Issue 6

April 2014

APSIPA NEWSLETTER

Asia-Pacific Signal and Information Processing Association Newsletter

APSIPA Social Net and Friend Labs

Kenneth Lam, Thomas Zhang and C.-C. Jay Kuo

One of the key missions of APSIPA is to provide education, research and development exchange platforms for both academia and industry. This mission is being achieved by presentations in APSIPA Annual Summit and Conferences, APSIPA Newsletters, publications in APSIPA Transactions on Signal and Information Processing (an open access journal), etc. Due to the recent emergence of social media such as Facebook and LinkedIn, we have another convenient platform for information sharing. This idea was first discussed in APSIPA ASC 2011 in Xi'an, China, with an overwhelming support from the APSIPA Board of Governors.

Immediately after APSIPA ASC 2011, the APSIPA Social Net program was launched in 2011 December. It is built upon the world most popular professional social network - LinkedIn. A group called "Asia Pacific Signal and Information Processing Association – APSIPA" has been set up in LinkedIn. Any people with a LinkedIn account can apply to join. There are currently around 2500 APSIPA e-members

Furthermore, we establish two types of sub-groups under the APSIPA main group.

- The first type of subgroups is created based on the 6 TC tracks. They are: APSIPA.SPS, APSIPA.SIPTM, APSIPA.SLA, APSIPA.BioSiPS, APSIPA.IVM and APSIPA.WCN. The 6 TC members are encouraged to use their subgroup to communicate
- The second type of subgroups is created based on countries/regions in the Asia-Pacific rim. Now, we have: APSIPA-USA (United States), APSIPA-CHN (China), APSIPA-JPN (Japan), APSIPA-KRO (South Korea), APSIPA-AUS (Australia), APSIPA-CAN (Canada), APSIPA-HKG (Hong Kong), APSIPA-IND (India), APSIPA-NZL (New Zealand), APSIPA-SGP (Singapore), APSIPA-TWN (Taiwan) and APSIPA-THA (Thailand). The country-based subgroup will serve as the basis to build up local chapters.

To enrich contents in the APSIPA Social Net, we have filmed a sequence of interviews. They are accessible via http://www.apsipa.org/social.htm. In 2014 April, this task is assigned to a newly formed APSIPA Social Net Committee (ASNC). The ASNC is led by APSIPA VP-Member Relations and Development, Professor Kenneth Lam, with the following members:

- Ms Summer Jia He, University of Southern California, USA (Secretary)
- Dr Iman Ardekani, Unitec Institute of Technology, New Zealand
- Dr Cheng Cai, Northwest A&F University, China
- Dr Lu Fang, University of Science and Technology of China

In this issue

- Dr Weiyao Lin, Shanghai Jiao Tong University, China
- Dr Chuang Shi, Kansai University, Japan
- Prof. Chia-Hung Yeh, National Sun Yat-sen University, Taiwan
- Dr Jiantao Zhou, University of Macau, Macau

Another related program called “APSIPA Friend Labs” was launched in 2013 August through the joint effort of APSIPA VP-Institutional Relations and Education, Professor Thomas Zhang, and APSIPA VP-Member Relations and Development, Professor Kenneth Lam. An academia or industrial lab is qualified to be an APSIPA friend lab if it has at least 10 current or former lab members who are full or associate members of APSIPA. A person can be an APSIPA associate member by joining the APSIPA Group in Linkedin. A person can be an APSIPA full member by clicking the "Join Us" button in the up-right corner of the APSIPA homepage and following the given instructions. All APSIPA Friend Labs will be listed in the APSIPA website. Each lab has one page to post lab information, photos and a link to the lab home page. The provided data in the on-line application form will be used to generate the friend lab page in the APSIPA website. There are about 150 APSIPA Friend Labs now and the number continues to grow. The list of current APSIPA Friend Labs is given in the following page: http://www.apsipa.org/friendlab/Application/lablist.asp.

The APSIPA Social Net and Friend Labs are new initiatives for a professional community. They are still in their early stage, demanding further dedication and innovation to reach the full potential. We sincerely invite you, your friends and colleagues to join the APSIPA group to become its e-member. We would also like to recruit more research labs to become APSIPA Friend Labs. Your participation will make APSIPA a warm and interactive professional community.

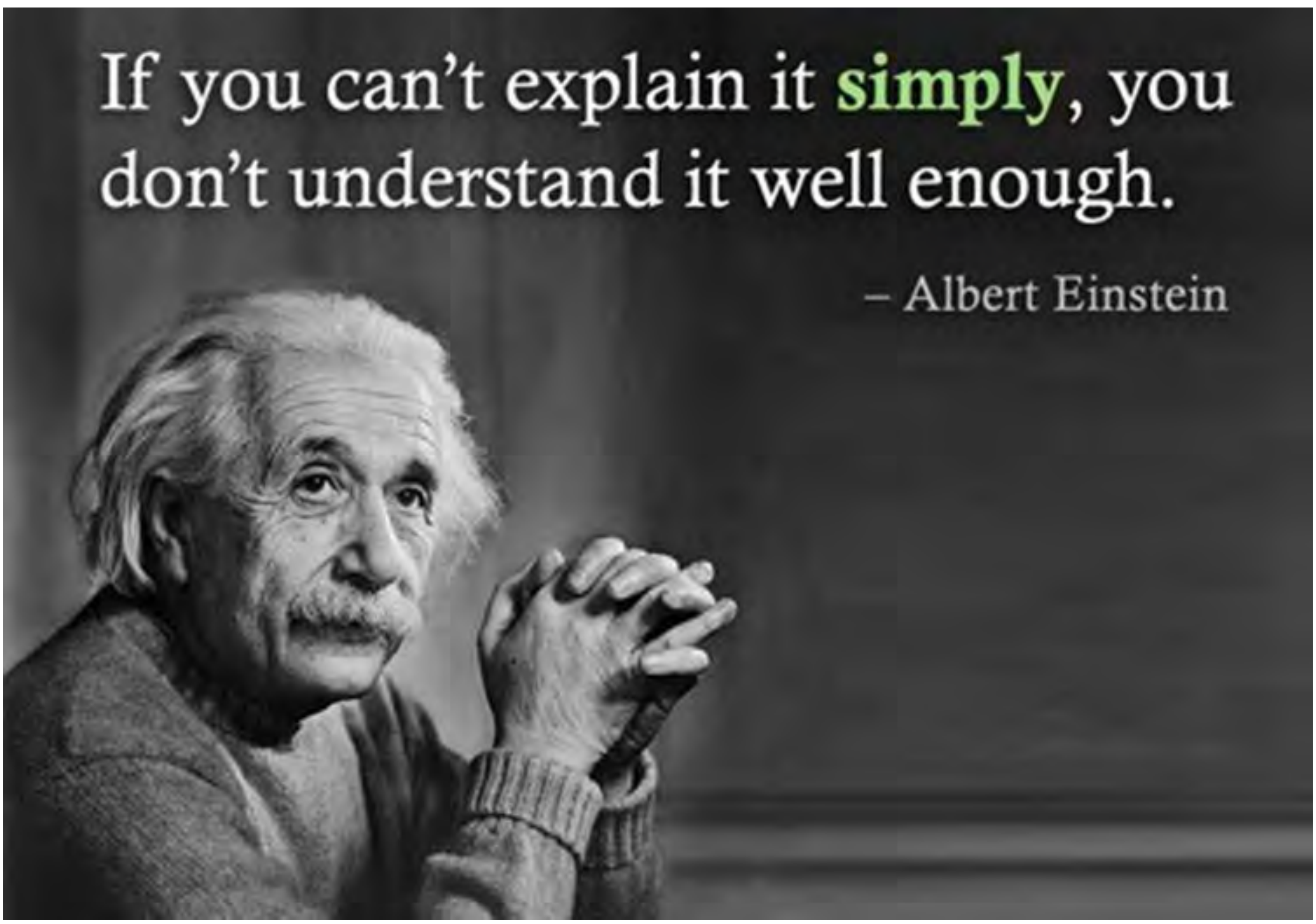
Recent APSIPA Distinguished Lecturer Activity

1) Lecturer: Tatsuya Kawahara (Kyoto University, Japan)
Title: Recent trend of spoken dialogue systems
Time: 3:00PM, March 26, 2014
Venue: Room 1-303 Info Sci&Tech Building, Tsinghua University, Beijing, China
Host: Prof. Thomas Fang Zheng

There has been significant progress in spoken dialogue systems both in research and deployment. This talk will give a brief history of spoken dialogue systems including their evolution and recent technical trend.
In the new perspective, back-end of spoken dialogue systems is extended from conventional relational databases (RDB) to general documents, incorporating information retrieval (IR) and question-answering (QA) techniques. This paradigm shift and the author's approaches are reviewed.

