to it by gate-level simulator. On the other hand, computation time is almost 350 times faster than it by gate-level simulator. It is competitive to the evaluation time by cycle-accurate simulation. As a future prospect, it is also necessary to develop a simulation environment for various topologies of asynchronous NoC.

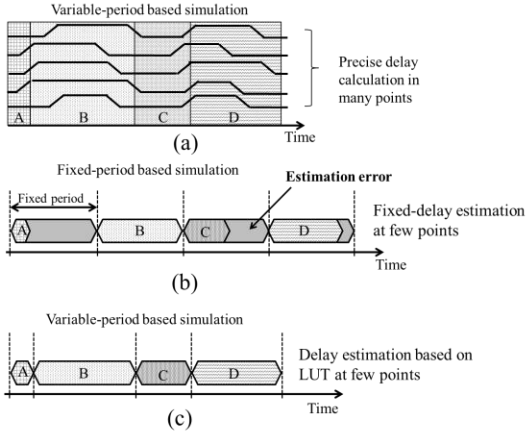


Figure 1. Timing chart by simulators: (a) gate-level, (b) cycle-accurate and (c) desirable (proposed)

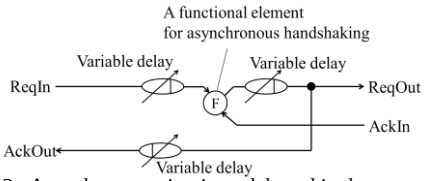


Figure 2. Asynchronous circuit model used in the proposed simulator

TABLE I. COMPARISON OF SIMULATORS

	Gate-level	Cycle-accurate	Proposed
Estimated Error Rate in Latency	-	53.1%	3.0%
Normalized Computation Time	1	0.0024	0.0028

Acknowledgment

This work is supported by VLSI Design and Education Center (VDEC), The University of Tokyo in collaboration with Synopsys, Inc.

References

[1] N. Onizawa, et al., IEEE Computer Society Annual Symposium on VLSI 2010, pp.357-362, 2010.

[2] A. Banerjee, et al., Very Large Scale Integration (VLSI) System, IEEE Transaction on, vol.17, pp.319-329, 2009

A Calibration for Gain and Phase Errors of Sensors Based on A Moving Sensor

Wang Tiedan
School of Electronic Engineering, University of Electronic Science and Technology China
ChengDu China
Email: fox113@126.com

Wei Ping
School of Electronic Engineering, University of Electronic Science and Technology China
ChengDu China

Abstract—**The problem on determining the Direction-of-Arrival(DOA) of multiple sources impinging upon a LES array under the sensor gain and phase errors was addressed. We propose a simple and fast calibration algorithm by using a un-calibration moving sensor. And this method does not require any a priori knowledge of the DOA of a signal. Computer simulation results illustrate the good performance of the proposed algorithm.**

Keywords-component; DOA; model uncertain; calibration;

I. INTRODUCTION

The estimation of DOA of multiple narrowband signals is a classical problem in array signal processing. Many subspace-based algorithms with high-resolution for array signal processing have attracted considerable interest in recent years. However, when array error such as the uncertainty of sensor's gain and phase exists, their performance degrades significantly. So it is essential to calibrate the sensors errors in practice.

A number of blind calibration algorithms have been developed to calibrate the sensor gain and phase errors using some presumed structures. Weiss and Friedlander presented the Eigen-structure method for direction finding in [1]. This method can estimate the signals' direction without knowing signals' DOA. But it has great computations and this method is not convergent when the signal to noise ratio (SNR) is not very high or the signals impinging on the array are not separated enough in space. In [2] a procedure is developed for LES array that calibrates the array by estimating the sensor gain and phase perturbations. The Y.H.Sng and Youming Li improved [2] in [3] and [4]. M.P.Wylie,S.Roy，H.Messer and others have also do some research in array calibration algorithms based on linear equi-spaced(LES) array[5]-[7].

This paper presents a method by using a moving sensor to calibrate the sensor gain and phase errors. This method can be achieved by the moving sensor slide parallel with array. It can be regarded as standard to calibrate the array sensors errors. Combined with method [3], the result is much more practical and better than before.

This paper is organized as follows: In section II, the data model used is introduced. The calibration method using moving sensor is given in section III. Section IV combined the blinding calibration to our method. Numerical computer simulation can be found in section V.

II. ARRAY MODEL

Consider a Uniform Linear Array (ULA) composed M sensor with no sensor errors. There are K $(1 < K < M)$ far-field incoherent signals impinging on the array from directions $\{\theta_1, \theta_2, \cdots, \theta_K\}$ with wavelength λ . The output of array is given by:

$$\mathbf{X(t)} = \mathbf{AS(t)} + \mathbf{n(t)} \tag{1}$$

where $\mathbf{S(t)} = [\mathbf{s_1(t)}, \mathbf{s_2(t)}, \cdots, \mathbf{s}_K\mathbf{(t)}]^T$ is signal matrix; $\mathbf{n(t)}$ is a complex white Gaussian noise with zero mean, t is a snapshot and $\mathbf{A}$ is the directional matrix:

$$\begin{aligned}\mathbf{A} &= [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \cdots, \mathbf{a}(\theta_N)] \\ &= \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{-j\phi_1} & e^{-j\phi_2} & \cdots & e^{-j\phi_N} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-j(M-1)\phi_1} & e^{-j(M-1)\phi_2} & \cdots & e^{-j(M-1)\phi_N} \end{bmatrix}\end{aligned} \tag{2}$$

where $\phi_k = 2\pi d \sin\theta_k / \lambda$.

With the presence of gain and phase errors, the array output has been changed as:

$$\mathbf{X}(t) = \mathbf{\Gamma AS}(t) + \mathbf{n}(t) \tag{3}$$

where $\mathbf{\Gamma} = diag(g_1 e^{j\varphi_1}, g_2 e^{j\varphi_2}, \cdots, g_M e^{j\varphi_M})$, g_i 、 φ_i are gain and phase error respectively.

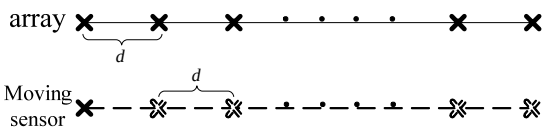


Figure 1. The moving sensor array model

The moving sensor is parallel with array, and slide with distance d, as it shown in figure 1.

We divide time t into M equivalently: $t_1, t_2, \cdots, t_M$. At the period of t_i $i = 1, 2, \cdots, M$, the moving sensor parallel the i th sensor of array, the output of the moving sensor is:

$$\mathbf{x}_e(t_i) = g_e \exp(j\phi_e)\mathbf{A}(i,:)\mathbf{S}(t_i) + n_e(t_i) \tag{4}$$

and the i th sensor of array output is:

$$\mathbf{x}_i(t) = g_i \exp(j\phi_i)\mathbf{A}(i,:)\mathbf{S}(t) + n_i(t_i) \tag{5}$$

III. Calibration Methods

When the process of sensor movement is over, the moving sensor has been paired with all the array sensors. We have:

$$\begin{aligned} d_i &= \frac{E\left[x_i(t_i)x_i^H(t_i)\right]}{E\left[x_e(t_i)x_e^H(t_i)\right]} \\ &= \frac{E\left\{\left[g_i \exp(j\phi_i)\mathbf{A}(i,:)\mathbf{S}(t_i) + n_i(t_i)\right]\right.}{E\left\{\left[g_e \exp(j\phi_e)\mathbf{A}(i,:)\mathbf{S}(t_i) + n_e(t_i)\right]\right.} \\ &\quad \frac{\left.\bullet\left[g_i \exp(j\phi_i)\mathbf{A}(i,:)\mathbf{S}(t_i) + n_i(t_i)\right]^H\right\}}{\left.\bullet\left[g_e \exp(j\phi_e)\mathbf{A}(i,:)\mathbf{S}(t_i) + n_e(t_i)\right]^H\right\}} \\ &= \frac{g_i^2 P + \sigma^2}{g_e^2 P + \sigma^2} \approx \frac{g_i^2}{g_e^2} (P \gg \sigma^2) \end{aligned} \tag{6}$$

where, $P = E\left\{\left[\mathbf{A}(i,:)\mathbf{S}(t_i)\right]\bullet\left[\mathbf{A}(i,:)\mathbf{S}(t_i)\right]^H\right\}$ is the power of sensor output. Consider the distance of each sensor is not far, we can suppose that the power of impinging signal of each sensor is equal.

Then we let:

$$\begin{aligned} \gamma_i &= E\left[x_i(t_i)x_e^H(t_i)\right] \\ &= E\left\{\left[g_i \exp(j\phi_i)\mathbf{A}(i,:)\mathbf{S}(t_i) + n_i(t_i)\right]\right. \\ &\quad \left.\bullet\left[g_e \exp(j\phi_e)\mathbf{A}(i,:)\mathbf{S}(t_i) + n_e(t_i)\right]^H\right\} \\ &= g_i g_e P \exp(j(\phi_i - \phi_e)) \end{aligned} \tag{7}$$

and

$$\rho_i = angle(\gamma_i) = \exp(j(\phi_i - \phi_e)) \tag{8}$$

And then define the calibration matrix as follows:

$$\mathbf{C} = diag\left(c_1, c_2, \cdots, c_M\right) \tag{9}$$

where $c_i = \sqrt{d_i} \bullet \exp(j\rho)$.

The received data is calibrated as

$$\begin{aligned} \mathbf{X}_e(t) &= \mathbf{C}^{-1}\left(\mathbf{\Gamma}\mathbf{A}\mathbf{S}(t) + \mathbf{n}(t)\right) \\ &= g_e \exp(j\phi_e)\mathbf{A}\mathbf{S}(t) + \mathbf{C}^{-1}\mathbf{n}(t) \end{aligned} \tag{10}$$

The estimation of DOA will not be affected by the constant $g_e \exp(j\phi_e)$ and the $\mathbf{C}^{-1}\mathbf{n}(t)$ can be regarded as new gauss noise.

IV. Combine with Blinding Calibration

The proposed method in section III does not using the array structure to calibrate the sensor errors. We can join the proposed method to some blinding calibration. In this way, we can improve the performance of blinding calibration and make proposed method more practice.

A number of auto-calibration algorithms have been proposed to calibrate the sensor errors using LEA. Paulraj and Kailath has propose a calibration method by using the structure of LEA[2]. Youming Li and M.H.Er simplify that method and have better performance[4]. But it has demonstrated that those methods phase error estimation are biased estimation[4]. Combining our method to those blind calibration methods could vastly improve the performance.

The covariance matrix of LES is given as:

$$\mathbf{R}_{xx} = \begin{bmatrix} r_0 & r_1 & r_2 & \cdots & r_{M-1} \\ r_1^* & r_0 & r_1 & \cdots & r_{M-2} \\ r_2^* & r_1^* & r_0 & \cdots & r_{M-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{M-1}^* & r_{M-2}^* & r_{M-3}^* & \cdots & r_0 \end{bmatrix} \tag{11}$$

With the presence of gain and phase errors, the array output covariance matrix has been changed as:

$$\tilde{\mathbf{R}} = (\tilde{r}_{ij}) = \mathbf{\Gamma R \Gamma^H} = (r_{ij} g_i g_j e^{j(\varphi_i - \varphi_j)}) \tag{12}$$

According to the Y.H. Sng and Youming Li's method, the gain error is estimated just by considering the amplitude differences between two adjacent elements in the main diagonal, and the phase errors is estimated by considering the phase differences between two adjacent elements in the first upper diagonal line as follows:

(1) Gain error estimation

Let $u_{ijkl} = \ln\left\{\left|\hat{r}_{ij}\right| / \left|\hat{r}_{kl}\right|\right\}$. Then the gain errors can be estimated by solving the following $(M-1)$ linear equations:

$$2\ln g_i - 2\ln g_{i+1} = u_i \quad (i = 1, 2, \cdots, M-1) \tag{13}$$

(2) Phase error estimation

Let $v_i = angle\left\{\left|\hat{r}_{i(i+1)}\right| / \left|\hat{r}_{(i+1)(i+2)}\right|\right\}$ be the phase difference between element $\hat{r}_{i(i+1)}$ and $\hat{r}_{(i+1)(i+2)}$ .Then the phase errors can be estimated by solving the following $(M-2)$ linear equations:

$$\varphi_i - 2\varphi_{i+1} + \varphi_{i+2} = v_i (i = 1, 2, \cdots, M-2)$$

When $\hat{g}_i$ and $\hat{\varphi}_i$ are estimated by above equations, let $\hat{\Gamma}_{ii} = \hat{g}_i e^{j\hat{\varphi}_i}$. Then the gain and phase errors can be calibrated by the transformation $\hat{\mathbf{\Gamma}}^{-1}\hat{\mathbf{R}}\left(\hat{\mathbf{\Gamma}}^{\mathbf{H}}\right)^{-1}$.

But we can see that the matrix of equations for the phase error estimated is:

$$\mathbf{A_{phase}} = \begin{bmatrix} 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & & \vdots \\ 0 & 0 & \cdots & 1 & -2 & 1 \end{bmatrix} \Bigg\} M-2 \tag{14}$$

We can easy known that $rank\left(\mathbf{A}\right) = M-2 < M$. So it is a underdetermine problem, and will bring large errors of phase error estimated.

Introduce the information of the moving sensor, see from (7) and (8), we can get:

$$\rho_i = angle\left(\gamma_i\right) = \exp(j(\phi_i - \phi_e)) \tag{15}$$

From every movement time, it can get that:

$$\Delta\varphi_{ij} = \frac{\rho_i}{\rho_j} = \exp(j(\phi_i - \phi_j)) \tag{16}$$

Every movement is equivalently to add one equation. Substituted the increase equation into (14), it can solve the underdetermine problem. Simulation results demonstrate the effectiveness of the proposed method

V. SIMULATION RESULTS

Defines: The array gain and phase errors were generated by the following formula:

$g_i = 1 + \sqrt{a}\delta_g, \varphi_i = \sqrt{b}\delta_\varphi$, where δ_g, δ_φ are the Gaussian-distributed random variables with zero mean and variance $\delta^2 = 1$, a, b are constants.

With N Monte Carlo runs, the root mean-square error(RMSE) is defined as

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\left(\hat{\theta}_{1i} - \theta_1\right)^2 + \left(\hat{\theta}_{2i} - \theta_2\right)^2\right]} \tag{17}$$

A. moving sensor method

We first present comprehensive results of simulations using the algorithm described in section III for DOA estimation for two source with $\theta_1 = 55^\circ$ and $\theta_2 = 60^\circ$ using LES arrays of sizes M=8 , respectively, with interelement spacing of the half wavelength of the signal carrier $\left(d = \frac{\lambda}{2}\right)$. Snapshot data was generated according to the observation model (3)

$$\mathbf{X}(t) = \mathbf{\Gamma A S}(t) + \mathbf{n}(t)$$

and the covariance matrix of the observations was estimated by a sample average based on 200 snapshots every movement.

$$\begin{aligned} \hat{d}_i &= \frac{\frac{1}{200}\sum_{i=1}^{200} x_i(t) x_i^H(t)}{\frac{1}{200}\sum_{i=1}^{200} x_e(t) x_e^H(t)} \\ \hat{\gamma}_i &= \frac{1}{200}\sum_{i=1}^{200} x_i(t) x_e^H(t) \\ \hat{R} &= \frac{1}{200}\sum_{i=1}^{200} \mathbf{x}(t)\mathbf{x}^H(t) \end{aligned} \tag{18}$$

All simulations were assuming the noise is a complex Gaussian distribution, independent of the signal. $a = 0.05$, $b = 9^\circ$, $N_k = 1000$ Monte Carlo runs. The performance RMSE versus the signal-to-noise ratio(SNR) of the two algorithms are shown in Fig.2:

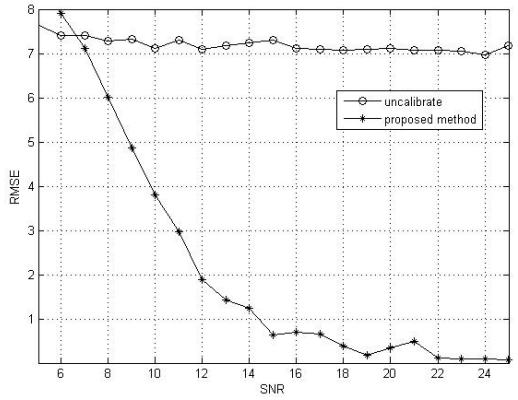


Figure 2. RMSE analysis for estimating DOA versus SNR

B. Combine with blinding calibration

The performance will great improvement when combine our method to Y.H. Sng and Youming Li's method. Supposed Two incoherent narrowband sources with $\theta_1 = 55^\circ$, $\theta_2 = 58^\circ$ incident on a uniform linear array with elements $M = 8$. the snapshot number is 200 and $a = 0.05;\ b = 64^\circ, N_k = 1000$. All the simulation results show that the performance of the improved method is better to the method of Y.H. Sng-Youming Li. Showing in Fig.3

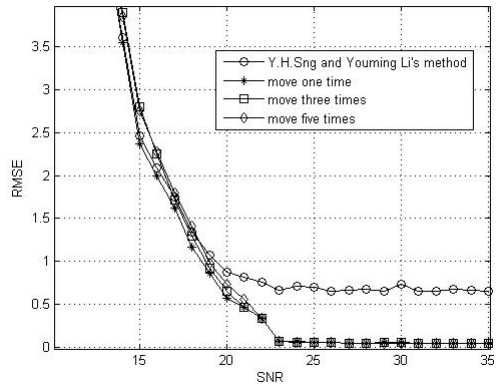


Figure 3. RMSE analysis for combine with Y.H. Sng and Youming Li's method for estimating DOA versus SNR

VI. CONCLUSION

This paper described a simplified method by using a move sensor to estimate the gain and phase error in LEA. And this method does not require iteration, so it with little computation. Simulation results verified the effectiveness of the proposed algorithm. Combine with Y.H. Sng and Youming Li's method, it will reduce the engineering difficulties and improve the performance of the original method in [3] and [4]。

REFERENCES

[1] A.J Weiss and B.Friedlander, "Eigenstructure methods for direction finding with sensor gain and phase uncertainties." Circuits, System & Signal Processing, vo1.9, Mar. 1990: 217-300,

[2] A. Paulraj and T. Kailath, "Direction of arrival estimation by eigenstructure methods with unknown sensor gain and phase," in Proc. Int. Conf. Acous., Speech, Signal Processing(ICASSP), Tampa, FL, March 1985. 640-643.