2) Lecturer: Tatsuya Kawahara (Kyoto University, Japan)
Title: Multi-modal sensing and analysis of poster conversations
Time: 3:00PM, March 28, 2014 (Friday)
Venue: Room 25-B407, Tianjin University, China
Host: Prof. Jianwu Dang

Conversations in poster sessions in academic events, referred to as poster conversations, pose interesting and challenging topics on multi-modal signal and information processing. This talk gives an overview of our project on the smart posterboard for multi-modal conversation analysis. The smart posterboard has multiple sensing devices to record poster conversations, so we can review who came to the poster and what kind of questions or comments he/she made. The conversation analysis combines speech and image processing such as head tracking, speech enhancement and speaker diarization. Moreover, we are also investigating high-level indexing of interest and comprehension level of the audience, based on their multi-modal behaviors during the conversation.



Text-dependent Speaker Verification and RSR2015 Speech Corpus

Anthony Larcher and Haizhou Li

RSR2015 (Robust Speaker Recognition 2015) is the largest publicly available speech corpus for text-dependent robust speaker recognition. The current release includes 151 hours of short duration utterances spoken by 300 speakers. RSR2015 is developed by the Human Language Technology (HLT) department at Institute for Infocomm Research (I2R) in Singapore. This newsletter describes RSR2015 corpus that addresses the reviving interest of text-dependent speaker recognition.

WHY ANOTHER CORPUS?

Speaker verification has reached a maturity that allows state-of-the-art engines to be used for commercial applications. During the last decade, the effort of the community has lead to great improvements in terms of performance and usability of speaker verification engines. Nevertheless, it is well known in the community that performance of text-independent speaker verification engines suffers from the lack of enrolment and training data.

In the context of short duration, text-dependency is known to improve accuracy of speaker verification when dealing with short duration speech segments. Despite its strong commercial potential, text-dependent speaker recognition lies on fringes of the main stream speaker recognition. As a result, the speech resources available for such research are either too small or inadequate to take advantage of the technologies develop for the mainstream text-independent speaker verification. In view of the fact that text-dependent speaker verification and user-customized command and control, that recognizes a user-defined voice command at the same time identifies the speaker, are useful application scenario. RSR2015 is developed for the following objectives.

- To provide a database of reasonable size that supports significance test of text-independent speaker recognition. RSR2015 contains recordings from 300 speakers during 9 sessions in order to create enough target and impostor trials.
- To have a gender-balanced database that allows fair analysis of gender influence in text-dependent speaker recognition. RSR2015 involves 143 female and 157 male speakers.
- To allow for analysis of the phonetic variability in the context of text-constrained speaker recognition [2]. RSR2015 protocol includes more than 60 different utterances spoken by all speakers.

WHAT CAN WE DO WITH RSR2015?

RSR2015 allows for simulation and comparison of different use-cases in terms of phonetic content. For example, the most extreme constraint is to fix a unique utterance for all users of the system all the time. In the case where a larger set of fixed pass-phrases is shared across users, the scenario becomes very similar to user-customized command and control application. On the other hand, it is possible to limit the phonetic content of the speech utterance by randomly prompting sequences of phones or digits. In this case, the context of the phone varies across sessions and especially between enrolment and test.

The choice of a specific scenario depends on what constraints we would like to impose on the users. Unfortunately, no existing database allows for a comparison of speaker recognition engines across scenarios in similar conditions. RSR2015 is designed to bridge the gap. It consists of fixed pass-phrases, short commands and random digit series recorded in the same conditions.

DATABASE DESCRIPTION

RSR2015 contains audio recordings from 300 speakers, 143 female and 157 male in 9 sessions each, with a total of 151 hours of speech. The speakers were selected to be representative of the ethnic distribution of Singaporean population, with age ranging from 17 to 42.

The database was collected in office environment using six portable devices (four smart phones and two tablets) from different manufacturers. Each speaker was recorded using three different devices out of the six. The speaker was free to hold the smart phone or tablet in a comfortable way.

To facilitate the recording, a dialogue manager was implemented on the portable devices as an Android© application. The speakers interact with the dialogue manager through a touch screen to complete the recording. Each of the 9 sessions for a speaker is organized into 3 parts:

PART 1 – Short-sentences for pass-phrase style speaker verification (71 hours)
All speakers read the same 30 sentences from the TIMIT database [3] covering all English phones. The average duration of sentences is 3.2 seconds.
Example: "Only lawyers love millionaires."

PART 2 – Short commands for user-customized command and control (45 hours)
All speakers read the same 30 commands designed for the StarHome applications. The average duration of short commands is 2 seconds.
Example "Light on"

PART 3 – Random digit strings for speaker verification (35 hours)
All speakers read the same 3 10-digit strings, and 10 5-digit strings. The digit strings are session dependent.

INTERSPEECH 2014 SPECIAL SESSION AND BENCHMARKING

During INTERSPEECH 2014, The Institute for Infocomm Research together with IBM Research, and the Centre de Recherche en Informatique de Montreal (CRIM) propose a special session on *Text-Dependent Speaker Verification with Short Utterance*. The purpose of this special session is to gather the research efforts from both the academia and industries toward a common goal of establishing a new baseline and explore new directions for text-dependent speaker verification. For ease of comparison across systems, the RSR2015 database, that comes with several evaluation protocols targeting at different scenarios, has been proposed to support research submitted for the INTERSPEECH 2014 special session [4, 5].

WHERE TO GET THIS DATABASE?

The license is available at [Link]

Please contact Dr Anthony Larcher at Email alarcher [at] i2r.a-star.edu.sg or Dr Kong-Aik Lee at kalee [at] i2r.a-star.edu.sg.

REFERENCES:

[1] K. A. Lee, A. Larcher, H. Thai, B. Ma and H. Li, "Joint Application of Speech and Speaker Recognition for Automation and Security in Smart Home", in *Annual Conference of the International Speech Communication Association (Interspeech),* 2011, pp 3317-3318.

[2] A. Larcher, P.-M. Bousquet, K. A. Lee, D. Matrouf, H. Li and J.-F. Bonastre, "I-vectors in the context of phonetically-constrained short utterances for speaker verification", In *ICASSP* 2012

[3] W. M. Fisher, G. R. Doddington and K. M. Goudie-Marshall, "The DARPA speech recognition research database: specifications and status", *DARPA Workshop on Speech Recognition,* 1986, pp. 93-99

[4] A. Larcher, K. A. Lee, B. Ma and H. Li, "RSR2015: Database for Text-Dependent Speaker Verification using Multiple Pass-Phrases", in *Annual Conference of the International Speech Communication Association (Interspeech),* **2012**, 1580-1583

[5] A. Larcher, K. A. Lee, B. Ma and H. Li, "Text-dependent Speaker Verification: Classifiers, Databases and RSR2015," *Speech Communication,* **2014,** in press

APSIPA Membership Board

The APSIPA Membership Board (AMB) was set up in February 2014. The AMB is led by APSIPA VP-Member Relations and Development, Professor Kenneth Lam, with the following members:

- Prof. Waleed Abdulla, The University of Auckland
- Prof. Woon-Seng Gan, Nanyang Technological University
- Prof. Hsueh-Ming Hang, National Chiao-Tung University
- Prof. Yo-Sung Ho, Gwangju Institute of Science and Technology
- Prof. Jiwu Huang, Shenzhen University
- Prof. Hitoshi Kiya, Tokyo Metropolitan University
- Prof. Chung-Nan Lee, National Sun Yat-sen University
- Dr Tan Lee, Chinese University of Hong Kong
- Prof. Nam Ling, Santa Clara University
- Prof. Ming-Ting Sun, University of Washington

The AMB members will serve a renewable 2-year term, with a maximum of two consecutive terms. The AMB will actively promote APSIPA to all researchers and academics in the fields of signal and information processing, and also organize APSIPA activities for the benefit of APSIPA members and e-members.