[3] Y. H. Sng and Youming Li, "Fast algorithm for gain and phase error calibration of linear equi-spaced(LES) Array", Proceedings of ICSP Volume 1, 21-25 Aug. 2000:441 - 444 vol.1

[4] Y. H. Sng and Youming Li, "Theoretical Analyses of Gain and Phase Error Calibration With Optimal Implementation for Linear Equispaced Array", IEEE Trans. Signal Process, vol.54,No.2, FERUARY 2006

[5] M.P.Wylie,S.Roy, and H.Messer, "Joint DOA estimation and phase calibration of linear equispaced(LES) arrays," IEEE Trans. Signal Process., vol. 42, pp.3449-3459, Dec.1994

[6] E.K.L.Huang, "A critical study of a self-calibrating direction-finding method for arrays," IEEE Trans. Signal Process., vol.42, pp.471-474, Feb.1994.

[7] V.C.Soon,L.Tong, Y.F.Huang, and R.Liu, "A subspace method for estimating sensor gains and phases," IEEE Trans. Signal Process., vol. 42, pp.973-976, Apr.1994.

A NOVEL DESIGN FOR FOUR ELEMENTS TIME DOMAIN ULTRA-WIDE BAND ANTENNA ARRAY

Pengcheng Li[1], Jin Pan[1], Deqiang Yang[1]
[1]*Univerty of Electronic Science and Technology of China, Chengdu, China*
[1]lpc_182020861@sohu.com

Abstract—***In the paper, a new design of four elements time domain ultra-wide band antenna array is applied to control the delay time of every element. By the design, automatic scan of synthetic pattern is realized. In four elements array, theoretical analysis and compute is employed to set delay time. The pattern was measured have error. Then, compensation computational method is used to modify the delay time. The theory is validated by renewed measured data. Between -40° and 40°, the angle of scanning beam is precisely controlled.***

Keywords-time domain antenna array; delay time; precisely controll; modified method

I. INTRODUCTION

Generally, in GPR(ground penetrating radar) system[1,2], antenna array is controlled to follow linear mechanical motion of radar mainframe. So, large area scanning can be realized. Although this method offers simple circuits design and low cost, complicate mechanical control is inevitable. Besides, it also has low scanning speed and high failure rate. But, these disadvantages are not belong to antenna array with electric scanning method. Otherwise, in some algorithm, just as time domain synthetic aperture method and time reversal method[2,5], GPR system of mechanic method could be replaced by electric scanning antenna array of time controlling. The higher gain is obtained. It can significantly increase detecting depth. In GPR system, many time domain UWB(ultra-wide band) antennas are be used[4-9]. The element antenna[11], used in the test, has good performance of time domain wide band. The parameters, just as array element distance and element delay time, are reasonably designed to insure precision of scanning angle.

In Theoretical analysis

II. THEORETICAL ANALYSIS

The concept of time domain ultra-wide band time control antenna array is derived from frequency domain phased-array antenna. In time domain antenna array, beam scanning of array is realized by accurately delay time controlling. Delay time controlling is easily achieved by SRT(software radio technology). With a simple system design, high reliability and low cost is offered. In our measurement, time delay of every time delay is calculated first.

When the scanning angle is θ, distance between transmit element and receive antenna can be expressed as fig.(1) (θ<0°).

$$L_{Am}=\sqrt{[(\frac{n+1}{2}-m)r]^2+d^2-2d(\frac{n+1}{2}-m)r\sin\theta},m\le\frac{n}{2} \tag{1}$$

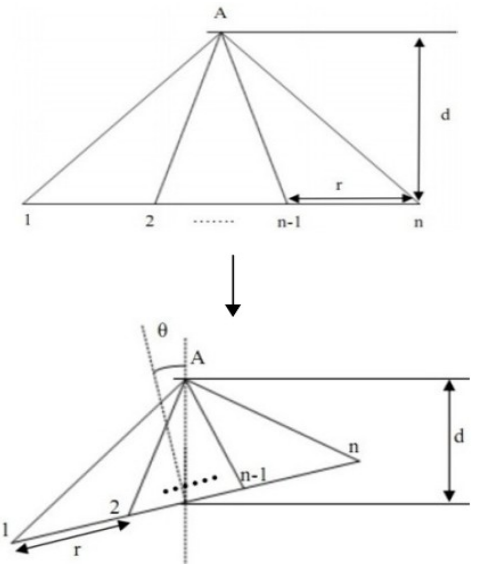
$$L_{Am}=\sqrt{[(m-\frac{n+1}{2})r]^2+d^2+2d(m-\frac{n+1}{2})r\sin\theta},m\ge\frac{n}{2} \tag{2}$$

The delay time for each antenna at a random angle can be calculated be equation (3),

$$t_{1m}=\frac{L_{A1}-L_{Am}}{c} \tag{3}$$

So, received signal can be defined as equation (4),

$$y_n(t) = A\sum_{m=1}^{n} x[t - t_{1m}] \qquad (4)$$



Figur.1 Array (Horizon and Deflection)

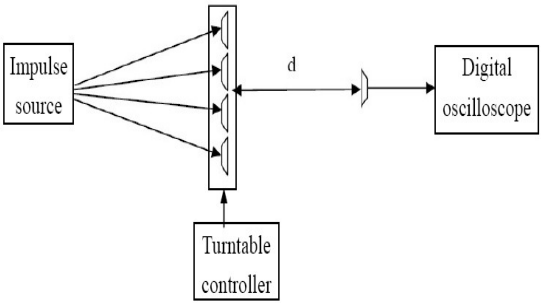
In equation (4), A means the synthesis of antenna gain and route decay. In our four element antenna array measurement, the r and d are *120mm* and *1960mm* respectively. θ is changed from -40° to 40° . The step is 10° . The delay time of each antenna element, calculated by equation (1)-(3), is shown in table 1.

III. MEASUREMENT

Test equipments and diagram are shown in fig.(2). The four element array is set up on the turntable. Digital oscilloscope is applied as receicer.

(a)

(b)



(c)

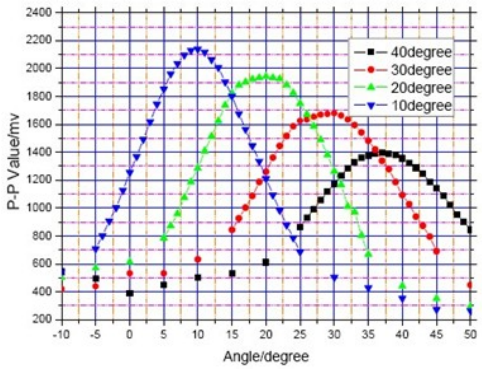
Figur.2

a. Antenna array and turntable

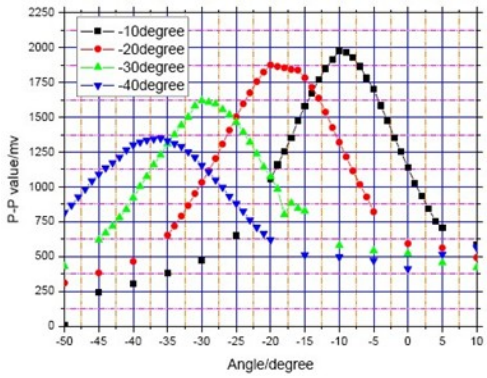
b. Digital oscilloscope

c. Measurement diagram

Just as fig.(3) shown, from -30 to 30 degree, the maximum radiation direction is satisfied the computed result. But, at ±40 degree, there is an error. The maximum radiation direction has an offset of 3 degree. So at this degree, algorithm should be modified.



(a)



(b)

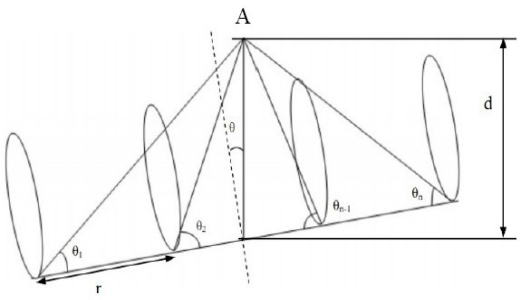
Figur.3 Measured pattern

a. 10～40 degree

b. -10～-40 degree

IV. ERROR ANALYSIS AND MODIFY

This error is introduced be directivity of the time domain antenna. Because, in equation (4), A means a constant value. So, equation (4) only be appropriate for the array, which is constituted by antennas with omnidirectional pattern. If element antenna has directivity, equation (4) needs be modified. Just as shown in fig.4.



Figur.4

The modified equation is represented as follow,

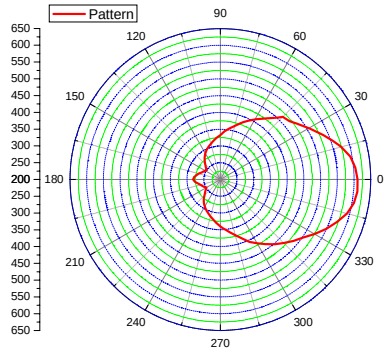
$$y_n(t) = \sum_{m=1}^{n} A(\theta_m) x[t - t_{1m}] \qquad (5)$$

$A(\theta_m)$ is delegated a directivity function.

The element pattern has been measured, just as [9] and fig.5.

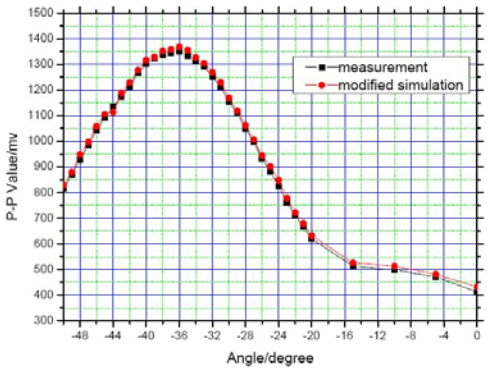
$$\theta_m = \frac{\arccos\{[L_{Am}^2 + \frac{(n-2m+1)^2}{4} r^2 - d^2]}{[(n-2m+1)L_{Am} r]\}}, m \le n/2 \qquad (4)$$

$$\theta_m = \frac{\arccos\{[L_{Am}^2 + \frac{(2m-n-1)^2}{4} r^2 - d^2]}{[(2m-n-1)L_{Am} r]\}}, m \ge n/2 \qquad (5)$$

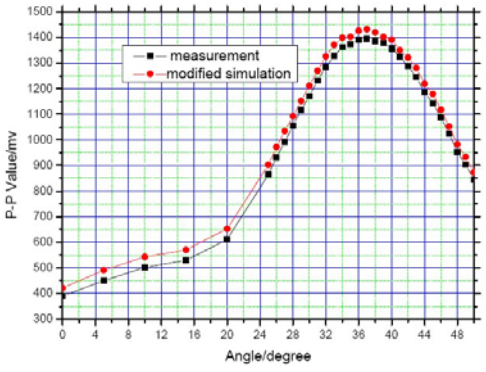


Figur.5 Element pattern

The new revised array pattern is obtained in fig.6. The measured results validate modified algorithm.



(a)



(b)

Figur.6 Modified pattern

a. -40 degree

b. 40degree

V. CONCLUSION

In this paper, time domain antenna array with time controller is analyzed. Between -30 to 30 degree, the scanning angle is accurately managed. But, while it is used to larger scanning angle, precisely control is hardly realized. Modified algorithm is proposed. The maximum radiation direction is computed according to the degree of deflection. The precision is 1 degree. The theory is validated by measurement result. So, with high accuracy, time domain UWB (ultra-wide bandwidth) of time delay controller can be used in GPR. Measurement speed is increased in large area detecting with less mechanical control. Besides, measurement distance could be improved by higher gain of the array in application

Reference

1 Serbin. G, Or. D, "Ground-penetration radar measurement of soil water content dynamics using a suspended horn antenna", IEEE Transactions on, Vol.42, pp.1695-1705,2004

2 Sun.Y, Li. X, Li. J, "Practical landmine detector using forward-looking ground penetrating radar," Electronics Letters, Vol. 41, pp.97-98, 2005

3 G. E. Atteia, A. A. Shaalan, and K. F. A. Hussein, "Practical landmine detector using forward-looking ground penetrating radar," Progress in Electromagnetics Research, PIER 71, 211-226, 2007

4 Ren, W., Z. G. Shi, and K. S. Chen, "Novel planar monopole UWB antenna with 5-GHz band-notched haracteristic", Journal of Electromagnetic Waves and Applications, Vol. 21, No. 12, 1645–1652, 2007.

5 Zheng, Wenjun, Zhiqin Zhao and Zaiping Nie, "Application of TRM in the UWB Through Wall Radar," Progress In Electromagnetics Research, PIER 87, pp. 279-296, 2008.

6 Chen, Ye, William T. Joines, Zhentian Xie, Guining Shi, Qing H. Liu, and Lawrence Carin, "Double-Sided Exponentially Tapered GPR Antenna and Its Transmission Line Feed Structure" ,IEEE Trans. on Antennas and Propagation, Vol. 54, pp.2615-2623, Sep. 2006

7 Shams, K. M. and M. Ali, "A planar inductively coupled bow-tie slot antenna for WLAN application," Journal of Electromagnetic Waves and Applications, Vol. 20, No. 7, 861–871, 2006.

8 Wu, T. T. and R. W. King, The Cylindrical Antenna with Nonreflecting Resistive Loading, IEEE Trans. on Antennas and Propagation, Vol. 13, pp.369-373, 1965.

9 Pan, Jin, Zai-Ping Nie, "The Loading Formula for V-type Traveling wave Antenna," Chinese Journal of Radio Science, Vol.11, No.2 June, 1996

10 Zhou, You, Jin Pan, and Zai-ping Nie, "Engineering Design of the Transient Back-Cavity Bowtie Antenna," Journal of University of Electronic Science and Technology of China, Vol. 34, pp. 1-3, 2005.

11 D. Yang, J. Pan, Z. Zhao, and Z. Nie, "Design of Trapezoidal Cavity-Backed Resistance Loaded Bow Tie Antenna with Ultra-wideband and High Directivity", Journal of Electromagnetic Waves and Applications, Vol. 24, No. 8, 1685–1695, 2010.

12 Glenn, S, Smith, "A Note on the Criteria for the Far Zone in the Time-Domain Analysis of Antennas," IEEE Transaction on. Antenna and Propagation, Vol.54, pp.292-297, January 2006.

Table1 Scanning angle and delay time of element antenna

Scanning Angle(θ)	-40°	-30°	-20°	-10°	0°	10°	20°	30°	40°
antenna1	-224ps	-224ps	-224ps	-224ps	-224ps	-224ps	-224ps	-224ps	-224ps
antenna2	20ps	-34ps	-104ps	-134ps	-200ps	-258ps	-384ps	-480ps	-444ps
antenna3	290ps	150ps	50ps	-96ps	-200ps	-335ps	-530ps	-600ps	-847ps
antenna4	532ps	376ps	172ps	-14ps	-224ps	-420ps	-580ps	-880ps	-1070ps

Evaluation of Unsupervised Language Model Adaptation using Retrieval of Relevant Documents

Souichi Sato
Graduate School of Engineering
Tohoku University
Sendai, Japan
souichi_sato@spcom.ecei.tohoku.ac.jp

Ryo Masumura
Graduate School of Engineering
Tohoku University
Sendai, Japan
ryo77373@spcom.ecei.tohoku.ac.jp

Akinori Ito
Graduate School of Engineering
Tohoku University
Sendai, Japan
aito@spcom.ecei.tohoku.ac.jp

Abstract— **To improve the accuracy of an LVCSR system, we gather topic-related documents from WWW, and adapt the language model using the gathered texts. We focus on an unsupervised method that adapts the language model without manual specification of the topic or keywords. In this paper, we conducted the language model adaptation using the retrieval of relevant documents. We compared a method based on a small corpus, a method based on Web search and a method using a large corpus downloaded from the Web. We investigated the performance of each method as well as the effect of iterative adaptation, and found that the method based on the large corpus combined with iterated adaptation gave the largest improvement.**

Keywords-component; speech recognition; language model adaptation

I. INTRODUCTION

An N-gram, one of statistical language models, is the most widely used one as a language model for large vocabulary continuous speech recognition (LVCSR). In general, an N-gram can be learned from a text corpus that contains a variety of topics. Such N-gram shows a high performance to typical speech inputs. However, sometimes the performance does not improve when the input is concerned with a specific topic. The reasons include the influence of out-of-vocabulary (OOV) words, or recognition errors raising from low linguistic probabilities. To solve these problems, many methods are investigated to adapt the language model to the recognition target [1][2]. Several previous works attempted to collect texts relevant to recognition target using Web search engine, and the retrieved relevant documents were used as a source of language model adaptation [3][4].