RECENT ADVANCES IN CONTRAST ENHANCEMENT AND THEIR APPLICATIONS TO LOW-POWER IMAGE PROCESSING

Chulwoo Lee, Chul Lee, and Chang-Su Kim

School of Electrical Engineering, Korea University, Seoul, Korea
E-mails: {chulwoo, chul.lee, cskim}@mcl.korea.ac.kr

1. INTRODUCTION

In spite of recent advances in imaging technology, captured images often fail to preserve scene details faithfully or yield poor contrast ratios due to limited dynamic ranges. Contrast enhancement (CE) techniques can alleviate these problems by increasing contrast ratios. Therefore, CE is an essential step in various image processing applications, such as digital photography and visual surveillance.

Conventional CE techniques can be categorized into global and local approaches. A global approach derives a single transformation function, which maps input pixel intensities to output pixel intensities, and applies it to all pixels in an image. Gamma correction, based on the simple power law, and histogram equalization (HE), which attempts to make the histogram of pixel intensities as uniform as possible, are two popular global contrast enhancement techniques [1]. A local approach, on the other hand, derives and applies the transformation function for each pixel adaptively. However, in general, a local approach demands higher computational complexity and its level of CE is harder to control. Therefore, global CE techniques are more widely used for general purposes than local ones. In this article, we review recent advances in global CE techniques and its applications.

2. CE TECHNIQUES AND ITS APPLICATIONS

HE is one of the most popular techniques to enhance low contrast images due to its simplicity and effectiveness. However, it has some drawbacks, such as contrast over-stretching, noise amplification, or contour artifacts. Various researches have been carried out to overcome these drawbacks. For example, several algorithms [2, 3, 4] divide an input histogram into sub-histograms and equalize them independently to reduce the brightness change between input and output images. Recently, histogram modification (HM) techniques, which manipulate an acquired histogram before the equalization, have been introduced. Let us review the generalization of HM algorithms [5], and introduce its applications to low power image processing.

2.1. Histogram Modification

For notational simplicity, let us consider a typical 8-bit imaging system, in which the maximum gray-level is 255. Let $\mathbf{x} = [x_0, x_1, \cdots, x_{255}]^T$ denote the transformation function, which maps gray-level k in the input image to gray-level x_k in the output image [5]. Then, conventional histogram equalization can be achieved by solving

$$\mathbf{D}\mathbf{x} = \overline{\mathbf{h}} \tag{1}$$

where $\overline{\mathbf{h}} = 255\mathbf{h}/(\mathbf{1}^T\mathbf{h})$ is the normalized input histogram and $\mathbf{D}$ is the differential matrix [4].

HE enhances the dynamic range of an image and yields good perceptual image quality. However, if many pixels are concentrated within a small range of gray levels, the output transformation function may have an extreme slope. This degrades the output image severely. To overcome this drawback, a recent approach to HM modifies the input histogram before the HE procedure to reduce extreme slopes in the transformation function. For instance, Wang and Ward [6] clamped large histogram values and then modified the resulting histogram using the power law. Also, Arici *et al.* [7] reduced the histogram values for large smooth areas, which often correspond to background regions, and mixed the resulting histogram with the uniform histogram. Lee *et al.* [5] employed a logarithm function to reduce large histogram values and prevent the transformation function from having too steep slopes.

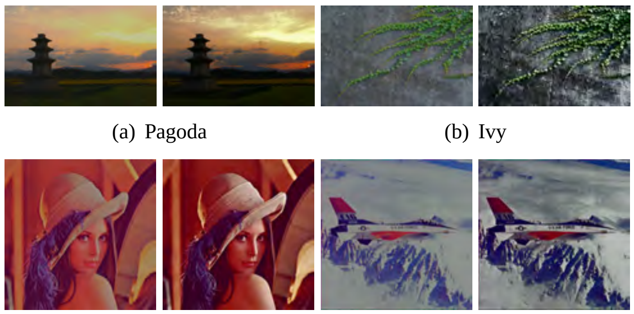
In this recent HM approach, the first step can be expressed by a vector-converting operation $\mathbf{m} = f(\mathbf{h})$, where $\mathbf{m}$ denotes the modified histogram. Then, the desired transformation function $\mathbf{x}$ can be obtained by minimizing the cost function

$$C_{\mathrm{H}} = \|\mathbf{D}\mathbf{x} - \overline{\mathbf{m}}\|^2 \tag{2}$$

where $\overline{\mathbf{m}}$ is the normalized column vector of $\mathbf{m}$.

2.2. Application to Low Power Image Processing

Whereas a variety of techniques have been proposed for the CE of general image, relatively little effort has been made to adapt the enhancement process to the characteristics of display devices. Since a large portion of power is consumed by display panels, it is essential to develop an image processing



(a) Pagoda (b) Ivy

(c) Lena (d) F-16

Fig. 1. Comparison of power-reduced output images. In each subfigure, the left image is obtained by the linear mapping method, and the right one by the PCCE algorithm in [5].

algorithm, which can save power in display panels as well as enhancing image contrast. Let us introduce display-specific CE algorithms by using the HM framework.

CE for emissive displays The power consumption of emissive displays, such as OLED devices, is proportional to squared pixel intensities [5]. By adding the power consumption term to the initial objective function in (2), we obtain a new cost function

$$C_{\mathrm{E}} = \|\mathbf{D}\mathbf{x} - \overline{\mathbf{m}}\|^2 + \lambda \mathbf{x}^t \mathbf{H}\mathbf{x} \tag{3}$$

where λ is a user parameter balancing between power and contrast. By minimizing the cost function, we can reduce the power while enhancing the contrast. If $\lambda = 0$, the power saving is not considered. On the other hand, as λ increase, the output image gets darker to achieve power reduction.

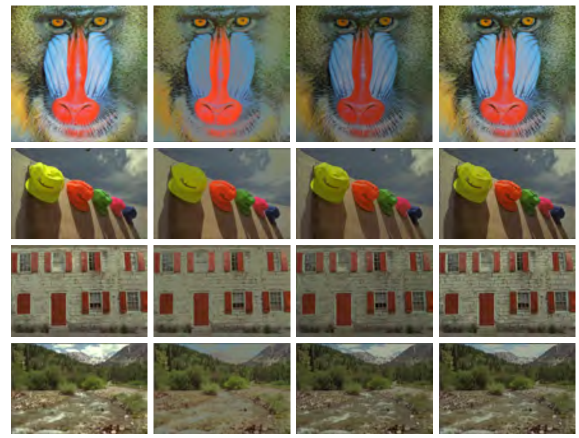
Fig. 1 compares the linear reducing method with the proposed power-constrained contrast enhancement (PCCE) algorithm [5]. The linear method provides hazy output images because of the low contrast. On the other hand, the proposed PCCE algorithm yields more satisfactory image qualities.

CE for non-emissive displays Conventional lower power image processing techniques for non-emissive displays, such as TFT-LCD devides, compensate a reduced backlight by increasing pixel intensities. Let us denote the backlight scaling factor as $b \in [0, 1]$. Then, output pixel intensities are scaled by factor of $1/b$. Since the maximum values are limited, the display shows $\min(255, \mathbf{x}/b)$. Therefore, the amount of quality loss at each grey level is given by

$$\mathbf{x}_b = \min(0, \min(255, \mathbf{x}/b) - 255) \tag{4}$$

The quality loss is modeled by $\mathbf{x}_b^t \mathbf{H}\mathbf{x}_b$ [8], where $\mathbf{H}$ is the diagonal matrix, obtained from the input histogram $\mathbf{h}$. By integrating this term into initial objective function (2), we obtain

$$C_{\mathrm{N}} = \|\mathbf{D}\mathbf{x} - \overline{\mathbf{m}}\|^2 + \mu \mathbf{x}_b^t \mathbf{H}\mathbf{x}_b \tag{5}$$



(a) (b) (c) (d)

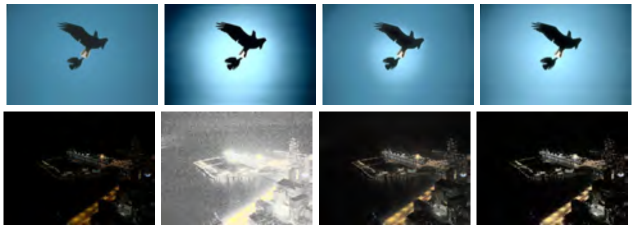
Fig. 2. Brightness-compensated contrast enhancement results on the “Baboon,” “Hats,” “Building,” and “Stream” images at $b = 0.5$. The input images in (a) are compensated by the linear compensation method in (b), the Tsai *et al.*’s algorithm [9] in (d), and the BCCE algorithm [8] in (d).

where μ is a user-controllable parameter.

Fig. 2 shows the result when the proposed BCCE algorithm [8] for non-emissive displays is applied. In case of the linear mapping, details in bright regions are lost. Although the conventional algorithm in [9] preserves the details more faithfully, the dynamic range is reduced and noise components are amplified. The proposed BCCE algorithm shows better image qualities.

3. CE TECHNIQUES USING HIGHER ORDER STATISTICS

Recently, several algorithms have been proposed to consider the joint distribution of neighboring pixel values for CE. These algorithms exploit the 2D histogram of neighboring pixel values, instead of the 1D histogram of individual pixel values. Celik and Tjahjadi [10] obtained a target 2D histogram by minimizing the sum of the differences from an input 2D histogram and the uniform histogram, and mapped the diagonal elements of the input histogram to those of the target histogram. Lee *et al.* [11] attempted to emphasize the gray-level differences of neighboring pixels that occur frequently in an input image. To this end, they proposed a tree-like data structure, called layered difference representation (LDR), and derived an efficient solution to the optimization problem. Shu and Wu [12] also exploited a 2D histogram of pixel values and computed a transformation function to maximize the expected image contrast. Among these new CE algorithms, let us briefly summarize our LDR work [11].



(a) (b) (c) (d)

Fig. 3. CE results on the "Eagle" and "Night view" image: (a) input image, (b) HE, (c) CVC [10], and (d) the LDR [11].

3.1. Layered Difference Representation

In LDR, a 2D histogram is used to find a desirable transformation function $\mathbf{x}$. Suppose that a pair of adjacent pixels in the input image have gray-levels k and $k+l$, then their difference l is mapped to the difference

$$d_k^l = x_{k+l} - x_k, \quad 0 \le k \le 255 - l \tag{6}$$

in the output image. Let the 2D histogram $h(k, k+l)$ count the number of pairs of adjacent pixels with gray-levels k and $k+l$ in the input image. Also, let $h_k^l = h(k, k+l) + h(k+l, k)$. The objective is to design the transformation function $\mathbf{x}$, which yields a large output difference d_k^l when h_k^l is a large number. In other words, frequently occurring pairs of pixel values in the input image should be clearly distinguished in the output image. To this end, d_k^l should be set proportionally to h_k^l:

$$d_k^l = \kappa_l h_k^l \tag{7}$$

where κ_l is a normalizing constant in [11].

The difference variables, d_k^l's, are grouped according to the input gray-level difference l, and each group is referred to as a *layer* in the hierarchical structure of LDR. Notice that d_k^l can be decomposed into the difference variables d_k^1's at layer 1 by

$$d_k^l = \sum_{i=k}^{k+l-1} d_i^1, \quad 0 \le k \le 255 - l. \tag{8}$$

Also, the transformation function is determined by the difference variables at layer 1,

$$\begin{aligned} x_0 &= 0, \\ x_k &= \sum_{i=0}^{k-1} d_i^1, \quad 1 \le k \le 255. \end{aligned} \tag{9}$$

From the relationships in (7) and (8), a difference vector $\mathbf{d} = [d_0^1, d_1^1, \ldots, d_{254}^1]^T$ is determined at each layer. Then, those difference vectors at all layers are aggregated into a unified vector, which is finally used to construct the transformation function $\mathbf{x}$ using (9).

In Fig. 3, HE and CVC cannot handle large histogram properly. Specifically, HE and CVC increase the contrast on the smooth region excessively, producing contour artifacts on "Eagle" image. On the contrary, LDR algorithm [11] exhibits better contrast by exploiting the full dynamic range. Image qualities are degraded when scenes are captured in very low light conditions. "Night view" image indicates a dark input image, which contains noise components. HE yields an extremely noisy image. Although CVC exploits the 2D histogram information, it still experiences the over-enhancement problem due to the high histogram peak. On the other hand, LDR algorithm [11] alleviates noise and clarifies the details of the buildings.

4. CONCLUSIONS

This article introduced the HM framework, formulated as constrained optimization problem. Also, we verified that HM techniques can be applied to the power-constrained image processing algorithms, which can enhance image quality and reduce display power consumption simultaneously. Moreover, we briefly reviewed one of the CE techniques using higher order statistics, which become more popular in recent years.

5. REFERENCES

[1] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice Hall, 3rd edition, 2007.

[2] Y.-T. Kim, "Contrast enhancement using brightness preserving bi-histogram equalization," vol. 43, no. 1, pp. 1–8, Feb. 1997.

[3] Y. Wang, Q. Chen, and B. Zhang, "Image enhancement based on equal area dualistic sub-image histogram equalization method," vol. 45, no. 1, pp. 68–75, Feb. 1999.

[4] Soong-Der Chen and A.R. Ramli, "Minimum mean brightness error bi-histogram equalization in contrast enhancement," vol. 49, no. 4, pp. 1310–1319, Nov. 2003.

[5] C. Lee, C. Lee, Y.-Y. Lee, and C.-S. Kim, "Power-constrained contrast enhancement for emissive displays based on histogram equalization," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 80–93, Jan. 2012.

[6] Q. Wang and R. K. Ward, "Fast image/video contrast enhancement based on weighted thresholded histogram equalization," vol. 53, no. 2, pp. 757–764, May 2007.

[7] T. Arici, S. Dikbas, and Y. Altunbasak, "A histogram modification framework and its application for image contrast enhancement," *IEEE Trans. Image Process.*, vol. 18, no. 9, pp. 1921–1935, Sept. 2009.

[8] C. Lee, J.-H. Kim, C. Lee, and C.-S. Kim, "Optimized brightness compensation and contrast enhancement for transmissive liquid crystal displays," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, pp. 576–590, Apr. 2014.

[9] P.-S. Tsai, C.-K. Liang, T.-H. Huang, and H. H. Chen, "Image enhancement for backlight-scaled TFT-LCD displays," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 4, pp. 574–583, Apr. 2009.

[10] T. Celik and T. Tjahjadi, "Contextual and variantional contrast enhancement," *IEEE Trans. Image Process.*, vol. 20, pp. 1921–1935, Sept. 2009.

[11] C. Lee, C. Lee, and C.-S. Kim, "Contrast enhancement based on layered difference representation of 2-D histograms," *IEEE Trans. Image Process.*, vol. 22, pp. 5372–5384, Dec. 2013.

[12] X. Shu and X. Wu, "Image enhancement revisited: from first order to second order statistics," in *Proc. IEEE ICIP*, Sep. 2013, pp. 886–890.

Robust Speech Recognition and its LSI Design

Yoshikazu MIYANAGA

Hokkaido University, Sapporo, 060-0814 Japan

1 Introduction

This letter introduces recent noise robust speech communication techniques and its implementation to a small robot. This speech communication system consists of automatic speech detection, speech recognition and speech rejection.

One of current speech recognition systems are considered as a system consisting of phoneme-based speech recognition and language model [1], [2]. As another speech recognition approach, a phrase-based speech recognition has been also explored [3] - [5]. Compared with the system with phoneme-based speech recognition and language model, the phrase- based speech recognition can provide higher recognition accuracy in case of various noise environments. On the other hand, it requires high calculation cost and a lot of training database for all target speech phrases. By using efficient LSI design of a speech recognition system, the author can realize real time speech recognition with low power consumption. The embedded speech communication module recognizes all target speech phrases with high recognition accuracy where non target speeches are automatically rejected. The smart robot system can answer only for the target speech phrases. About the recognition accuracy of this system, the system can realizes 85% - 99% under 0 - 20 dB SNR and echo environments.