There have been a couple of works of language model adaptation using the WWW so far; Bulyko et al. combined the language model learned from the typical dialogue and the information obtained from Web by filtering [5]; Ito et al. proposed language model adaptation from news stories and automatically extracted similar articles from present news on the Web [6]; Nemoto et al. proposed a language model adaptation using Web documents downloaded by a search query generated from slides of the lecture [7].

In this paper, we investigate an unsupervised language model adaptation method. Especially, effect of iterative adaptation on accuracy of LVCSR is examined. We used three methods of retrieval of relevant documents. The first one uses small-scale language resources. The second one extracts keywords from the preliminary recognition result and downloads adaptation text using an Web search engine using the keywords. In the third method, we first download large number of Web documents beforehand, and choose documents relevant to the preliminary recognition result among the downloaded text database.

II. LANGUAGE MODEL ADAPTATION USING WWW

A. *Unsupervised Adaptive Procedure*

Table 1 shows the details of the baseline corpus. Figure 1 shows the basic framework of the unsupervised language model adaptation.

Step 1 The baseline language model is trained from a task-independent baseline corpus, and that language model is used to obtain the preliminary speech recognition results from the spoken document.

Step 2 Retrieve documents relevant to the preliminary recognition result using the above-mentioned methods, and create a corpus for the adaptation using the retrieved documents.

Step 3 Create an adapted language model by N-gram count mixture using the baseline corpus and corpus for the adaptation.

Step 4 Speech recognition is performed again using the adapted language model, and the second recognition result is obtained. This recognition result is used as the preliminary recognition result for the second adaptation, and repeat the procedure from Step 2.

Table.1 The baseline language model

Baseline corpus	CSJ 2536 lectures
Vocabulary size	59310
Total number of words	7652530
Backoff smoothing	Witten-Bell
Number of unigram entries	41708(cut-off 1)
Number of bigram entries	874059(cut-off 0)
Number of trigram entries	2869943(cut-off 0)

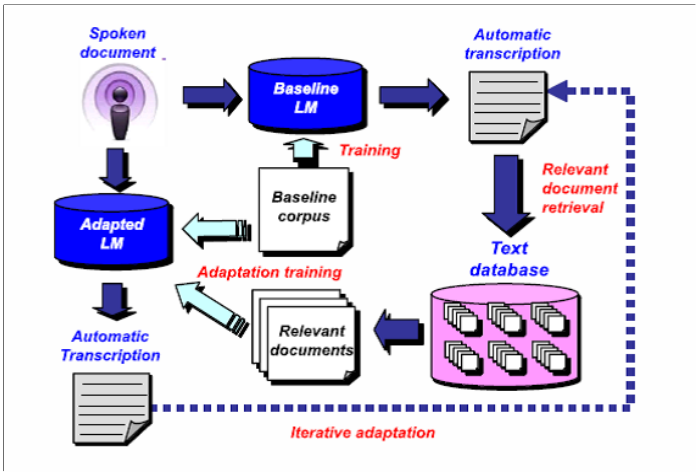


Figure.1 Unsupervised language model adaptation using retrieval of relevant documents

B. Retrieval of Relevant Documents

We used vector space model [13] for retrieval of relevant documents from a local database. Let a feature vector of document D be $\mathbf{I}(D)$, calculated as follows:

$$\mathbf{I}(D) = (t_D(w_1), \ldots, t_D(w_V)) \quad (1)$$

Here, $t_D(w)$ denotes a TF-IDF value of word w in document D. Then the similarity of two documents is measured using the cosine distance:

$$sim(D_1, D_2) = \frac{\mathbf{I}(D_1) \cdot \mathbf{I}(D_2)}{|\mathbf{I}(D_1)| \cdot |\mathbf{I}(D_2)|} \quad (2)$$

When choosing documents relevant to a given document D_{in} from a set of documents $D_1, \ldots D_N$ (i.e. the local database), we choose the documents that have the largest similarity to D_{in}.

C. Search Engine

In this study, we used Yahoo! Japan [8] as a search engine, and Yahoo API [9] was used to send search queries to the server and get search results. We can obtain up to 1000 URLs from one search query by using the Yahoo API.

Table.2 Experimental conditions

Acoustic model	Continuous Density HMM
Sampling frequency	16 kHz
Analysis window	Hamming Window
Window length	25 ms
Frame shift	10 ms
Features	MFCC (12dim.)+ Δ MFCC (12dim.)+ Δ Power
CMS	Performed utterance by utterance

III. Speech Recognition Experiment

We used 40 lectures from the Corpus of Spontaneous Japanese (CSJ) as the test set. Experimental conditions of the recognition experiment is shown in Table 2. We employed Julius [10] as a speech recognizer.

In the experiments in this section, we observed four indices as the performance measure: the word correct rate (Cor), the word accuracy (Acc), the adjusted perplexity (APP), and the out-of-vocabulary rate (OOV). The word correct rate is the ratio of correctly recognized words without considering word insertions. The word accuracy is that considering word insertions. The adjusted perplexity is the average number of distinct words predicted at a certain linguistic context; the lower the APP is, the better the performance of a language model is. The OOV rate is the ratio of words that are not contained in the vocabulary of the language model.

A. Adaptation Result Using Small-Scale Language Resources

First, we used a small corpus as a source of relevant document retrieval. We used Mainichi Shimbun newspaper corpus as the database as the small-scale language resources to evaluate the basic performance of unsupervised language model adaptation [11]. Mainichi Shimbun corpus is a database of newspaper articles. We used two-year data from articles of 2000 and 2001, which contain about 200 thousand articles, around 100 words per 1 article, and the total number of words are 21 million words. Table 3 shows the results of the iterative adaptation when 1000 or 5000 relevant documents (articles) were used as the corpus for adaptation. Number of iteration was changed from one to five.

It is clear from Table 3 that the recognition performance improved by repeating the adaptation. When top 1000 documents were used, word accuracy, adjusted perplexity and OOV improved by 0.31 points, 0.92 and 0.04 point, respectively. When top 5000 documents were used, the improvements were 0.16 points, 1.06 and 0.01 point, respectively.

Table.3 Adaptation Result (Small-Scale Language Resources)

iteration	Cor[%]	Acc[%]	APP	OOV[%]
baseline	67.51	62.45	101.5	1.54
1000-1	69.11	64.74	94.04	0.75
1000-2	69.35	65.03	93.19	0.71
1000-3	69.36	65.03	93.19	0.71
1000-4	69.37	65.05	93.25	0.72
1000-5	69.36	65.05	93.12	0.71
5000-1	69.15	65.1	92.48	0.65
5000-2	69.17	65.23	91.45	0.64
5000-3	69.21	65.26	91.45	0.64
5000-4	69.18	65.26	91.39	0.64
5000-5	69.2	65.26	91.42	0.64

B. Adaptation Result Using a Web Search Engine

When using a Web search engine, we can search relevant documents without keeping large amounts of data in the local storage because we download the relevant data from external resources. The problem is how to compose a search query from the preliminary recognition result. We examined the following two methods for choosing keywords.

- Method 1
  - Step 1 We calculate the TF-IDF for each word that appear in the input document [12]. Then, the words with the highest score are selected as keywords.
  - Step 2 Keyword clustering is performed by similarity between words based on the document co-occurrence frequency [12], and each of the cluster is used as a query.
  - Step 3 The goodness of the composed search queries are measured [12], and the amount of data downloaded by each search query is determined.
- Method 2
  - Step 1 Let the input document be D_{in} . Each word appearing in D_{in} is used as a query of Web search, and a small number of documents are downloaded. Note that the downloaded Web documents are associated with a specific keyword used as the query. We denote the i-th Web document as D_w^i.
  - Step 2 Compose feature vectors for all of the downloaded documents as well as the preliminary recognition result. We calculate

$$\mathbf{f}_{\text{in}} = \mathbf{I}(D_{\text{in}}) \quad (1)$$

and

$$\mathbf{f}_i(w) = \mathbf{I}(D_w^i). \quad (2)$$

Then a feature vector for word w is calculated as

$$\mathbf{F}(w) = \sum_{i=1}^{N} \frac{\mathbf{f}_i(w)}{|\mathbf{f}_i(w)|}. \quad (3)$$

  - Step 3 Similarity between $\mathbf{F}(w)$ and $\mathbf{f}_{\text{in}}$ are calculated for all w:

$$R(w, D_{\text{in}}) = sim(\mathbf{F}(w), \mathbf{f}_{\text{in}}) \quad (4)$$

Then we choose words that give the highest $R(w, D_{\text{in}})$ as the set of keywords.

  - Step 4 Keyword clustering is performed by similarity between words using cosine similarity between feature vectors of words, and queries are constructed.
  - Step 5 The goodness of the composed search queries are measured [12], and the amount of data downloaded by each query is determined.

Table 4 shows the result of iterative adaptation using top 1000 documents as a corpus for adaptation. It is clear for both Method 1 and Method 2 that the improvement by the adaptation is almost converge by five iterations. Also, we obtained larger adaptation effects by Method 2 than by Method 1. We compared the results obtained by five iterations. The improvement of Method 2 for word accuracy, adjusted perplexity and OOV were 0.37 point, 1.02 and 0.04 point, respectively.

Table.4 Adaptation Result (Web search engine)

iteration	Cor[%]	Acc[%]	APP	OOV[%]
baseline	67.51	62.45	101.5	1.54
Method1-1	69.37	64.92	92.18	0.62
Method1-2	69.84	65.51	87.7	0.53
Method1-3	69.99	65.78	87.0	0.47
Method1-4	70.01	65.78	86.56	0.46
Method1-5	70.05	65.8	86.36	0.45
Method2-1	69.86	65.49	88.83	0.56
Method2-2	70.3	66.01	85.32	0.45
Method2-3	70.29	66.14	84.25	0.42
Method2-4	70.39	66.21	84.51	0.4
Method2-5	70.44	66.19	85.29	0.41

C. Adaptation Results Using Downloaded Data

Finally, we examined a method that uses a large amount of downloaded data for the adaptation. The adaptation method based on Web search has an advantage that we do not need to store a large amount of data in a local storage, but it has a

drawback that quality of the adaptation corpus depends on the extracted keywords. In this experiment, we prepared a large corpus downloaded from the Web, and that corpus was used to choose relevant documents.

On creating the large corpus, we prepared about 280 thousands nouns, and each of those nouns was used as a keyword of Web search. We downloaded 50 documents from the search result of one keyword. Table 5 shows the details of the created data.

After creating the large corpus, the corpus was used as a source of relevant document retrieval. We used 1000 and 5000 relevant documents as the adaptation data. Table 6 shows the result of language model adaptation after five iterations.

Table 6 shows that the Web-corpus-based adaptation gave the best improvement compared with the adaptation based on the small corpus or the Web search. Comparing the result of the first adaptation and the fifth iteration using 1000 adaptation documents, improvement of the word accuracy, adjusted perplexity and OOV improved by 0.7 point, 2.22 and 0.06 point, respectively. When 5000 documents were used, the improvements were 0.67 point, 3.93 and 0.06 point, respectively.

Table.5 Details of the Downloaded Data

Target number of keywords	287715 Nouns
Download URL	50 URL/query
Download epoch	Feb. 2010-Apr. 2010
Total number of documents	About 15 million
Total number of words	About 10 billion
Total number of distinct words	395809

IV. CONCLUSIONS

To realize highly accurate speech recognition by retrieval of relevant documents using the Web, we examined unsupervised language model adaptation using Web documents. By repeating adaptation procedure using relevant documents, we confirmed the improvement in recognition performance.

In the experiment, the largest improvement was obtained when adaptation was performed repeatedly using 5000 documents selected from the downloaded data. Here, the word accuracy, adjusted perplexity and OOV was improved by 4.1 points, 5.6 and 1.3 point, respectively.

As a future work, we will collect relevant documents using OOV candidates as search queries. We expect further improvement of recognition performance using this method.

Table.6 Adaptation Results (Downloaded Data)

iteration	Cor[%]	Acc[%]	APP	OOV[%]
baseline	67.51	62.45	101.5	1.54
1000-1	70.73	66.67	81.7	0.33
1000-2	71.37	67.33	79.48	0.27
1000-3	71.38	67.38	79.55	0.27
1000-4	71.39	67.39	79.43	0.27
1000-5	71.38	67.37	79.48	0.27
5000-1	70.95	67.36	81.15	0.27
5000-2	71.51	67.94	78.2	0.22
5000-3	71.61	68.02	77.3	0.21
5000-4	71.63	68.04	77.19	0.21
5000-5	71.61	68.03	77.23	0.21

REFERENCES

[1] A.Ito, M.Kohda, "Evaluation of Task Adaptation Using N-Gram Count Mixture", Trans. IEICE (D-II), Vol.J83-DII, No.11, pp.2418-2427, 2011.

[2] H.Nanjo, T.Kawahara, A.Yamada, K.Uchimoto, "Unsupervised Language Model Adaptation for Lecture Speech Recognition", Tech. Rep. IEICE, NLC2002-75, (SLP44-32), 2002.

[3] A.Sethy, P.G.Georgiou, S.Narayanan, "BUILDING TOPIC SPECIFIC LANGUAGE MODELS FROM WEBDATA USING COMPETITIVE MODELS", In Proc. Interspeech, pp.1293-1296, 2005.

[4] M.Suzuki, Y.Kajiura, A.Ito, S.Makino, "Unsupervised language model adaptation based on automatic text collection from WWW", In Proc. Interspeech, pp.2202-2205, 2006.

[5] Ivan Bulyko, Mari Ostendorf and Andreas Stolcke, "Getting More Mileage from Web Text Sources for Vondversational Speech Language Modeling using Class-Dependent Mixtures", In Proc. HLT-NAACL, 2003.

[6] Y.Ito, H.Nishizaki, Y.Sekiguchi, "Improving recognition performance of spoken documents using similar documants on the Internet", Tech. Rep. IEICE, NLC2005-65, 2005.

[7] Y.Nemoto, T.Kawahara, Y.Akita, "Automatic Lecture Transcription by Exploiting Slide Information for Language Model Adaptation", Tech. Rep. IEICE, NLC2007, 2007.

[8] "Yahoo! JAPAN", http://www.yahoo.co.jp/.

[9] "Yahoo! Developer's network", http://developer.yahoo.com/.

[10] A.Lee, T.Kawahara and K.Shikano, "Julius – an open source real-time large vocabulary recognition engine", Proc. Eurospeech, pp.1691-1694, 2001.

[11] "CD – Mainichi Data Collection", http://www.nichigai.co.jp/sales/mainichi/mainichidata.html.

[12] A.Ito, Y.Kajiura, M.Suzuki and S.Makino, "Automatic query generation and query relevance measurement for unsupervised language model adaptation of speech recognition", EURASIP Journal on Audio, Speech, and Music Processing, 2010.

[13] C.J.Van Rijsbergen, Information Retrieval, Butterworth-Heinemann, Newton, Mass, 1979.

Two-Dimensional Simulation Approach on the Mechanism of Microwave Induced Thermo-Acoustic System

Jian Song, Zhiqin Zhao
School of electronic engineering
University of Electronic Science and Technology of China, UESTC
Chengdu, China
songjian1817@gmail.com

Abstract—An approach based on the combination of Finite-Difference Time-Domain (FDTD) and *K*-wave is developed to simulate thermo-acoustic signal production and propagation in microwave induced thermo-acoustic tomography (MITAT) system. In the MITAT system, the relationship between electric field and acoustic field is a very important issue because of its multiple physics mechanism. For the thermo-acoustic signal production, the rationality to make the Specific Absorption Rate (SAR) values as initial pressure value of the thermo acoustic signal is discussed. The relationship between the sample boundaries and the amplitude of the signal is investigated. Two processes, the absorption of microwave energy which is realized by FDTD and the acoustic propagation which is realized by *K*-wave, are involved in the simulation. Some discussions are made based on the simulation results.

Keywords-microwave-induced thermo-acoustic; FDTD ; acoustic propagation; K-wave

I. INTRODUCTION

The microwave induced thermo-acoustic tomography (MITAT) is an innovative technique for early breast tumor detection [1]. It utilizes the modulated microwave pulse to irradiate on a biologic tissue, then thermo-deposition is produced due to the electromagnetic energy absorbed by the biologic tissue. For the malignance, large magnitude thermo-acoustic signal is produced due to the higher permittivity and conductivity than normal tissue. The mechanism of microwave induced thermo-acoustic is an important part of MITAT. A model of the electromagnetic and acoustic wave propagation using the Finite-Different Time-Domain is presented [2]. With the time steps increasing, the error of FDTD is accumulated and affects the following accuracy. Additionally, the relationship between thermo-acoustic signal and sample boundaries has not been reported. However, a relationship between the thermo-acoustic waveform and sample boundaries is investigated in our work.