2 Speech Communication System

When we consider an embedded automatic speech recognition (ASR) system, e.g., ROBOT with ASR, many poor surroundings and long distance microphone conditions should be considered. In case of exhibition rooms, out-door, and general house environments, the SNR should be always lower than 20dB. In case of free-hands and general human communications, the 10 cm - 5 m distances should be also considered. The introduced ASR can recognize speech under the above various conditions accurately. The conditions used in our development are considered as (1) white noise, speech-like noise and factory noise with 5dB SNR, (2) echo environment less than 150ms and (3) medium distance ($< 5m$) from a microphone.

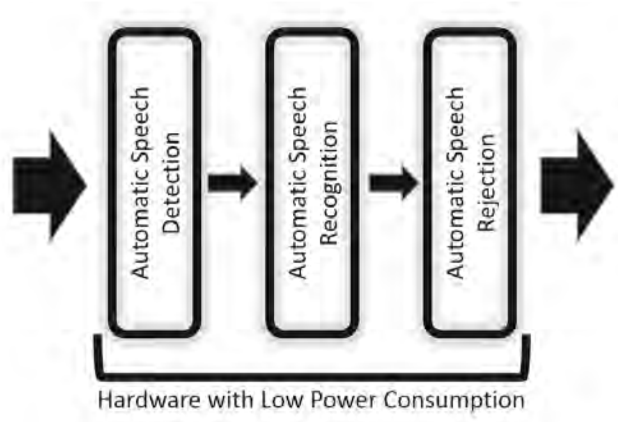


Figure 1: Our Speech Communication System

In order to keep the high performance of ASR, our system only recognizes the set of selected phrases. Any other words and phrases except target speech are rejected by our system. In Fig.1, the overview of our speech communication system is described. The system consists of automatic speech detection, i.e., automatic voice activity detection (VAD), automatic speech recognition (ASR), and automatic speech rejection.

The ASR in this system has been designed as a noise robust ASR. In order to make the system robust against various noises, the combination of both Dynamic Range Adjustment (DRA) and Running Spectrum Filtering (RSF) has been developed [3] - [5]. The noise-biased speech features can be compensated by using these methods and thus its recognition accuracy can be improved under many noise circumstances.

The automatic speech rejection has been used for the elimination of speech-like sounds and non-target speech phrases. The ASR in this system only concentrates a target phrase, e.g., a command word, and thus the set of these command words has a limited number of phrases. The system considers they are target and registered into the dictionary. Any other phrases are regarded as the out-of-dictionary words and they are rejected by this system. This mechanism has been realized

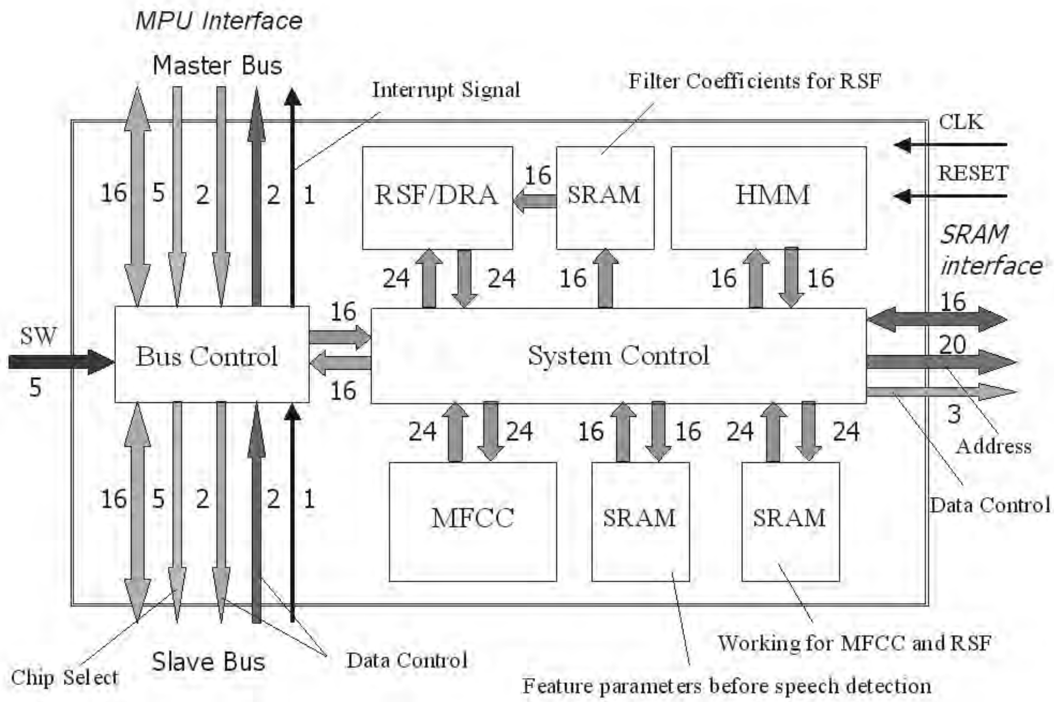


Figure 2: Block Diagram of Speech Communication System

by using several criterions for the likelihood values observed from HMM in the recognition stage. These criterions are applied for each target phrase and the optimum conditions are determined by using the speech database used in a training stage.

Environment	Noise Level	Accuracy
Meeting Room	50 dB	96.4%
Elevator	50 dB	95.0%
Stairs	45 dB	85.1%
Car A (Idling, No-Moving)	50 dB	99.4%
Car B (High Speed, Open Window)	75 dB	93.3%
Car C (High Speed, Audio ON (FM))	75 dB	88.9%

Figure 3: Accuracy of Speech Communication

3 LSI Design of Speech Communication System

As another important property in human-machine interface, a response time should be considered. When a general human interaction is considered, 100—200msec responses can keep a real-time response. In other words, if the system can respond its answer to a user within 150msec, a user feels its response becomes real-time. Compared with other real-time processing, i.e., parallel/pipeline computer, wireless communication and multi-media machine to machine communications, the time period of 150msec seems to be slow and thus it is not difficult to design such systems. However, when we would like to design the system with low power consumption and real-time processing at the same time, its design becomes complicated. In order to realize both of them, we have developed a full custom LSI for our speech communication system. Fig.2 shows its block diagram.

The architecture consists of HMM calculation module, noise robust processing module, and MFCC calculation module. In addition, the master and slave connections are given through BUS CONTROL and thus several same LSIs can be connected in a system.

The first designed LSI was fabricated with Rohm CMOS 0.35 □*m* in 2000. It can respond its recognition result within 0.180msec/phrase, 10MHz clock and 93.2mW power consumption where 1,000 phrases are recognized by using this LSI. The real time processing has been realized for 1,000 target phrases. Using parallel connections of these LSIs, the system can recognize more than 1,000 phrases within the same response time. Using this LSI, a small ASR evaluation board has been developed. Its specification is as follows:

1. Phrase ASR system where 1,000 maximum number of phrases can be registered.

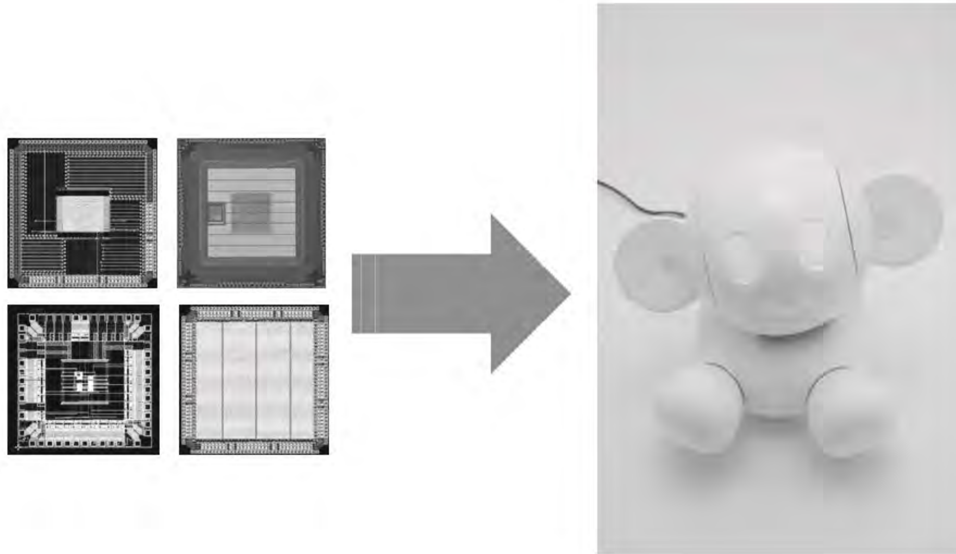


Figure 4: Chapit

2. FLASH memory of NOR type 32 bit and 16 bit ADC with at most 44.4 kHz sampling frequency are embedded.
3. The host connections are SPI, I2C, and UART.
4. The maximum clock is 22.5792 MHz and 3.3v power voltage is supplied. The 180 msec response time can be realized with 10MHz clock.
5. The small scale size is L: 55mm x W: 44mm x H: 12mm.

Using the evaluation board, the results of the field experiments are shown in Fig.3. In these experiments, the microphone distances are 30cm, 60cm and 90cm where these distances are generally required by many users. The total number of phrases is 15. For the experiments on echo robustness, the circumstances on a meeting room, an elevator and a stairs hall have been used. For the experiments of noise robustness, the circumstances on three different conditions of cars have been used.

As the total performance on echo and noise circumstances, the developed ASR system can realize 93.0% correctness (speech recognition accuracy and rejection accuracy). When we use this system in exhibitions, the system can recognize all words under the condition on 70 - 75 noisy level and 2 - 5 m microphone distance.

4 Speech Communication Module Embedded Smart Info-media System

The introduced speech communication module is embedded into a small robot. It is shown in Fig.4. The total system of this small robot has been designed by Raytron [6]. The two microphones at the ear of this robot are integrated and thus a sound in front of this robot is mainly recorded and analyzed.

The name of this robot is CHAPIT and it was announced in the late year of 2007. It can recognize only 300 words at the first time. However, a user can speak any words against the robot with 30cm—5m distance and 5 - 30 dB noisy/echo circumstances and then the robot answer for the registered phrases. Its interface style seems to be quite familiar among human to human communications.

5 Conclusion

In this letter, a speech communication system has been introduced. It can provide high accuracy of speech recognition under noisy and echo circumstances. The robust techniques are implemented into automatic speech detection, automatic speech recognition and automatic speech rejection. As current speech recognition accuracy, the system can realize 85% - 99% under 0 - 20 dB SNR and 150msec echo environments.

References

[1] Sadaoki Furui, "Speaker-Independent isolated word recognition using dynamic features of speech spectrum", IEEE Trans. on Acoust., Speech, and Signal Process., vol.ASSP-34, no.1 p.p.52-59, Feb. 1986.

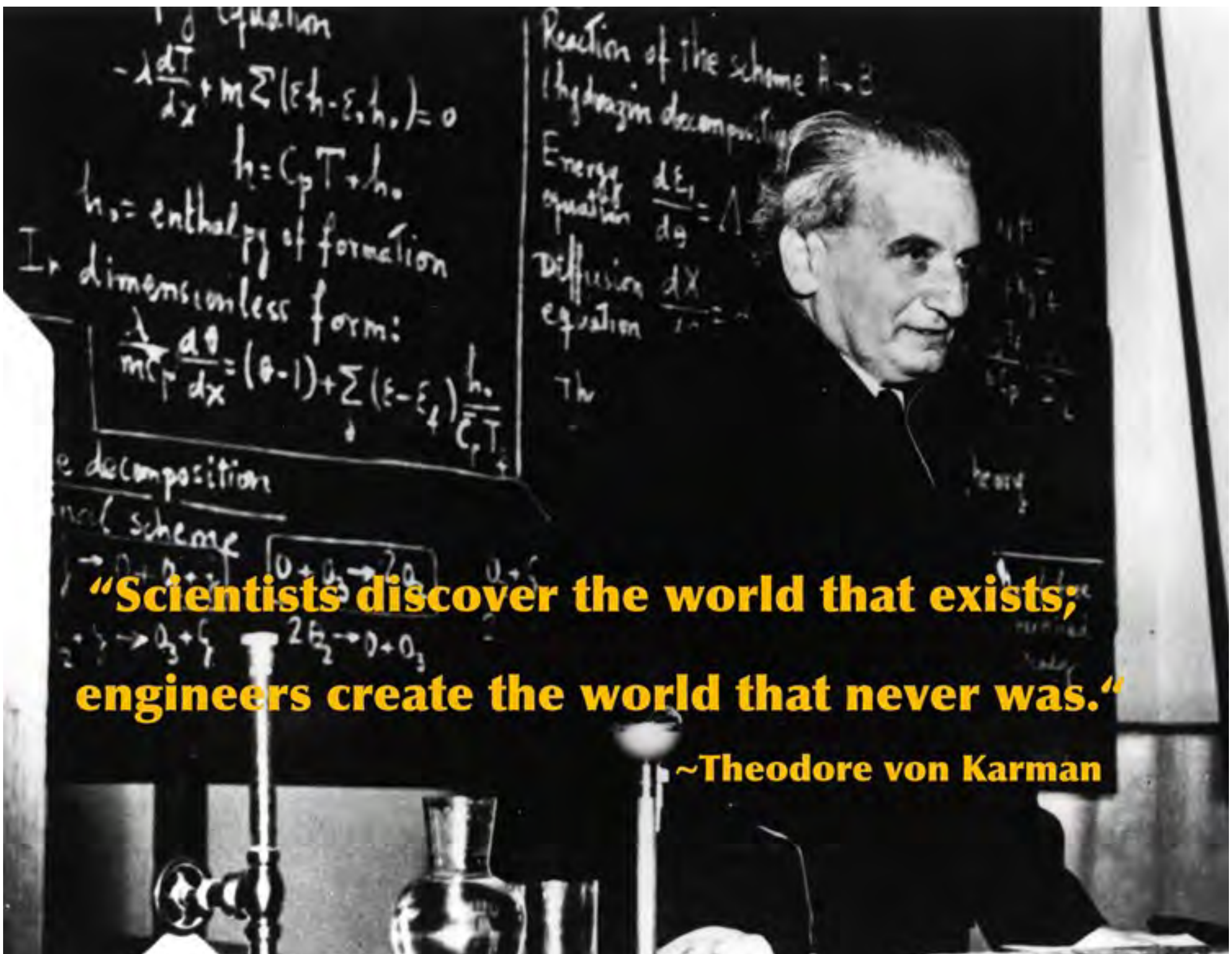
[2] A. Lee and T. Kawahara, " Recent Development of Open-Source Speech Recognition Engine Julius, " Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, pp.131-137, Oct., 2009.

[3] S.Yoshizawa, Y.Miyanaga, N.Wada and N.Yoshida: "A Low- Power LSI Design of JapaneseWord Recognition System", Proceedings of IEICE International Technical Conference on Circuits/Systems, Computers and Communications 2002, 1 (1): 98-101 (2002)

[4] Shingo Yoshizawa, Naoya Wada, Noboru Hayasaka, Yoshikazu Miyanaga, "Scalable Architecture for Word HMM-Based Speech Recognition and VLSI Implementation in Complete System", IEEE Transactions on Circuits and Systems I, Vol.53, No.1, pp.70-77, January 2006.

[5] Yoshikazu Miyanaga, Wataru Takahashi, Shingo Yoshizawa, "A Robust Speech Communication into Smart Info-Media System," IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, Vol. E96-A, No. 11, pp. 2074-2080, Nov. 2013.

[6] Yoshiyuki Miyazaki, Yoshikazu Miyanaga,"New Development and Enterprise of Robust Speech Recognition Systems ", 2010 International Workshop on Information Communication Technology, August 2010.



- Book Review -

The Road to Scientific Success:

Inspiring Life Stories of Prominent Researchers

Edited by **Deborah D L Chung** (University at Buffalo, State University of New York, USA); email: ddlchung@buffalo.edu

This book series describes the road to scientific success, as experienced and described by prominent researchers. The focus on research process (rather than research findings) and on personal experience is intended to encourage the readers, who will be inspired to be dedicated and effective researchers.

The objectives of this book series include the following.