The program includes two parts, the first part is the process of absorption of microwave energy and the second one is acoustic propagation process. In the first process, the distribution of electric fields in sample zone is obtained and then the SAR values are calculated. Then the SAR values are applied to the initial acoustic pressure [3] and the propagation process is simulated by *K*-wave [4] method. In [2], the propagation process of acoustic is implemented by FDTD. *K*-wave method has better computing speed and accuracy than FDTD.

The research emphasis is the multiple physics mechanism in MITAT system. The research on mechanism makes the physical processes of microwave-induced thermo-acoustic clearer and accelerates the study of the MITAT.

The remainder of the paper is organized as follows. In section Ⅱ, the theory of MITAT is discussed and the mathematical derivation form electric field to acoustic field is finished. In section Ⅲ, we present and analysis the simulation results and then compare with the experiment result. Section Ⅳ is the conclusion.

II. THEORETCIAL ANSLYSIS OF MITAT

Microwave-induced thermo-acoustic can be summarized as two processes: the unequal absorption of microwave energy due to different biological tissues; the other one is acoustic production due to thermal expansion caused by the heat energy. Microwave-induced thermo-acoustic includes the electromagnetic energy to heat, heat to pressure two physical processes. There is the conversion among the three physical quantities that is electric fields, heat energy and acoustic pressure. Chen [1] uses dynamic model and the generalized Hooke theorems to obtain the relationship of acoustic pressure *P* and normal stress σ_{ii} such as (1) shows.

$$\mathrm{P}_i = \sigma_{ii} \bullet \vec{n}, \qquad (1)$$

where $\vec{n}$ is the unit normal vector in *i* direction, *i* is *x, y* or *z*.

The acoustic pressure *P* depends on the medium increased temperature *T* due to the microwave heating [1]. Assuming that there is no thermal diffusion in the point source, the thermal diffusion and heat conduction equation is shown by

$$\nabla^2 T(\vec{r},t) + \frac{h(\vec{r},t)}{\lambda} = \frac{1}{k}\frac{\partial T(\vec{r},t)}{\partial t}, \qquad (2)$$

where k is the coefficient of thermal diffusion, $h(\vec{r},t)$ is the heat function. It also can be written as (3)

$$h(\vec{r},t) = dW/(\rho dV) \quad . \tag{3}$$

The width of the electromagnetic pulse used in MITAT is often 0.2~1.0 microsecond and the thermal diffusion in tissue can be neglected. The first part of (2) is zero, i.e., $\nabla^2 T(\vec{r},t)=0$. Equation (4) can be obtained from the second derivative with respect to time of (2).

$$\nabla^2 \mathrm{P}(\vec{r},t) - \frac{1}{c^2}\frac{\partial^2 \mathrm{P}(\vec{r},t)}{\partial t^2} = \frac{k}{2R_T}\frac{1}{C_V}\frac{\partial h(\vec{r},t)}{\partial t} \quad , \tag{4}$$

where R_T is infinitesimal torque, C_V is material specific heat capacity, k is the thermo diffusion coefficient.

By solving second order linear differential equations, the microwave-induced thermo-acoustic source solution is derived.

$$\mathrm{P}(\vec{r},t) = \frac{k}{2R_T}\frac{1}{C_V}\iiint_{|\vec{r}-\vec{r}'|=t\bullet c}\frac{d\vec{r}'}{|\vec{r}-\vec{r}'|}\frac{\partial h(\vec{r}',t)}{\partial t}, \tag{5}$$

where $\vec{r}$ is the field location, $\vec{r}'$ is the source location.

In the lossy medium, the absorption of electromagnetic energy can be measured by power absorption rate per unit mass i.e., Specific Absorption Rate (SAR).

$$SAR = \frac{d}{dt}\left(\frac{dW}{\rho dV}\right) = \frac{\partial h(\vec{r}',t)}{\partial t} = \frac{\sigma}{2\rho}\left|\vec{E}\right|^2 \quad . \tag{6}$$

Equation (7) can be obtained which is based on the combination of (5) and (6),

$$\mathrm{P}(\vec{r},t) = \frac{k}{2R_T}\frac{1}{C_V}\iiint_{|\vec{r}-\vec{r}'|=t\bullet c}\frac{\overrightarrow{dr'}}{|\vec{r}-\vec{r}'|}\bullet SAR = \frac{k}{2R_T}\frac{1}{C_V}\iiint_{|\vec{r}-\vec{r}'|=t\bullet c}\left|\vec{E}(\vec{r}',t)\right|^2\frac{\overrightarrow{dr'}}{|\vec{r}-\vec{r}'|} \quad . \tag{7}$$

The connection between electric fields and acoustic pressure is built across to (7). Through (5), (6) and (7), it is reasonable for the SAR values are equivalence as the acoustic source.

III. SIMULATION RESULTS

In order to investigate the effect of the specimen itself on the signal, the comparison of the signal from the specimen and ones without from specimen is carried out. The time delay depends on the distance between the specimen and probe is the main criterion to judge the thermo-acoustic signal which is from the specimen. The FDTD method is employed to simulate the electromagnetic fields distribution in the two-dimensional surface which the specimen is involved. The process of acoustic propagation is simulated by the *K*-wave method.

A. Time Delay Simulation

In this part, the time delay corresponding to the different specimen and probe distances is simulated. Two simulations with different distances are carried out. The distances between the sample and probe are 5.0 cm and 3.0 cm, respectively. The acoustic speed is set to 1500 m/s, the calculated arrival time is 3.3×10-5s and 2.0×10-5s. In Fig.1 (a) the start time of signal is about 3.3×10-5s and the start time of Fig.1 (b) is 2.0×10-5s. The simulation result matches well with the theoretical results. Further more, the time delay depends on the specimen and probe distance is validated.

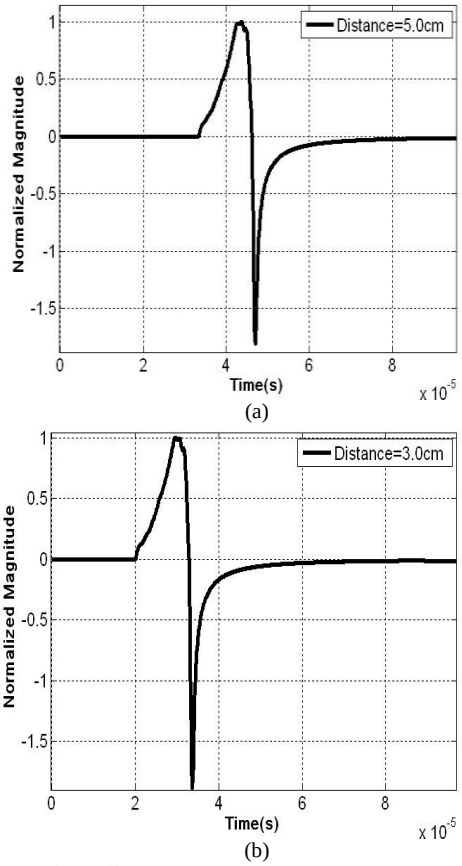


Fig.1 Signals for different distances. (a) is the acoustic signal for the distance between sensor and sample is 5.0 cm. (b) is the acoustic signal for the distance between sensor and sample is 3.0 cm.

B. Relationship Between Sample and Signal

In this part, the relationship between sample and signal is studied, especially the corresponding relationship between the sample boundary and signal amplitude. From (7), the acoustic pressure P has relationship with the electric fields in the sample zone. In the homogeneous medium, the acoustic pressure value of a particle depends on the electric fields in the sample zone.

The *K*-wave method is used to simulate the process of acoustic propagation. The one of basis equations of acoustic propagation is show in (8).

$$-\nabla \mathrm{P} = \frac{\partial u}{\partial t}, \tag{8}$$

where u is the particle velocity.

From (8), the time derivative of particle velocity equals the gradient of acoustic pressure. Moreover, the particle which has the fast pressure space change rate is the one who has the largest acceleration. In the waveform, the peaks are produced.

Figure 2 shows the kinds of signals obtained from the various samples.

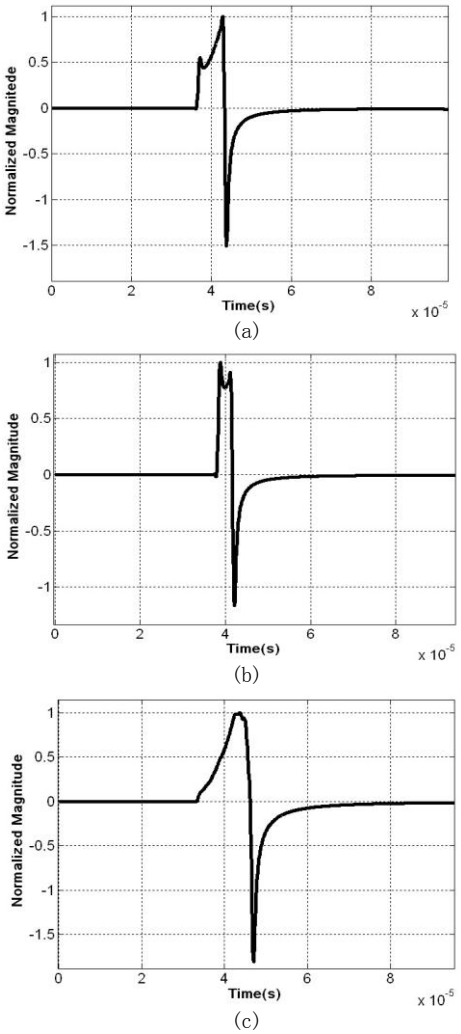


Fig.2 Thermo-acoustic signals for various samples. (a) is the signals for the rectangle samples with the size of 1.0cm×1.0cm, (b) is the signals for the rectangle samples with the size of 0.5cm×1.0cm, (c) is the signal for circle sample with radius=1.0cm.

In Fig.2, the distances between sample boundaries and probes are 5.5 cm, 5.75 cm and 5 cm, respectively. (a), (b), (c) are the corresponding signals. The rising edge appears at the moment that is corresponding with the sample boundary.

The pressure distribution curve along the acoustic propagation direction within the sample is shown in Fig.3.

In Fig.3, the acoustic pressure at the boundary location changes dramatically whatever in the right boundary or left one. The pressure change rate is flat within the sample. Across (8), the dramatic space change rate of pressure induces the large particle acceleration, i.e., the sharp rising edge of waveform. That reason is why the existence of boundaries induces signal peaks.

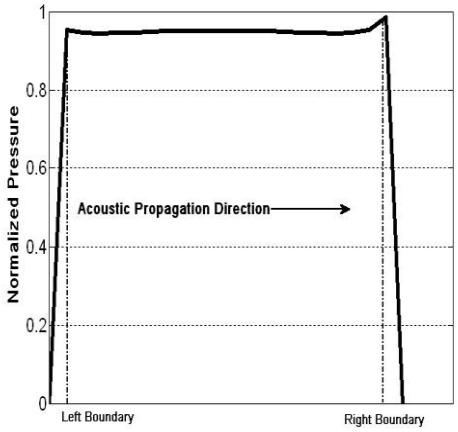


Fig.3 Pressure Distribution Curve along the Acoustic Propagation Direction

C. Experiment Result

One of the experiment results is shown in Fig.4(a). In the experiment, the size of rectangle sample is 0.5 cm×1.0 cm. In Fig.2 (b), the comparison of the signals from simulation and experiment is carried out. The experiment signal has oscillating tails compared with the one in simulation. The reason of this phenomenon is the acoustic wave reflection due to the sample container. Additionally, the heterogeneity within the specimen also causes this condition. The experiment signal matches well with the simulation signal on the signal waveform whether the oscillating tails are not considered.

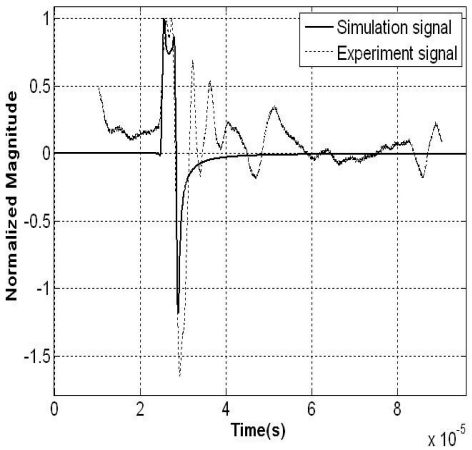


Fig.4 Comparison of signals from simulation and experiment

IV. CONCLUSION

A hybrid approach based on the combination FDTD and *K*-wave is proposed to simulate thermo-acoustic production and propagation. The mechanism of MITAT system is discussed from the view of multiple physics. The rationality to make the Specific Absorption Rate (SAR) value as initial pressure value of the thermo acoustic signal is verified. The existence of the sample boundaries induces the signal peaks. This conclusion provides the basis for imaging the sample contour used by thermo-acoustic signal. Further investigation in the condition of the layered sample is underway, and we hope to report the progress in the near future.

ACKNOWLEDGMENT

This work was supported by National Natural Science Foundation of China (No. 60927002）.

REFERENCES

[1] G. P. Chen, Z. P. Nie, "Critical technologies research of the microwave-induced thermo-acoustic tomography system", Dissertation, University of Electronic Science and Technology of China,UESTC,2009

[2] K. G. Zhu, M. Popovic, "FDTD Electromagnetic-Acoustic Model: A 2-D Numerical Coding Framework," IEEE APSURSI,PP.1-4,2010

[3] Y. Xie, B. Guo, J. Li, G. Ku, L. V. Wang, "Adaptive and Robust Methods of Reconstruction (ARMOR) for Thermo-acoustic Tomography," IEEE Trans. Biomed. Eng., vol. 55, pp. 2741-2752, 2008.

[4] B. E.Treeby, B. Y.Cox, "k-Wave: MATLAB toolbox for the simulation and reconstruction of photo-acoustic wave fields," Journal of Biomedical Optics. 2010

[5] D. Feng, Y. Xu, G. Ku, and L. V. Wang, "Microwave-induced thermo-acoustic tomography: Reconstruction by synthetic aperture," Med. Phys., vol. 28, pp. 2427-2431, Dec. 2001.

[6] Y. Xu, D. Feng, and L. V. Wang, "Exact frequency-domain reconstruction for thermo-acoustic tomography I. Planar geometry," IEEE Trans. Med. Imag., vol. 21, no. 7, pp. 823-828, Jul. 2002.

[7] Y. Xu, M. Xu, and L. V. Wang, "Exact frequency-domain reconstruction for thermo-acoustic tomography: II. Cylindrical geometry," IEEE Trans. Med. Imag., vol. 21, no. 7, pp. 829-833, Jul. 2002.

[8] M. Xu and L. V. Wang, "Time-domain reconstruction for thermo-acoustic tomography in a spherical geometry," IEEE Trans. Med. Imag., vol. 21, no. 7, pp. 814-822, Jul. 2002.

[9] M. Xu, Y. Xu, and L. V. Wang, "Time-domain reconstruction algorithms and numerical simulations for thermo-acoustic tomography in various geometries," IEEE Trans. Biomed. Eng., vol. 50, no. 9, pp. 1086-1099, Sep. 2003.

[10] G. Ku, B. D. Fornage, X. Jin, M. Xu, K. K. Hunt, and L. V. Wang, "Thermo-acoustic and photo-acoustic tomography of thick biological tissues toward breast imaging," Technol. Cancer Res. Treat., vol. 4, pp. 1-7, Oct. 2005.

[11] X. Jin, and L. V. Wang, "Thermo-acoustic tomography with correction for acoustic speed variations," Phys. Med. Biol., vol. 51, pp. 6437-6448, 2006.

[12] J. H. Kim, S. B. Park, and S. A. Johnson, "Tomography imaging of ultrasonic reflectivity with correction for acoustic speed variations," Ultrasonic. Imaging., vol. 6, pp. 304-316, 1984.

[13] J. R. Jago, T. A. Whittingham, "Experimental studies in transmission ultrasound computed tomography," Phys. Med. Biol., vol. 36, pp. 1515-1542, 1991.

[14] G. P. Chen, Z. Q. Zhao, Z. P. Nie, Q. H. Liu, "A computational study of time reversal mirror technique for microwave-induced thermo-acoustic tomography," Journal of Electromagnetic Waves and Applications., vol. 22, no. 12, pp. 2191-2204, 2008.

[15] G. P. Chen, W. B. Yu, Z. Q. Zhao, Z. P. Nie, Q. H. Liu, "The prototype of microwave-induced thermo-acoustic tomography imaging by time reversal mirror," Journal of Electromagnetic Waves and Applications., vol. 22, no. 11, pp.1565-1574, 2008.

A Tight Coupling Directional Coupler from Composite Right/Left-Handed Transmission Lines

Hao Wang and Xiaohong Tang
School of Electronic Engineering
University of Electronic Science and Technology of China
Chengdu 610054, China

Hong Zhang
College of Electronics and Information Engineering
Sichuan University
Chengdu 610064, China

***Abstract*—The coupled line structure of the directional coupler designed using composite right/left handed transmission lines (CRLH TLs) is presented. The directional coupler, composed of two identical CRLH TLs, has tight coupling characteristic. Measured results show that the coupling of the directional coupler is 3dB from 3.3 to 3.6 GHz. The coupling can be adjusted by varying the gap between coupled CRLH TLs or the number of unit cell of CRLH TLs.**

Keywords-Directional coupler; tight coupling; CRLH TLs

I. INTRODUCTION

Conversional directional couplers with coupled line structures are usually used for weak couplings. But as for tight coupling cases, they require lines that are too close together to be practical. Lange couplers can provide tight coupling, but it is difficult to fabricate the necessary bonding lines [1]. In order to achieve tight coupling, CRLH TLs have been applied to design coupled line directional couplers successfully [2] [3].

In this paper, a directional coupler based on CRLH TLs is analyzed, simulated and measured. The simulated and measured results show that the directional coupler has tight coupling characteristic. The directional coupler can even achieve arbitrary coupling level by varying the gap between coupled CRLH TLs or the number of unit cells of CRLH TLs.

II. DESIGN

A. CRLH TL

The LC-based CRLH TL is shown in Fig. 1 and the equivalent circuit of unit cell of the LC-based CRLH TL, shown in Fig. 2, consists of an inductance L_R in series with a capacitance C_L and a shunt capacitance C_R in parallel with an inductance L_L [4]. For the balanced case, the series and shunt resonances are equal

$$L_R C_L = L_L C_R \tag{1}$$

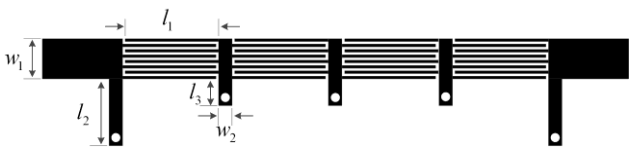


Figure 1. Four-cell of LC-based CRLH TL

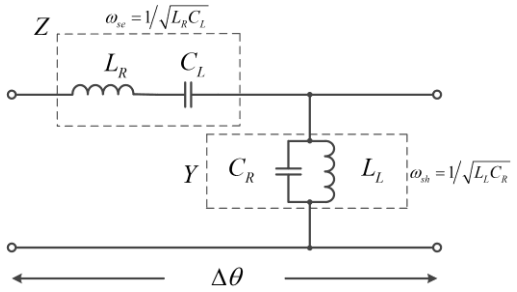


Figure 2. Equivalent circuit of unit cell of the LC-based CRLH TL

The characteristic impedance of the CRLH TL is given by

$$Z_0 = \sqrt{Z/Y} = \sqrt{L_L/C_L} = \sqrt{L_R/C_R} \tag{2}$$

The propagation constant of the CRLH TL is

$$\beta = \beta_R + \beta_L = \omega\sqrt{L_R C_R} - \frac{1}{\omega\sqrt{L_L C_L}} \tag{3}$$

The transition frequency is given by

$$\begin{aligned} \omega_0 &= 1/\sqrt[4]{L_R C_R L_L C_L} \\ &= \sqrt{\omega_R \omega_L} \\ &= \sqrt{\omega_{se}\omega_{sh}} \end{aligned} \tag{4}$$

where $\omega_R = 1/\sqrt{L_R C_R}$, $\omega_L = 1/\sqrt{L_L C_L}$, $\omega_{se} = 1/\sqrt{L_R C_L}$, $\omega_{sh} = 1/\sqrt{L_L C_R}$.

Left handed cutoff frequency ω_{cL}^{Bal} and right handed cutoff frequency ω_{cR}^{Bal} are given by

$$\omega_{cL}^{Bal} = \omega_R\left(\sqrt{1+\frac{\omega_L}{\omega_R}} - 1\right) \tag{5}$$

$$\omega_{cR}^{Bal} = \omega_R\left(\sqrt{1+\frac{\omega_L}{\omega_R}} + 1\right) \tag{6}$$

The fractional bandwidth can be written as

$$FBW = 2\frac{\omega_{cR} - \omega_{cL}}{\omega_{cR} + \omega_{cL}} \tag{7}$$

If the number of cells N is chosen for some specific applications, the CRLH TL will be determined ultimately.

A cylindrical via hole in microstrip shown in Fig. 3 will generate inductance effect. The inductance can be represented as [6]

$$L_{via} = \frac{\mu_0}{2\pi}\left[h \cdot \ln\left(\frac{h+\sqrt{r^2+h^2}}{r} \right) + \frac{3}{2}\left(r - \sqrt{r^2+h^2} \right) \right] \tag{8}$$

where $\mu_0 = 4\pi \times 10^{-7}\,(H/m)$ is the permeability of free space, h is substrate height, and r is the radius of the via hole.

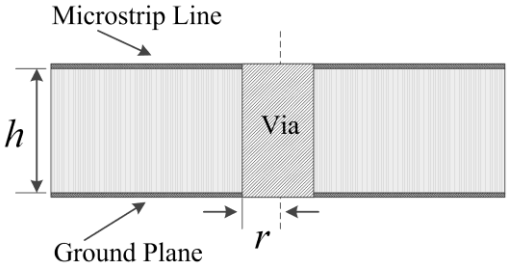


Figure 3. Side view of a cylindrical via hole in microtrip

B. Directional Coupler Based on CRLH TLs

The directional coupler shown in Fig. 4 is composed of two identical CRLH TLs. Generally, the electrical length of the CRLH TL is not quarter wavelength. Here, $\vec{k}$ denotes wave number and $\vec{S}$ is Poynting vector. It can be seen that the direction of energy is antiparallel to the direction of wave on the CRLH TL. Power supplied to port 1 is coupled to port 3, while the remainder of the input power is delivered to port 2 and port 4. In an ideal directional coupler, no power is delivered to port 4.

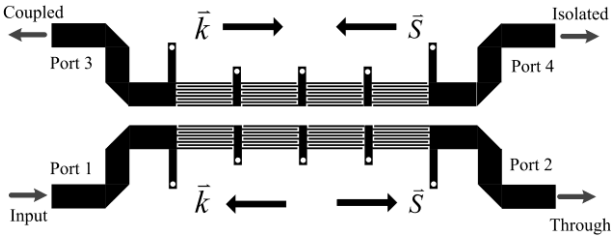


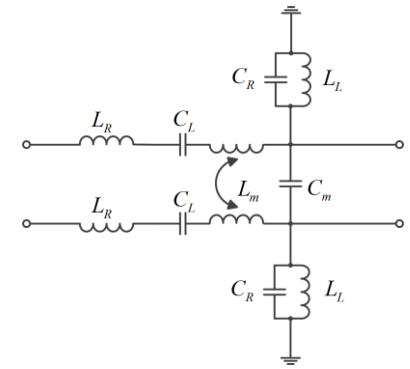
Figure 4. Layout of the directional coupler

The equivalent circuit of the directional coupler is shown in Fig. 5 with decomposition into the corresponding even and odd equivalent circuits. L_m and C_m are mutual inductance and coupling capacitance, respectively. Therefore, the even series and shunt resonance frequencies are [3]

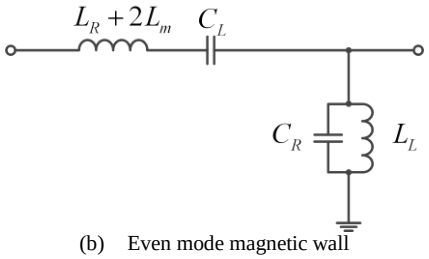
$$\omega_{se,e} = \frac{1}{\sqrt{(L_R + 2L_m)C_L}} \tag{9}$$

$$\omega_{sh,e} = \frac{1}{\sqrt{L_L C_R}} \tag{10}$$

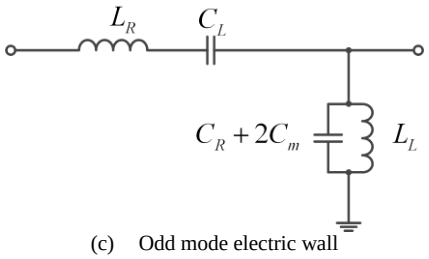
The odd series and shunt resonance frequencies are



(a) Equivalent circuit



(b) Even mode magnetic wall



(c) Odd mode electric wall

Figure 5. Equivalent circuit for unit cell of the directional coupler

$$\omega_{se,o} = \frac{1}{\sqrt{L_R C_L}} \tag{11}$$

$$\omega_{sh,o} = \frac{1}{\sqrt{(C_R + 2C_m)L_L}} \tag{12}$$

The even characteristic impedance is given by

$$Z_{0e} = \sqrt{\frac{L_L}{C_L}}\sqrt{\frac{1-\omega^2(L_R+2L_m)C}{1-\omega^2 L_L C_R}} \tag{13}$$

The odd characteristic impedance is

$$Z_{0o} = \sqrt{\frac{L_L}{C_L}}\sqrt{\frac{1-\omega^2 L_R C_L}{1-\omega^2 L_L (C_R+2C_m)}} \tag{14}$$

The coupling coefficient is

$$C = \frac{(Z_{0e}-Z_{0o})\tanh\left[(\alpha+j\beta)l\right]}{2Z_0 + (Z_{0e}+Z_{0o})\tanh\left[(\alpha+j\beta)l\right]} \tag{15}$$

III. SIMULATION AND MEASUREMENT

As shown in Fig. 1, the parameters of the CRLH TL are given: $w_1 = 3.0mm$, $l_1 = 7.0mm$, $w_2 = 1.0mm$, $l_2 = 8.0mm$, $l_3 = 3.8mm$. The radii of via holes are 0.25mm. The substrate used in the

design is F4B-2 with a relative dielectric constant $\varepsilon_r = 2.65$, a thickness of 1.0 mm, and dissipation factor $\tan\delta = 0.0009$.

The directional coupler is simulated using the full-wave electromagnetic simulation package IE3D. Some simulated results of coupling levels of the CRLH directional coupler versus gaps are shown in Table I, where the coupling levels of the conventional coupled line directional coupler with corresponding gaps are also shown for comparison. From the Table I, it can be seen that CRLH directional coupler has tight coupling characteristic, compared to the conventional coupled line directional coupler.

TABLE I
COUPLING LEVELS VERSUS GAPS FOR THE 4 CELLS CRLH DIRECTIONAL COUPLER AND CONVENTIONAL COUPLED LINE DIRECTIONAL COUPLER

Gap (mm)	Coupling of CRLH Couplers (dB)	Coupling of Conventional Couplers (dB)
0.3	3.0	14.6
0.8	6.0	18.7
1.5	10.0	23.4
4.0	20.0	34.7

The size of the fabricated directional coupler shown in Fig. 6 is 63.8mm × 24.3mm. The gap between the coupled CRLH TLs is 0.3mm corresponding to 3dB coupling.

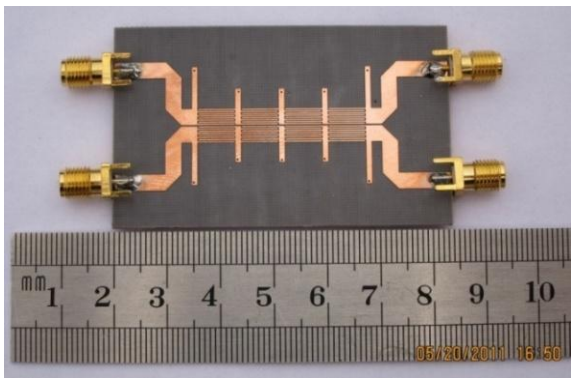
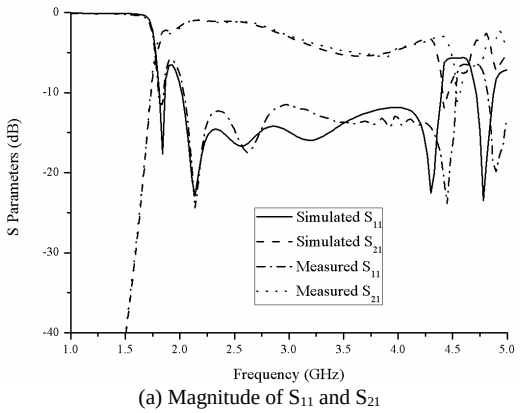
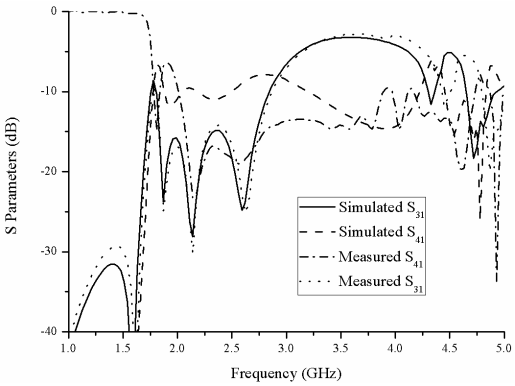


Figure 6. Picture of the fabricated directional coupler

The fabricated directional coupler is measured using a vector network analyzer (Aglient N5230). The simulated and measured results are shown in Fig. 7. The measured results



(a) Magnitude of S_{11} and S_{21}



(b) Magnitude of S_{31} and S_{41}

Figure 7. Simulated and measured S parameters

show that 3dB coupling is achieved from 3.3-3.6 GHz. Coupling flatness is less than 1dB over the designed bandwidth.

IV. CONCLUSION

In this paper, a 3dB directional coupler based on CRLH TLs is analyzed, simulated and measured. This directional coupler can be applied to antenna transceiver switches and microwave power monitoring. In addition, by varying the gap or the number of unit cells of the CRLH TLs, tight or loose coupling can be achieved for some specific applications.

REFERENCES

[1] J. Lange, "Interdigitated Stripline Quadrature Hybrid," IEEE Transactions on Microwave Theory and Techniques, vol. MTT-17, pp. 1150-1151, December 1969.

[2] C. Caloz and T. Itoh, "A Novel Mixed Conventional Microtrip and Composite Right/Left-Handed Backward-Wave Directional Coupler with Broadband and Tight Coupling Characteristics," IEEE Microwave and Wireless Components Letters, vol. 14, pp. 31-33, January 2004.

[3] C. Caloz, A. Sanada and T. Itoh, "A Novel Composite Right-/Left-Handed Coupled-Line Directional Coupler with Arbitrary Coupling Level and Broad Bandwidth," IEEE Transactions on Microwave Theory and Techniques, vol. 52, pp. 980-992, March 2004.

[4] A. Lai, C. Caloz and T. Itoh, "Composite right/left-handed transmission line metamaterials," IEEE Microwave Magazine, pp. 34-50, September 2004.

[5] C. Caloz and T. Itoh, Electromagnetic Metamaterials: Transmission Line Theory and Microwave Applications. pp. 120-121, New York: Wiley, 2004.

[6] Marc E. Goldfarb and Robert A. Pucel, "Modeling Via Hole Grounds in Microstrip," IEEE Microwave and Guided Wave Letters, Vol. 1, NO. 6, pp.135-137, June 1991

RF Magnetic Near Field Measurement using a Magnetic Force Microscope for the Potential Application to Evaluate LTE-Class RFIC

Yasushi Endo[1], Masaaki Fukushima[1], Masatoshi Watanabe[1], Kaoru Arai[1], Yutaka Shimada[1], Masahiro Yamaguchi[1,2]
[1] Department of Electrical and Communication Engineering, Graduate School of Engineering, Tohoku University, Japan
[2] New Industry Creation Hatchery Center, Tohoku University, Japan
{endo, m-fuku, m-wata, k-arai, simayuta, yamaguti}@ecei.tohoku.ac.jp

Abstract— **This paper proposes a new measurement technique of the radio frequency (RF) magnetic near field from a LTE-class RFIC chip employing a cantilever tip fabricated for magnetic force microscope as a sensor. When both an AC and amplitude modulated currents with a sweeping-frequency were applied to a coplanar waveguide (CPW) and a MFM tip approached the CPW surface, the magnetic force interaction between the tip and the RF field occurred, and the tip apparently oscillated with maximum amplitude at the resonance frequency of the tip. These results suggest that the measurement system has potential as a micron scale RF field detector to evaluate LTE-class RFIC, but a few technical problems must be solved.**

Keywords-magnetic force microscope; coplanar waveguide; RF magnetic near field; amplitude modulation;

I. INTRODUCTION

Recently, the cellular phones and the laptop computers are progressing continuously, for the realization of larger information systems and higher capacitance in communication. In particular, the semiconductor design rules are miniaturizing with lighter weight, further downsizing, and more functions in cellular phones. ITRS reported that the design rule in radio frequency integrated circuits (RFIC) arrives at 65 nm with a minimum pitch of 130 nm in 2007 [1]. When the pitch gets shorter, crosstalk appears among the lines and the electromagnetic noise problems become more distinct. The noise moves from the digital circuits to the analog circuits on RFIC chips which process radio-frequency signals, then the electromagnetic immunity (EMI) problem gets more serious [2]. To enable the communication with a sufficient receiving sensitivity, it is necessary to identify the origin, propagation, and mixing of noise in RFIC chips [3]. Therefore, noise measurement techniques with a higher spatial resolution and sensitivity are required.

In the previous works, some measurement methods of RF magnetic near field on RFIC chips were reported. For example, superconducting quantum interference device (SQUID) can measure with the spatial resolution of approximately 50 μm [4]. The technique using shielded loop coils can measure with the spatial resolution of approximately 40 μm, and measurement frequency of 1 GHz [5]. However, a trade-off between the spatial resolution and the detection sensitivity exists in such techniques. On the other hand, magneto-optic probes, which can measure magnetic near field on IC chips [6], have difficulty to achieve high sensitivity because the magento-optic crystals need to be bigger without considerable improvement in the spatial resolution.