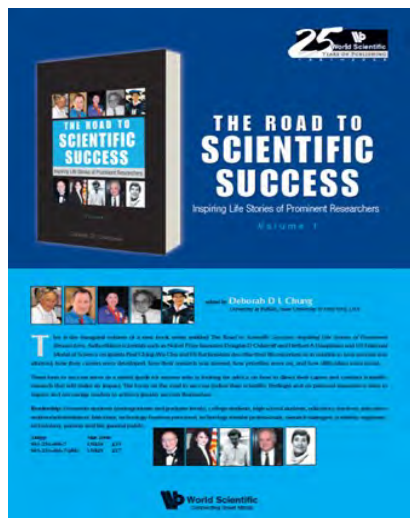
- To motivate young people to pursue their vocations with rigor, perseverance and direction
- inspire students to pursue science or engineering
- To enhance the scientific knowledge of students that do not major in science or engineering
- To help parents and teachers prepare the next generation of scientists or engineers
- To increase the awareness of the general public to the advances of science
- To provide a record of the history of science

Due to the historic significance of the life stories and the impact of the scientific advances behind each story, this book is expected to be valuable from the viewpoint of the history of science.

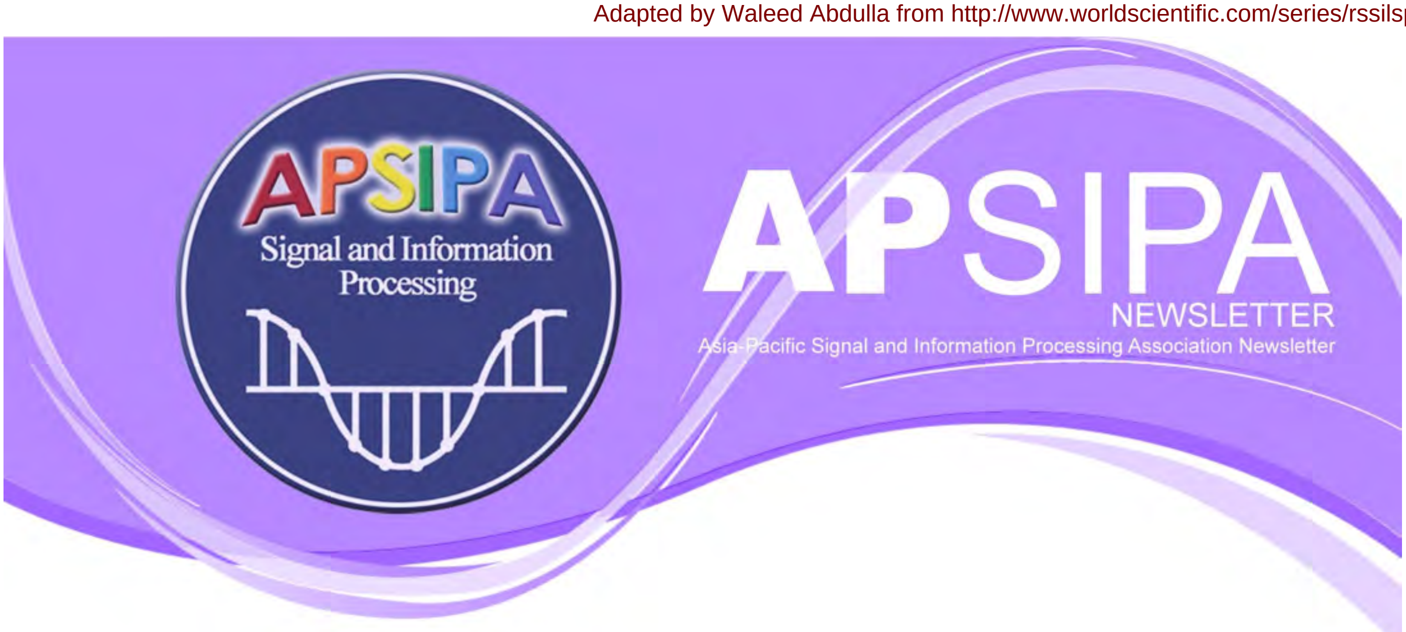
The uniqueness of the book series relates to the following.

- Untold life stories in the words and photos of world-class scientists
- Keys to scientific success elucidated
- Scientific research strategies disclosed
- Career preparation methods demonstrated
- History of science revealed

Up-to-date science covered in a way that readers with little or no science background can understand. This book series is suitable for use in courses on introduction to science or engineering, professional success methodology, scientific research methodology and history of science. Suitable readers include university students (undergraduate and graduate levels), college students, high school students, educators, teachers, education workers/administrators, historians, technology business personnel, technology transfer professionals, research managers, scientists, engineers, technicians, parents and the general public.



Adapted by Waleed Abdulla from http://www.worldscientific.com/series/rssilspr



Asia-Pacific Signal and Information Processing Association Annual Summit and Conference 2014

December 9-12, 2014, Chiang Mai, Thailand

Call for Papers

Welcome to the APSIPA Annual Summit and Conference 2014 located in Chiang Mai, the most culturally significant city in northern Thailand. Chiang Mai is a former capital of the Kingdom of Lanna (1296-1768) and is well known of historic temples, arresting scenic beauty, distinctive festivals, temperate fruits and invigorating cool season climate. The sixth annual conference is organized by Asia-Pacific Signal and Information Processing Association (APSIPA) aiming to promote research and education on signal processing, information technology and communications. The annual conference was previously held in Japan (2009), Singapore (2010), China (2011), USA (2012) and Taiwan (2013). The field of interest of APSIPA concerns all aspects of signals and information including processing, recognition, classification, communications, networking, computing, system design, security, implementation, and technology with applications to scientific, engineering, and social areas.

The regular technical program tracks and topics of interest include (but not limited to):

1. Biomedical Signal Processing and Systems (BioSiPS)
- 1.1 Biomedical Imaging
- 1.2 Modeling and Processing of Physiological Signals (EKG, MEG, EKG, EMG, etc.)
- 1.3 Biologically-inspired Signal Processing
- 1.4 Medical Informatics and Healthcare Systems
- 1.5 Genomic and Proteomic Signal Processing

2. Signal Processing Systems: Design and Implementation (SPS)
- 2.1 Nanoelectronics and Gigascale Systems
- 2.2 VLSI Systems and Applications
- 2.3 Embedded Systems
- 2.4 Video Processing and Coding
- 2.5 Signal Processing Systems for Data Communication

3. Image, Video, and Multimedia (IVM)
- 3.1 Image/video Coding
- 3.2 3D image/video Processing
- 3.3 Image/video Segmentation and Recognition
- 3.4 Multimedia Indexing, Search and Retrieval
- 3.5 Image/video Forensics, Security and Human Biometrics
- 3.6 Graphics and Animation
- 3.7 Multimedia Systems and Applications

4. Speech, Language, and Audio (SLA)
- 4.1 Speech Processing: Analysis, Coding, Synthesis, Recognition and Understanding
- 4.2 Natural Language Processing: Translation, Information Retrieval, Dialogue
- 4.3 Audio Processing: Coding, Source Separation, Echo Cancellation, Noise Suppression
- 4.4 Music Processing

5. Signal and Information Processing Theory and Methods (SIPTM)
- 5.1 Signal Representation, Transforms and Fast Algorithms
- 5.2 Time Frequency and Time Scale Signal Analysis
- 5.3 Digital Filters and Filter Banks
- 5.4 DSP Architecture
- 5.5 Statistical Signal Processing
- 5.6 Adaptive Systems and Active Noise Control
- 5.7 Sparse Signal Processing
- 5.8 Signal Processing for Communications
- 5.9 Signal Processing for Energy Systems
- 5.10 Signal Processing for Emerging Applications

6. Wireless Communications and Networking (WCN)
- 6.1 Wireless Communications: Physical Layer
- 6.2 Wireless Communications and Networking: Ad-hoc and Sensor Networks, MAC, Wireless Routing and Cross-layer Design
- 6.3 Wireless Networking: Access Network and Core Network
- 6.4 Security and Cryptography
- 6.5 Devices and Hardware

Submission of Papers

Prospective authors are invited to submit either full papers, up to 10 pages in length, or short papers up to 4 pages in length, where full papers will be for the single-track oral presentation and short papers will be mostly for poster presentation. The conference proceedings of the main conference will be published, available and maintained at the APSIPA website.