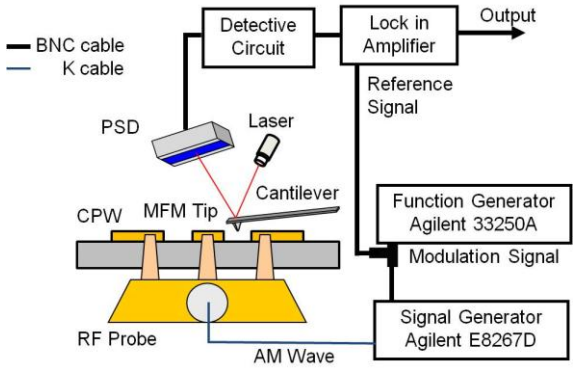


Fig. 1. Schematic diagram of the measurement system of the magnetic near field on a CPW.

This research focuses on a magnetic force microscopy (MFM) as a new measurement technique of RF magnetic near field measurement on RFIC chips with a submicron resolution [7]. In particular, a high frequency MFM (HF-MFM) is applied to evaluate magnetic field originated from a write head of hard disk drive (HDD) [8-9]. The frequency of the signal and the spatial resolution are 2 GHz and 20 nm, respectively [10]. Furthermore, the magnetic near field measurement on transmission line using a MFM tip is suggested. The measurement of RF magnetic near field on a CPW, which is a common transmission line on RFIC chips, was already reported in our previous work [7]. The goal of this research is the measurement of RF magnetic near field on RFIC chips. As for the basic study, an amplitude modulated (AM) current with the carrier frequency, f_c, of 0.01-40 GHz was supplied to the CPW and the magnetic near field generated on the CPW was successfully measured.

II. EXPERIMENTAL PROCEDURE

A. *Measurement system* [11]

The RF magnetic near field from a CPW was evaluated by the oscillation of a MFM cantilever tip at the mechanical resonance frequency. As shown in Fig. 1, an AC current or an amplitude modulated (AM)-current was supplied to a CPW through a ground-signal-ground pin type wafer probe. In the case of AM current, the carrier and the modulated frequency were 0.01-40 GHz and 23-28 kHz, respectively. A reference signal was supplied by a lock in amplifier. When the MFM tip

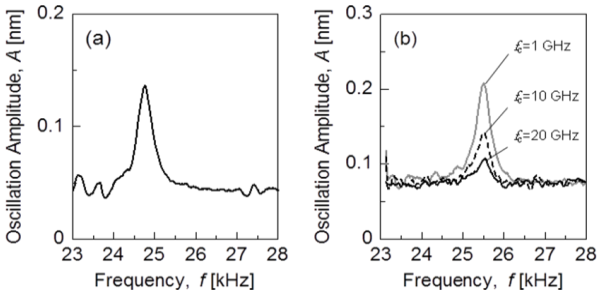


Fig 2. Resonance amplitude of a MFM tip as a function of frequency. (a) An AC current of 31.6 mA is supplied to the CPW. (b) An AM-modulated current of 14.1 mA is applied to the CPW at the carrier frequency f_c of 1, 10, and 20 GHz.

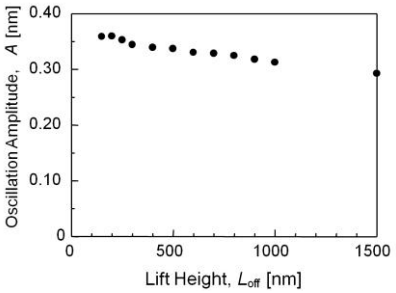


Fig 3. Change in the oscillation amplitude A of the MFM tip at the resonance frequency of the tip with a lift height L_{off}.

approaches the CPW surface, the magnetic interaction between the tip and RF magnetic near field occurs, and the tip oscillates with maximum amplitude at the resonance frequency of the tip. The oscillation amplitude of the tip changing with frequency was detected by a position sensitive detector (PSD) and the signal was analyzed by a lock-in amplifier. The measurement error for the oscillation amplitude of the tip was evaluated as approximately ±1.2 % as all curves of oscillation amplitude vs frequency were measured with a very narrowband-width of 1 Hz. Figure 2 shows the examples of the measurements where an oscillation peak is clearly seen at the resonance frequency of the tip (24.7 kHz). Note that the cantilever was driven solely by the RF near field and other physical forces were not applied.

Although the interaction area of the tip with the magnetic field was difficult to determine, the monopole approximation about the tip magnetization provided a relation between the vertical magnetic field (H_z) and the tip parameters as [12]

$$H_z = \frac{1}{\pi\ \mu_0 M_{tip} R^2} \cdot \frac{kA}{2Q}, \tag{1}$$

where A, Q, k, M_{tip} and R are the maximum oscillation amplitude at the resonance frequency of the tip, quality factor, spring constant of the cantilever, tip magnetization and tip apex radius, respectively.

B. *Magnetic force microscope (MFM) tip*

In this measurement, the MFM tip was a Si cantilever coated with a 30-nm-thick CoCrPt film, which had a low magnetic moment and was initially magnetized in the perpendicular direction. The spring constant of the cantilever, quality factor of the tip, and tip apex radius were approximately 1.4 N/m, 100-150, and 35, respectively.

C. *Coplanar waveguide (CPW)*

A CPW was fabricated by electron-beam lithography, DC magnetron sputtering, and a lift-off technique onto glass substrates (permittivity, ε_r=7.0). The CPW was composed of a Cu layer sandwiched in between Cr layers, that is, Cr (5 nm)/Cu (300 nm)/Cr (5 nm)/glass. The signal line, gap, and ground line widths were 5.0, 6.0, and 50.0 μm, respectively. These widths provided the characteristic impedance for the CPW of approximately 44.9 Ω. The length of the CPW was 2.5 mm.

III. THEORETICAL PREDICTION

A. *Magnetic near field on a CPW by an AC current*

When an AC current is supplied to the CPW, the force of the MFM tip, F_M, caused by vertical component of the magnetic field on the CPW is given as

$$F_M = q_{tip} H_z(t) \propto \frac{I}{h^2} \cos(2\pi f_r t), \tag{2}$$

where q_{tip}, H_z, and f_r are the end of the tip, the vertical magnetization, and resonance frequency of the tip, respectively.

On the other hand, the force of the MFM tip, F_E, caused by the electric field on the CPW is given as

$$H_z = \frac{1}{\mu_0 \pi M_{tip} R^2} \frac{\sqrt{2}kA}{Q}, \tag{3}$$

where m, A, h, and V_e are the magnetization of the tip, detection area of electric field above the CPW, lift height of the tip, and difference of voltages between signal line of the CPW and the tip, respectively. As can be seen in Eq. (2) and (3), the MFM tip detects only the RF magnetic near field on the CPW at the resonance frequency of the tip when AC current is supplied to CPW.

B. *Magnetic near field on a CPW by an AM-modulated current*

When an AM-modulated current is supplied to the CPW, a force of the MFM tip, F_M, caused by vertical component of the magnetic field on the CPW is expressed as

$$F_M \propto \frac{I}{h^2} \cos(2\pi f_r t), \tag{4}$$

where h, I, and f_c is the lift height of the tip, amplitude of current, and resonance frequency of the cantilever.

On the other hand, the force of the MFM tip, F_E, caused by electric field on the CPW is expressed as [13]

$$F_E = \frac{1}{2} \varepsilon \frac{A}{h^2} (V_{CPW} - V_{Tip})^2 \cos(2\pi f_r t), \tag{5}$$

where A, h, V_{CPW}, V_{Tip}, and f_c are the detection area of an electric field, lift height, potential of CPW, potential of the cantilever, and resonance frequency of the cantilever, respectively. This case is different from that in an AC current. That is, these formulae show that the tip detects both RF

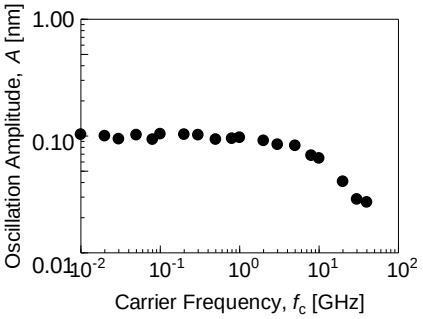


Fig. 4 Oscillation amplitude *A* of the MFM tip at the resonance frequency of the tip as a function of carrier frequency f_c.

magnetic and electric near fields at the resonance frequency of the tip.

IV. Experimental Result and Discussion

A. Measurement of magnetic near field on a CPW by an AC current [7]

Fig. 3 shows the variation in the oscillation amplitude as a function of lift height above the center of the gap on the CPW. The oscillation amplitude is not observed below a height of approximately 150 nm. The reason for this may be that the tip is absorbed onto the substrate surface. The oscillation amplitude appears above a height of 150 nm. The amplitude remains nearly constant at 150-200 nm lift height, but then monotonically decreases above 200 nm. The vertical component of the RF field was evaluated using the monopole approximation shown in Eq. (1). For an AC current of 31.6 mA, the vertical component is approximately 2.4 kA/m at a height of 150 nm.

B. Measurement of magnetic near field on a CPW by an AM-modulated current

Fig. 4 shows the oscillation amplitude of the tip as a function of carrier frequency. The oscillation amplitude was maintained constant up to approximately 5.0 GHz, but then logarithmically decreases above 5.0 GHz. This results may be attributed to the impedance mismatch of the CPW.

Then the oscillation amplitude of the tip was measured at a lift height of 300 nm as a modulation depth of an AM current was varied at the carrier frequencies of 1, 10, and 20 GHz. As shown in Fig. 5, for all carrier frequencies, the oscillation amplitude monotonically increases as the modulation depth increases. This means that the oscillation cycle of the MFM tip follows an envelope one of an AM-modulated field on the CPW. Consequently, the MFM tip oscillates by AM-modulated field on the CPW.

On the other hand, a relation between the lift height and the oscillation amplitude of the tip was measured as a modulation depth of an AM current was fixed as 14.1 mA at the carrier frequencies of 1, 10, and 20 GHz. As seen in Fig. 6, for all carrier frequencies, the oscillation amplitude remains constant as the lift height is varied between 200 to 3000 nm. This is very different from that in the case of an AC current, which may be attributed to both RF magnetic and electric near-fields on the CPW.

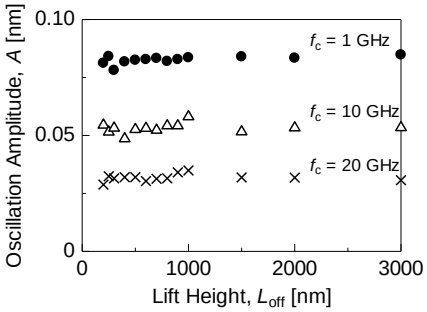


Fig. 5 Change in the oscillation amplitude *A* of the MFM tip at the resonance frequency of the tip with a modulation depth *m*. Solid circles(●), open triangles(Δ), and cross marks(×) represent a carrier frequency f_c of 1,10, 20 GHz, respectively.

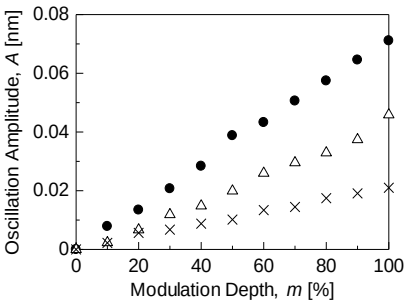


Fig 6. The relationship between the oscillation amplitude *A* of the MFM tip and lift height L_{off}. Solid circles(●), open triangles(Δ), and cross marks(×) represent a carrier frequency f_c of 1, 10, 20 GHz, respectively.

Furthermore, the oscillation amplitude of the tip was measured from the center of one ground line to the center of the other ground line at the lift height of 600 nm. The result is summarized in Fig. 7 as a function of the distance from the center of the signal line (0 nm in this figure). An AM-modulated current of 14.1 mA with *m*=100 % was supplied to the CPW. For all carrier frequencies, the oscillation amplitude monotonically decreases as the distance is varied from -30 to 30 μm. This result agrees with the theoretical prediction in Eq. (5), which may be due to that the cantilever works as both the magnet and plate type condenser [13]. Therefore, this result suggests that the tip interacts closely with both RF magnetic and electric near fields on the CPW as the AM-modulated current in the GHz range is supplied to the CPW.

C. Separation of magnetic near field and electric near field

To separate RF magnetic and electric near fields, our proposed system was improved as shown in Fig. 8. Namely, the improvement points are that each potential of the CPW and the MFM tip is equalized, and that a permanent magnet of 191 kA/m is equipped above the tip to supply DC magnetic field to the MFM tip. The oscillation amplitudes of the tip were measured with DC magnetic field and without DC magnetic field, respectively, as a function of the distance from the center of the signal line. The data can be obtained by subtracting the result without DC field from that with DC field. As can be seen in Fig. 9, the dips are observed at approximately the edge of the ground and the single lines, and are apparently different from that in Fig. 7. This means that the RF magnetic field on CPW

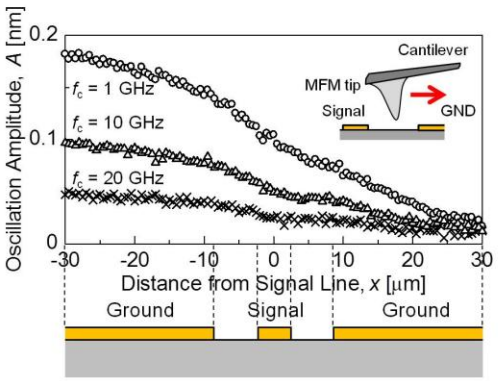


Fig. 7 Change in an oscillation amplitude *A* of a MFM tip with a position on CPW. Open circles(○), open triangles(Δ), and cross marks(×) represent a carrier frequency f_c of 1,10, 20 GHz, respectively.

occurs. Therefore, it is demonstrated that our improved system can measure RF magnetic near field on the CPW.

V. Conclusion

The RF magnetic near field on a CPW has been measured using a commercial type MFM tip. The result provides evidence that the RF magnetic field above 10 MHz can be detected with the high spatial resolution by fabricating a tip specifically designed for this application. This result suggests that this measurement technique may be applicable to measure the RF field distribution near a RF signal and power lines downsized to the micro-scale for a submicron scale IC chips such as LTE-class RFIC.

Acknowledgment

The authors would like to thank Professor S. Sugimoto and Associate Professor N. Tezuka of the Department of Metallurgy, Materials Science, and Materials Processing, Graduate School of Engineering, Tohoku University for performing e-beam lithography. This work was partly supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and by Special Coordination Funds for Promoting Science and Technology from the Formation of Innovation Center for Fusion of Advanced Technologies. This work was partly supported in part by a government project from the Japanese Ministry of Internal Affairs and Communications (Development of Technical Examination Services Concerning Frequency Crowding). This work was also partly supported by Priority Assistance for the Formation of Worldwide Renowned Centers of Research - The Global COE Program (Project: Center of Education and Research for Information Electrics Systems) from MEXT, Japan.

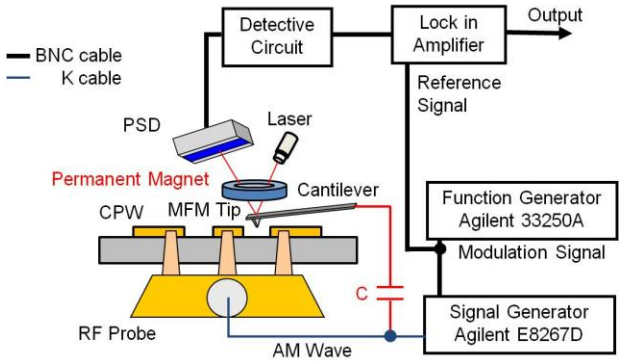


Fig. 8 Schematic diagram of the improved measurement system of the RF magnetic near field on the CPW.

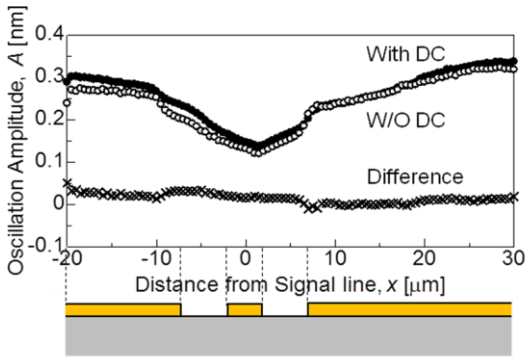


Fig. 9 Oscillation Amplitude *A* of MFM tip on CPW. Solid circles(●), open circles (○), and cross marks(×) represent with a DC magnet, without a DC magnet, and difference of them.

References

[1] International Technology Roadmap for Semiconductors (ITRS) 2010 Edition, p. 7, Figure 4, 2010.

[2] K. Chuang *et al*.,"A 17pJ/bit Broadband Mixed-Signal Demodulator in 90nm CMOS," 2010 IEEE MTT-S International Microwave Symposium, 904 2010.

[3] M. Yamaguchi, Y. Endo, Y. Shimada, K. Arai.,"R&D of IC Chip Level Low Noise Technology to Attain high-Speed and high-Quality Telecommunication," IEEE EMC Society Sendai Chapter Colloquim, CEWS-1-3, 2011.

[4] S. Chatraphorn, E. F. Fleet, and F. C. Wellstood.,"Scanning SQUID microscopy of integrated circuits," Appl. Phys. Lett., **76**, 2304, 2000.

[5] N. Ando.,"Miniaturizaed thin-film magnetic field probe with special resolution for LSI chip measurement," 2004 IEEE International Symposium on Electromagnetic Compatibility, **2**, 357, 2004.

[6] K. Nakajima *et al*.,"Polarization effect in scanning near-field optic/atomic-force microscopy (SNOM/AFM)," Ultramicroscopy, **71**, 257, 1998.

[7] Y. Endo, M. Watanabe, Y. Mitsuzuka, Y. Shimada, and M. Yamaguchi., "Measurement of magnetic near field on a coplanar waveguide using a MFM tip," J. Appl. Phys. **109**, 07D326, 2011.

[8] H. Saito *et al*., "Magnetic force microscopy of alternating magnetic field gradient by frequency modulation of tip oscillation," J. Appl. Phys. **105**, 07D524, 2009.

[9] M. R. Koblischka, J. D. Wei, T. Sulzbach, U. Hartmann., "High frequency MFM characterization of magnetic recording writer poles," Appl. Phys. A. **94**, 235, 2009.

[10] M. R. Koblischka, J. D. Wei, T. Sulzbach, A. D. Johnston, U. Hartmann., "Observation of stray fields from hard disk writer poles up to 2GHz," IEEE. Trans. Magn., **43**, 6, 2007.

[11] Y. Endo *et al*.,"Measurement of Amplitude modulated Electromagnetic near Field Using a MFM Tip," MAG-11-036, The Papers of Technical Meeting on Magnetics, IEE Japan., 2011

[12] Hans J. Hug, B. Stiefel, P. J. A. van Schendel, A. Moser, R. Hofer, S. Martin, and H.-J. Gu¨ntherodt, "Quantitative magnetic force microscopy on perpendicularly magnetized Samples," J. Appl. Phys., **83**, 5609, 1998.

[13] H. Saito, and S. Ishio: "Magnetic Force microscope Analysis of Fine Magnetic Domain Structure and Its Application to high-Density Magnetic Recording Media," J. Magn. Soc. Jpn., **24**, 1221, 2000.

Two elements time domain antenna research based on trapezoidal cavity-backed design

Deqiang Yang, Jin Pan, Pengcheng Li, Zhiqin Zhao, Zaiping Nie

Univerty of Electronic Science and Technology of China, Chengdu, China
Dqyang@ee.uestc.edu.cn

Abstract—An array of two time domain antenna elements is introduced in the paper. With the element of trapezoidal cavity-backed resistance loaded bow tie antenna, it can constitute an array. According the effect of distance between elements, the near-field of Vp-p is accurately measured. Then, the parameter of the array is designed for optimizing the gain. It is pointed that, with this designed method, the gain is significantly increased and the near-field of pattern become narrow. The measurements validate the analysis data.
Keywords: Time domain antenna; trapezoidal cavity-backed element; the near-field of Vp-p

I. INTRODUCTION

In traditional theory and technology about time domain pulse antenna array, the far-field pattern-array scan is already grown up. The near-field research is few reported. Time domain pulse antenna is widely employed GPR(Ground Penetration Radar), UWB(Ultra-wide band) communication, time reverse imaging and concealed target detection[1-3]. But, in the impulse antenna, the most difficult challenges are to preserve the waveform well and keep the trailing small[4-7]. Compared to conventional antenna, the detecting distance is usually limited by the efficiency of time domain antenna. So, enlarging the impulse amplitude and elevating the sensitivity of receiver are usually applied to make up for the disadvantage. Therefore, if the gain is increased directly, the problem of detecting distance can be resolved easily. With wide band of time domain, the time domain antenna array is applied for higher gain. In the paper, the near-field pattern of two elements array will be investigated.

II. THEORETICAL ANALYSIS

Compared to traditional antenna array, the pattern of time domain antenna array should be defined from time domain. So, time-domain antenna array scan is realized through different delay between antenna units. In this paper, two elements are used. With precise delay controlling, the gain of array is significantly elevated. Besides, the pattern becomes narrow.

For two element array, the operating distance is short. So, normal analysis method is not suitable for our test. In the paper, the transceiver distance is limited.

As shown in fig.1, receiver antenna is placed at A. B and C represent transmit element. Between two transmit elements, there is no delay. Because of AB=AC, receive signal can be expressed as equation (1),

$$y(t) = 2 * x(\theta) * z(t) \tag{1}$$

In the equation, $x(\theta)$ represents synthesis of the gain and route loss. $z(t)$ means the signal of antenna element.

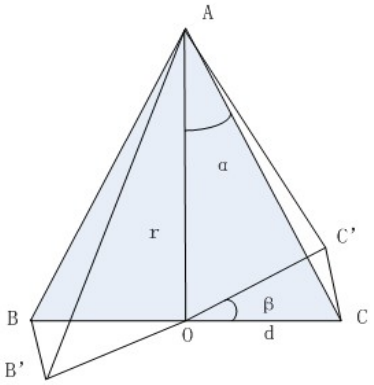


Figure 1 The route of received and transmitted antenna

If the two element array declines the special angle β, the route can be indicated AB'>AC'. Now, at point A, the received signal should be express as equation (3),

$$y_1(t) = x(\theta)z(t) + x(\theta)z(t - \frac{d_{AB'} - d_{AC'}}{c}) \tag{2}$$

In the equation, $d_{AB'}$ and $d_{AC'}$ represent the distance of AB' and AC', respectively. "c" means rate of microwave in free space. With triangle substitution, the distance is calculated,

$$d_{AB'} = \sqrt{d^2 + r^2 + 2dr\sin\beta} \tag{3}$$

$$d_{AB'} = \sqrt{d^2 + r^2 - 2dr\sin\beta} \tag{4}$$

If, delay time t_0 is inserted before the transmitting, t_0 is expressed as follow,

$$t_0 = \frac{d_{AB'} - d_{AC'}}{c} \tag{5}$$

So, the inserted time can offset ahead time, which caused by routine difference. Equation(5) is synchronism requirement. The maximum gain is also obtained by the requirement. It

shows that the higher gain can be acquired by two element time domain antenna array.

III. MEASUREMENT

A. *Receiver Antenna*

The measurement for time-domain antenna can be achieved in ordinary laboratory. For receiving a signal from the two elements array, a trapezoidal cavity-backed resistance loaded bow tie antenna, the similar to transmit element, is used as receiver antenna, just as shown in fig.(2). Receive and transmit antenna both use the vertical polarization.

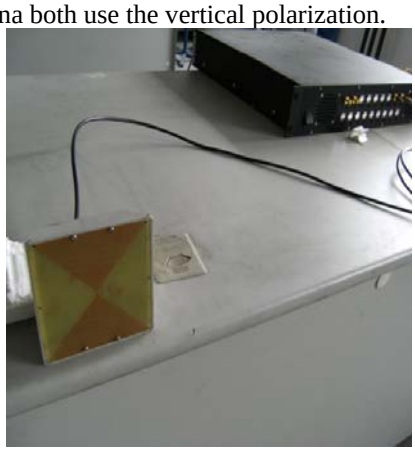


Figure 2 Receiving antenna

Fig.(3) shows the array. Distance between two elements is 20cm. The array is positioned on the antenna turntable rod. It can be rotated with any angle in horizontal plane. They are kept vertical polarization. With dividing the pulse signal from one source, cable's precise length and delay-controller can be used to keep delay of two separated signals same.

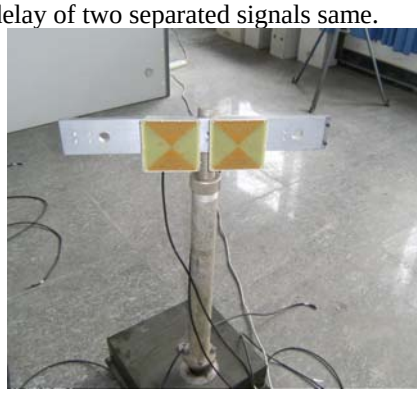


Figure 3 Transmitting two element antenna array

B. *Turntable Controller*

In fig.(4), the turntable controller is connected to computer. By programming, rotating order is sent to turntable controller. With enough driving current, engine could be worked as the angle of rotating order (accuracy=0.1°).

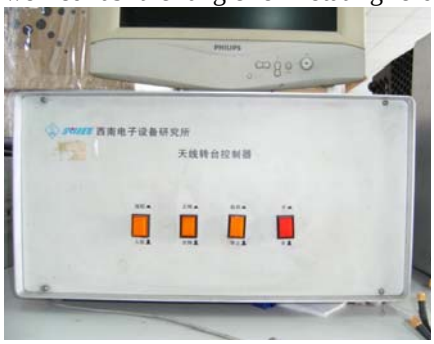
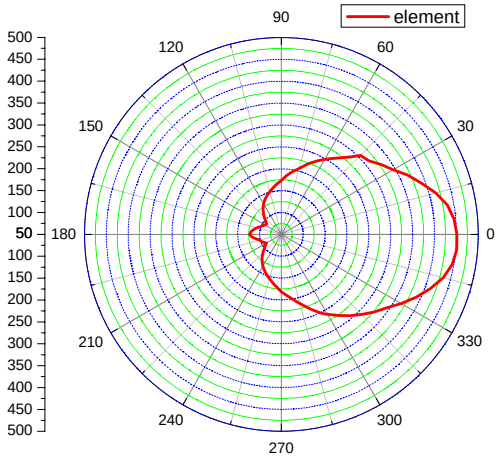


Figure 4 Turntable controller

Fig.(5) shows the digital oscilloscope.(Lecroy company's Wavepro 760Zi). It can quickly catch and analyze the signal in long packet. Its clock precision is 5ppm. All accuracy requires are satisfied by these experiment instrument.

Figure 5 Digital Oscilloscope

C. *Comparison of measurement and simulation*



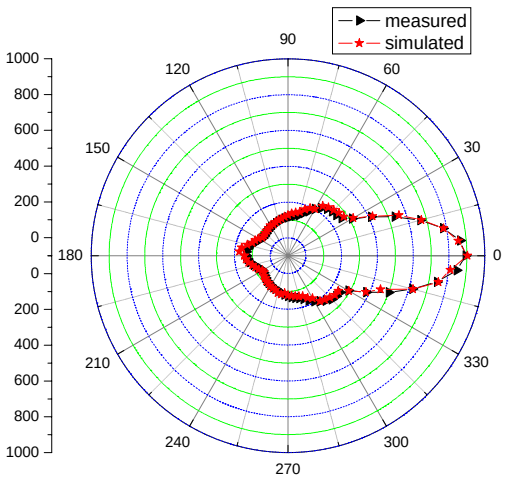


Figure 6 (a) element pattern (b) Simulated and measured results

In fig.7.b, it is shown that simulated result, with calculated by the modified computation, can be validated by the measured one.

IV. CONCLUSION

The array, based the design of trapezoidal cavity-backed resistance loaded bow tie antenna element, has higher gain and more narrow pattern than antenna element. Low efficiency, normally exiting in time domain antenna application, can be resolved by this method of constituting an array. In the paper, an effective technology routine is obtained higher time domain antenna gain. These disadvantages with increasing pulse magnitude have been avoided. If the time domain array is applied for GPR, target detection of longer distance is realized. In next step, the research and exploration will be aimed at application of multi-element time domain antenna array in GPR.

REFERENCES

[1] Zheng, Wenjun, Zhiqin Zhao and Zaiping Nie, “Application of TRM in the UWB Through Wall Radar,” Progress In Electromagnetics Research, PIER 87, pp. 279-296, 2008.

[2] Liu, Dehong, Gang Kang, Ling Li, Ye Chen, Sathyanarayanan Vasudevan, William Joines, Qing Huo Liu, Jeffrey Krolik, and Lawrence Carin, “Electromagnetic time-reversal imaging of a target in a cluttered environment,” IEEE Trans. on Antennas and Propagation, 2005, vol. 53, pp.3058-3066, Sep. 2005

[3] Chen, Ye, William T. Joines, Zhentian Xie, Guining Shi, Qing H. Liu, and Lawrence Carin, “Double-Sided Exponentially Tapered GPR Antenna and Its Transmission Line Feed Structure” ,IEEE Trans. on Antennas and Propagation, Vol. 54, pp.2615-2623, Sep. 2006

[4] Shams, K. M. and M. Ali, “A planar inductively coupled bow-tie slot antenna for WLAN application,” Journal of Electromagnetic Waves and Applications, Vol. 20, No. 7, 861–871, 2006.

[5] Wu, T. T. and R. W. King, The Cylindrical Antenna with Nonreflecting Resistive Loading, IEEE Trans. on Antennas and Propagation, Vol. 13, pp.369-373, 1965.

[6] Pan, Jin, Zai-Ping Nie, “The Loading Formula for V-type Traveling wave Antenna,” Chinese Journal of Radio Science, Vol.11, No.2 June, 1996

[7] D. Yang, J. Pan, Z. Zhao, and Z. Nie, “Design of Trapezoidal Cavity-Backed Resistance Loaded Bow Tie Antenna with Ultra-wideband and High Directivity”, Journal of Electromagnetic Waves and Applications, Vol. 24, No. 8, 1685–1695, 2010.

Analysis of Convergence Characteristic of Adaptive Notch Filters

Atsuhiro UEKI, Shunsuke KOSHITA, Masahide ABE and Masayuki KAWAMATA
Department of Electronic Engineering, Graduate School of Engineering
Tohoku University, Sendai, Japan
Email: {ueki, kosita, masahide, kawamata}@mk.ecei.tohoku.ac.jp

***Abstract*— In this paper, we discuss adaptive notch filters that are used in removal of sinusoidal noises. First, we derive a cost function which is the output variance to an adaptive notch filter with an input sinusoidal signal. Simulation results are presented to confirm the validity of the obtained analytical results of a cost function. Next, we derive the adaptive algorithm without approximation to analyze the convergence characteristic of adaptive notch filters. Finally, we compare the convergence characteristic of the proposed algorithm with that of the RPE algorithm. Simulation results show that the convergence speed of the RPE algprithm becomes slower when the theoretical convergence speed is faster.**

***Index Terms*— adaptive notch filter; convergence characteristic; RPE algorithm**

I. INTRODUCTION

Adaptive IIR notch filters are designed for signal environments consisting of sinusoidal components of unknown frequency immerged in background noise. In various telecommunications systems, there is a great need for adaptive systems that are capable of automatically enhancning or suppressing sinusoidal signals immersed in a broadband signal[1].

Many techniques have been proposed for design of the adaptive notch filters and one of the well-known adaptive algorithms is the recursive prediction error(RPE) algorithm. Their methods are approximated in designing their update equations. However, the effects of approximation on convergence characteristic remain to be revealed. Thus, it is important to get an adaptive algorithm without approximation.

In this paper, we analyze convergence characteristic of adaptive notch filters. First, we derive the cost function whose minimum point forces the notch filter to identify one of the input signal frequency. Next, we derive the update equation of the adaptive notch filter without approximation. Finally, we compare the proposed algorithm with the RPE algorithm.

II. PRELIMINARIES

A. Adaptive Notch Filters

Adaptive notch filters have been used in removal of sinusoidal noises. Fig. 1 shows the block diagram of an adaptive notch filter $H(z,\beta)$, where β is the variable parameter for tuning the notch frequency $\omega_{\rm o}$ of this filter. We consider the transfer function of a second-order IIR notch filter with the notch frequency $\omega_{\rm o}$, the cutoff frequnecies ($\omega_{\rm c_1}$, $\omega_{\rm c_2}$), and

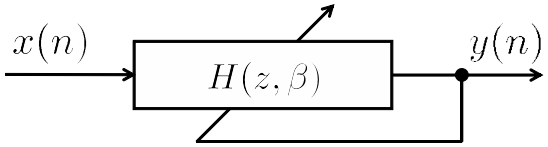


Fig. 1. Adaptive notch filter.

notch bandwidth $B_\omega = \omega_{\rm c_2} - \omega_{\rm c_1}$, given by

$$H(z,\beta) = \frac{1+\alpha}{2} \cdot \frac{1 - 2\beta z^{-1} + z^{-2}}{1 - \beta(1+\alpha)z^{-1} + \alpha z^{-2}} \tag{1}$$

where the notch bandwidth B_ω and the notch frequency $\omega_{\rm o}$ are expressed by parameters α and β such as

$$B_\omega = \omega_{\rm c_2} - \omega_{\rm c_1} = \cos^{-1}\left(\frac{2\alpha}{1+\alpha^2}\right) \tag{2}$$

$$\omega_{\rm o} = \cos^{-1}(\beta) \tag{3}$$

respectively[2]. Here, for stability, $|\alpha| < 1$ and $|\beta| < 1$. The frequency magnitude response of the above notch filter is plotted in Fig. 2 for various values of α and β.

B. Recursive Prediction Error Algorithm

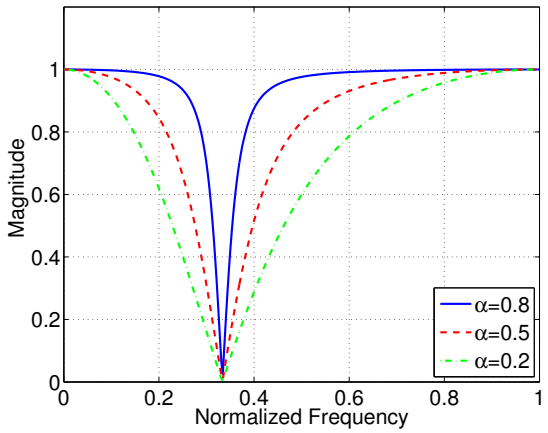
The recursive prediction error (RPE) algorithm adjusts the notch freqency of the notch filter to minimize the cost function $J(n) = E[y^2(n)]$, where $y(n)$ is the output signal[3]. The cost function $J(n) = E[y^2(n)]$ is approximated by the instantaneous value $J(n) = y^2(n)$, so that it can be applied to the case of nonstationary output signal $y(n)$.

The RPE algorithm updates the notch frequency $\omega_{\rm o}(n)$ along the negaive gradient of $J(n)$, defined as

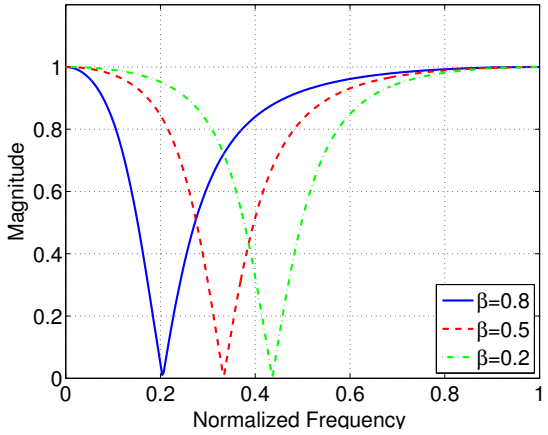
$$\nabla J(n) = \frac{\partial J(n)}{\partial \omega_{\rm o}(n)} = 2y(n)\frac{\partial y(n)}{\partial \omega_{\rm o}(n)} = 2y(n)\nabla y(n) \tag{4}$$

where the gradient $\nabla y(n)$ is given by

$$\begin{aligned}\nabla y(n) &= \frac{\partial y(n)}{\partial \omega_{\rm o}(n)} \\ &= -(1+\alpha)\sin\{\omega_{\rm o}(n)\}y(n-1) \\ &\quad +(1+\alpha)\cos\{\omega_{\rm o}(n)\}\frac{\partial y(n-1)}{\partial \omega_{\rm o}(n)} \\ &\quad +(1+\alpha)\sin\{\omega_{\rm o}(n)\}x(n-1) \\ &\quad -\alpha\frac{\partial y(n-2)}{\partial \omega_{\rm o}(n)}.\end{aligned} \tag{5}$$



(a) Three specific values of α with $\beta = 0.5$.



(b) Three specific values of β with $\alpha = 0.5$.

Fig. 2. Frequency magnitude response of the second-order IIR notch filter of Eq. (1).

Since $\omega_o(n-2)$, $\omega_o(n-1)$ and $\omega_o(n)$ stay close to each other for small step-sizes, we assume

$$\frac{\partial y(n-1)}{\partial \omega_o(n)} \approx \frac{\partial y(n-1)}{\partial \omega_o(n-1)}, \quad \frac{\partial y(n-2)}{\partial \omega_o(n)} \approx \frac{\partial y(n-2)}{\partial \omega_o(n-2)}. \tag{6}$$

Substituting Eq. (6) in Eq. (5) yields

$$\begin{aligned} \nabla y(n) &= -(1+\alpha)\sin\{\omega_o(n)\}y(n-1) \\ &\quad +(1+\alpha)\cos\{\omega_o(n)\}\frac{\partial y(n-1)}{\partial \omega_o(n-1)} \\ &\quad +(1+\alpha)\sin\{\omega_o(n)\}x(n-1) - \alpha\frac{\partial y(n-2)}{\partial \omega_o(n-2)} \\ &= -(1+\alpha)\sin\{\omega_o(n)\}\{x(n-1)-y(n-1)\} \\ &\quad +(1+\alpha)\cos\{\omega_o(n)\}\nabla y(n-1) \\ &\quad -\alpha\nabla y(n-2). \end{aligned} \tag{7}$$

Finally, the update equation for $\omega_o(n)$ is given by

$$\begin{aligned} \omega_o(n+1) &= \omega_o(n) - \mu\nabla J(n) \\ &= \omega_o(n) - 2\mu y(n)\nabla y(n) \end{aligned} \tag{8}$$

where μ is the step size parameter.

III. MAIN RESULTS

In this section, we devise the adaptive algorithm without approximation. First, we derive the cost function which is the output variance of an adaptive notch filter to the input sinusoidal signal. Next, we derive the update equation of an adaptive notch flter.

A. Cost Function

We review in this section the input signal model, the ideal notch filter, and a cost function whose minimum point forces the notch filter to identify one of the input signal frequencies.

We suppose that the input signal $x(n)$ is of the form

$$x(n) = A\sin(\omega_s n + \phi) \tag{9}$$

where A is the amplitude of the sinusoid, ω_s is the signal frequency and ϕ is its initial phase. We assume that the phase term ϕ is a randam variable which is uniformly distributed in $[0, 2\pi]$.

We find that its correlation function $r(l)$ is given by

$$\begin{aligned} r(l) &= E[x(n+l)x(n)] \\ &= A^2E[\{\sin(\omega_s(n+l)+\phi)\}\{\sin(\omega_s n+\phi)\}] \\ &= -\frac{A^2}{2}E[\cos(2\omega_s n+\omega_s l+2\phi)-\cos(\omega_s l)] \\ &= -\frac{A^2}{2}\left[\int_{-\pi}^{\pi}\cos(2\omega_s n+\omega_s l+2\phi)\cdot\frac{1}{2\pi}d\phi - \cos(\omega_s l)\right] \\ &= \frac{A^2}{2}\cos(\omega_s l). \end{aligned} \tag{10}$$

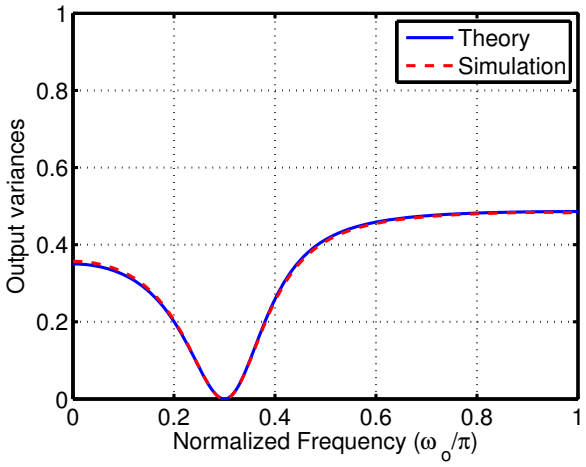
The spectral density function $S(\omega)$ is taken as

$$\begin{aligned} S(\omega) &= \sum_{l=-\infty}^{\infty} r(l)e^{-j\omega l} \\ &= \sum_{l=-\infty}^{\infty}\frac{A^2}{2}\cos(\omega_s l)e^{-j\omega l} \\ &= \frac{A^2}{4}\left[\sum_{l=-\infty}^{\infty} e^{j\omega_s l}\cdot e^{-j\omega l} + \sum_{l=-\infty}^{\infty} e^{-j\omega_s l}\cdot e^{-j\omega l}\right] \\ &= \frac{\pi A^2}{2}\delta(\omega-\omega_s) + \frac{\pi A^2}{2}\delta(\omega+\omega_s) \end{aligned} \tag{11}$$

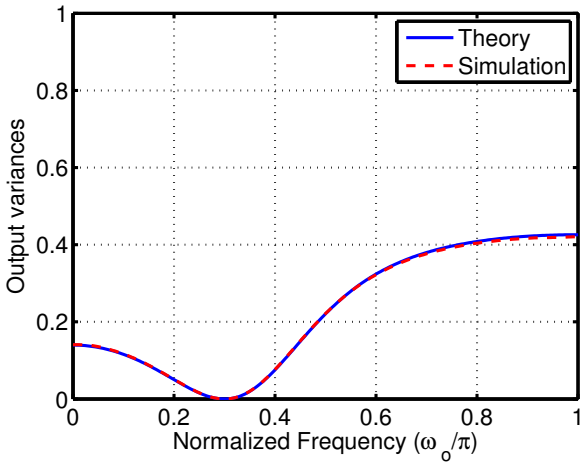
where $\delta(\cdot)$ is the Dirac delta function.

The output variance $E[y^2(n)]$ is taken as

$$E[y^2(n)] = r_y(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi}|H(e^{j\omega})|^2 S(\omega)d\omega. \tag{12}$$



(a) $\alpha = 0.5\pi$.



(b) $\alpha = 0.1\pi$.

Fig. 3. The output variances of the adaptive notch filters.

Substituting Eq. (11) in Eq. (12) yields the cost function $J(n) = E[y^2(n)]$, given by

$$\begin{aligned} E[y^2(n)] &= \frac{1}{2\pi}\int_{-\pi}^{\pi} |H(e^{j\omega})|^2 S(\omega) d\omega \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|\frac{1+\alpha}{2}\cdot\frac{1-2\beta e^{j\omega}+e^{j\omega\cdot 2}}{1-\beta(1+\alpha)e^{j\omega}+\alpha e^{j\omega\cdot 2}}\right|^2 \\ &\quad \times\left(\frac{\pi A^2}{2}\right)\{\delta(\omega-\omega_s)+\delta(\omega+\omega_s)\}\, d\omega \\ &= \frac{A^2}{4}\left|\frac{1+\alpha}{2}\cdot\frac{1-2\beta e^{j\omega_s}+e^{j\omega_s\cdot 2}}{1-\beta(1+\alpha)e^{j\omega_s}+\alpha e^{j\omega_s\cdot 2}}\right|^2 \\ &\quad +\frac{A^2}{4}\left|\frac{1+\alpha}{2}\cdot\frac{1-2\beta e^{j\omega_s}+e^{j\omega_s\cdot 2}}{1-\beta(1+\alpha)e^{j\omega_s}+\alpha e^{j\omega_s\cdot 2}}\right|^2 \\ &= \frac{1-4\beta\cos(\omega_s)+2\beta^2+\cos(2\omega_s)}{(1+\alpha)^2(1-2\beta\cos(\omega_s)+\beta^2)+2\alpha\cos(2\omega_s)} \\ &\quad \times\frac{A^2(1+\alpha)^2}{4}. \end{aligned} \tag{13}$$

B. Without Approximation Update Equation

We derived in the previous section the cost function $J(n) = E[y^2(n)]$. If $\omega_o(n)$ is the notch frequency at n, then its updated value $\omega_o(n+1)$ is given by

$$\begin{aligned} \omega_o(n+1) &= \omega_o(n) - \mu\nabla J(n) \\ &= \omega_o(n) - \mu\frac{\partial E[y^2(n)]}{\partial\omega_o(n)}. \end{aligned} \tag{14}$$

The gradient vector is equal to

$$\begin{aligned} \frac{\partial E[y^2(n)]}{\partial\omega_o(n)} &= 4(1-\alpha)^2[\cos(\omega_s)-\cos\{\omega_o(n)\}] \\ &\quad \times\sin^2(\omega_s)\sin(\omega_o(n)). \end{aligned} \tag{15}$$

Hence, substituting Eq. (15) in Eq. (14) yields the update equation for $\omega_o(n)$, given by

$$\begin{aligned} \omega_o(n+1) &= \omega_o(n) - \mu\frac{\partial E[y^2(n)]}{\partial\omega_o(n)} \\ &= \omega_o(n) - 4\mu(1-\alpha)^2 \\ &\quad \times[\cos(\omega_s)-\cos\{\omega_o(n)\}] \\ &\quad \times\sin^2(\omega_s)\sin\{\omega_o(n)\} \end{aligned} \tag{16}$$

IV. SIMULATION RESULTS

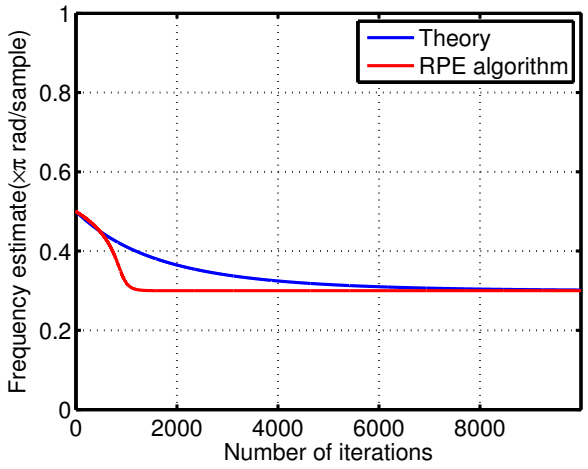
This section shows simulation results for the adaptive notch filters. The input signal used here is given by

$$x(n) = \sin(0.3\pi n + \phi) \tag{17}$$

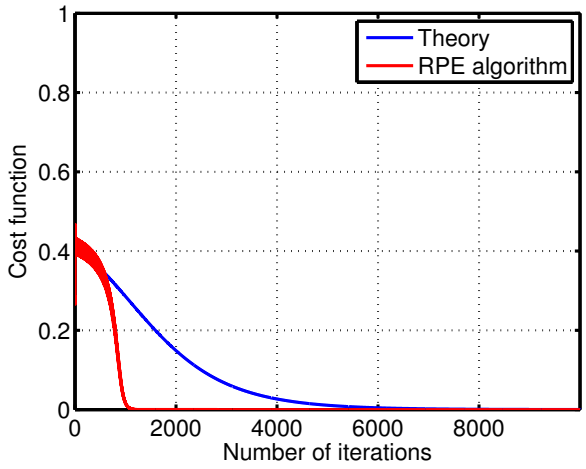
where ϕ is an initial phase which is a randam variable, uniformly distributed in $[0, 2\pi]$. Fig. 3 shows comparison between theoretical and simulated output variances versus the notch frequency ω_o with different choices of notch bandwidth parameter α. We see that the simulation results confirm the theoretical results in Fig. 3.

Fig. 4 shows comparison between theoretical and RPE algorithm frequency estimate and output variances. Each result is obtained with the specifacation of $\alpha = 0,5$, $\mu = 0.001$ and $\omega_o(0) = 0.5\pi$. In this case, the output variances versus the notch frequrncy ω_o are shown in Fig. 3(a). Fig. 4 shows that the convergence speed of RPE algorithm becomes faster when the theoretical convergence speed is relatively slower.

Fig. 5 shows comparison between theoretical and RPE algorithm frequency estimate and output variances. Each result is obtained with the specifacation of $\alpha = 0,1$, $\mu = 0.001$ and $\omega_o(0) = 0.5\pi$. In this case, the output variances versus the notch frequrncy ω_o are shown in Fig. 3(b). Fig. 5 shows that the convergence speed of RPE algorithm becomes slower when the theoretical convergence speed is relatively faster.

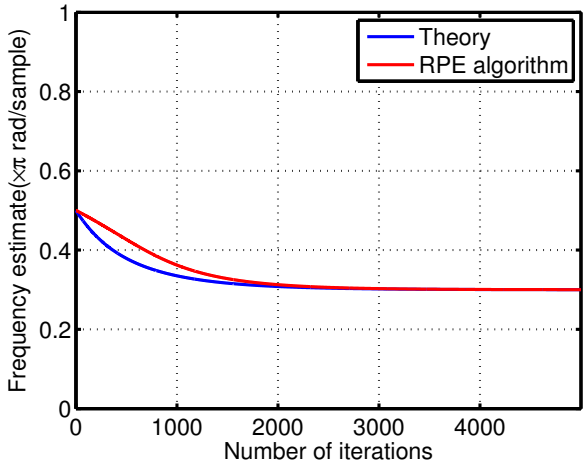


(a) Frequency estimates for adaptive notch filters.

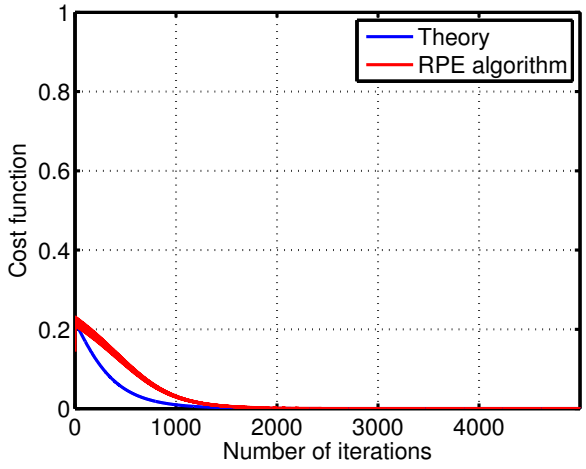


(b) The cost functions of adaptive notch filters.

Fig. 4. Comparison between theoretical and RPE algprithm($\alpha = 0.5\pi$).



(a) Frequency estimates for adaptive notch filters.



(b) The cost functions of adaptive notch filters.

Fig. 5. Comparison between theoretical and RPE algprithm($\alpha = 0.1\pi$).

V. CONCLUSION

This paper has analyzed convergence characteristic of adaptive notch filters. First, we derived the cost function which is the output variance of an adaptive notch filter to the input signal sinusoidal. Next, we derived the adaptive algorithm of adaptive notch filters without approximation. We compared the proposed adaptive algorithm with RPE algorithm. One of our future works is an improvement of the convergence characteristic of an adaptive notch filter.

REFERENCES

[1] P. A. Regalia, *Adaptive IIR Filtering in Signal Processing and Control*. Marcel Dekker, 1995.

[2] S. K. Mitra, *Digital Signal Processing - A Computer-Based Approach, 3rd Edition*. McGraw-Hill, 2006.

[3] J. J. Shynk, *Adaptive IIR fltering*. IEEE ASSP Magazine, vol. 6, no. 2, pp.4-21, April 1989.