Important Dates

Submission of Proposals for Special Sessions, Forum, Panel & Tutorial Sessions	May 9, 2014
Submission of Full and Short Papers	June 6, 2014
Submission of Papers in Special Sessions	July 4, 2014
Notification of Papers Acceptance	Aug. 29, 2014
Submission of Camera Ready Papers	Sep. 26, 2014
Author Registration Deadline	Sep. 26, 2014
Tutorial Session Date	Dec. 8, 2014
Summit and Conference Dates	Dec. 9-12, 2014

E-mail : secretary@apsipa2014.org

Organizer
Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI) Association of Thailand

Academic Sponsor
Asia-Pacific Signal and Information Processing Association (APSIPA)

Organizing Committees

Honorary Co-Chairs
Sadaoki Furui, Tokyo Institute of Technology, Japan
K. J. Ray Liu, University of Maryland, USA
Prayoot Akkaraekthalin, KMUTNB, Thailand

General Co-Chairs
Kosin Chamnongthai, KMUTT, Thailand
C.-C. Jay Kuo, University of Southern California, USA
Hitoshi Kiya, Tokyo Metropolitan University, Japan

Technical Program Co-Chairs
Pornchai Supnithi, KMITL, Thailand
Takao Onoye, Osaka University, Japan
Hsueh-Ming Hang, National Chiao Tung University, Taiwan
Anthony Kuh, University of Hawaii at Manoa, USA
Takeshi Ikenaga, Waseda University, Japan
Chung-Hsien Wu, National Cheng Kung University, Taiwan
Yodchanan Wongsawat, Mahidol University, Thailand
Oscar Au, HKUST, Hong Kong
Tomoaki Ohtsuki, Keio University, Japan

Forum Co-Chairs
Antonio Ortega, University of Southern California, USA
Waleed Abdulla, The University of Auckland, New Zealand
Homer Chen, National Taiwan University, Taiwan
Vorapoj Patanavijit, Assumption University, Thailand

Panel Session Co-Chairs
Mark Liao, IIS, Academia Sinica, Taiwan
Li Deng, Microsoft Research, USA
Jiwu Huang, Sun Yat-Sen University, China
Kazuya Takeda, Nagoya University, Japan

Special Session Co-Chairs
Minoru Okada, Nara Institute of Science and Technology, Japan
Gan Woon Seng, Nanyang Technological University, Singapore
Mrityunjoy Chakraborty, IIT Kharagpur, India

Tutorial Session Co-Chairs
Kenneth Lam, The Hong Kong Polytechnic University, Hong Kong
Toshihisa Tanaka, TUAT, Japan
Tatsuya Kawahara, Kyoto University, Japan
Sumei Sun, I^2R, A*STAR, Singapore

Publicity Co-Chairs
Yoshio Itoh, Tottori University, Japan
Yo-Sung Ho, Gwangju Institute of Science and Technology, Korea
Thomas Fang Zheng, Tsinghua University, China
Chung-Nan Lee, National Sun Yat-sen University, Taiwan
Chalie Charoenlarpnopparut, Thammasat University, Thailand

Publication Co-Chairs
Yoshinobu Kajikawa, Kansai University, Japan
Nipon Theera-umpon, Chiangmai University, Thailand,

Financial Chairs
Rujipan Sampanna, Bangkok University, Thailand
Pairin Kaewkuay, ECTI, Thailand

Local Arrangement Chairs
Suttichai Premrudeeprechacharn, Chiangmai University, Thailand
Sermsak Uatrongjit, Chiangmai University, Thailand
Sathaporn Promwong, KMITL, Thailand

General Secretaries
Werapon Chiracharit, KMUTT, Thailand
Boonserm Kaewkamnerdpong, KMUTT, Thailand

19th DSP 2014

International Conference on Digital Signal Processing

Imperial College London

HONG KONG

IET The Institution of Engineering and Technology · IEEE · IEEE Signal Processing Society · EURASIP · APSIPA · The Hong Kong Polytechnic University

20-23 August 2014

dsp2014.org

Important Deadlines

Submission of full papers
March 31, 2014

Notification of acceptance:
April 21, 2014

Submission of camera-ready papers:
May 21, 2014

The 2014 19th International Conference on Digital Signal Processing (DSP 2014), sponsored by Imperial College, London, and co-sponsored by the IEEE Signal Processing Society, the Hong Kong Polytechnic University, EURASIP, APSIPA, and IET, will be held in August 20-23, 2014 in Hong Kong. It is the longest in existence Conference in the area of DSP and belongs to a series of events, which commenced in London in 1968 and continued to Florence, Nicosia, Limassol, Santorini, Cardiff, Corfu. The last meeting took place in Santorini, Greece in July 2013. DSP 2014 addresses the theory and application of filtering, coding, transmitting, estimating, detecting, analysing, recognising, synthesising, recording, and reproducing signals by means of digital devices or techniques. The term "signal" includes audio, video, speech, image, communication, geophysical, sonar, radar, medical, musical, and other signals.

Indicative topics of interest:

Biomedical Signal Processing
- Biomedical Signal and Image Processing
- Brain-Computer Interface
- Genomic Signal Processing
- Signal Processing in Genomics and Proteomics

Digital and Multirate Signal Processing
- Adaptive Signal Processing
- Digital and Multirate Signal Processing
- Digital Filter Design and Implementation
- Multidimensional Filters and Transforms
- Multiresolution Signal Processing
- Multiway Signal Processing
- Theory and Applications of Transforms
- Time-Frequency Analysis and Representation
- Statistical Signal Processing

Sensor Array and Multichannel Processing
- Array Signal Processing
- Signal Processing for Smart Sensors and Systems
- Compressive Sensing

Signal Processing for Communications
- Geophysical/Radar/Sonar Signal Processing
- MIMO Signal Processing

Signal Processing for Audio/Image /Video
- Audio/Speech/Music Processing & Coding
- Digital Photography
- HDR Imaging
- Image and Multidimensional Signal Processing
- Image/Video Indexing, Search and Retrieval
- Image/Video Compression and Coding Standards
- Image/Video Content Analysis
- Image/Video Processing Techniques
- 3D Image Processing and Applications
- Mobile Imaging and Image Quality
- Real-Time Signal/Image/Video Processing
- Video Surveillance and Transportation Imaging
- Digital Watermarking and Data Hiding

Other Areas and Applications
- Big Data
- Cognitive Signal Processing
- DSP Education
- Nonlinear Signals and Systems
- Information Forensics and Security
- Internet of Things (IoT)
- Social Signal Processing & Affective Computing
- Signal and System Modelling
- Signal Processing of Financial Data
- VLSI Architectures and Implementations for DSP

The language of the Conference is English. Prospective authors are invited to electronically submit the full camera-ready paper. The paper must not exceed six (6) pages, including figures, tables and references, and should be written in English. Authors should indicate one or more of the above listed categories that best describe the topic of the paper, as well as their preference (if any) regarding lecture or poster sessions. Lecture and poster sessions will be treated equally in terms of the review process. Submitted papers will be peer-reviewed by at least two experts in the field. All accepted and presented papers will be included in the Conference Proceedings and will be submitted for inclusion into IEEE Xplore and EI Compendex. Please refer to the DSP 2014 website www.dsp2014.org for additional information. In addition to the technical program, a social program will be offered to the participants and their companions. It will provide an opportunity to meet colleagues and friends against a backdrop of outstanding natural beauty and rich cultural heritage in one of the best known international tourist destinations.

Honorary Chair
A. G. Constantinides, Imperial College, London, UK

General Chairs
Yong Ching Lim, Nanyang Technological Univ. of Singapore
Daniel P.K. Lun, Hong Kong Polytechnic University, Hong Kong
A. N. Skodras, U of Patras, Greece
Danilo Mandic, Imperial College London, UK

Technical Program Chairs:
Jonathon Chambers, Loughborough University, UK
H.K. Kwan, University of Windsor, Canada
Yui-Lam Chan, Hong Kong Polytechnic University, HK (Coordinator)

Conference Secretary:
Bonnie N.F. Law, Hong Kong Polytechnic University, HK

Constantinides Workshop Chair:
Yong Ching Lim, Nanyang Technological University of Singapore

Special Session Chair:
Kenneth K.M. Lam, Hong Kong Polytechnic University
Shing Chau Chan, University of Hong Kong, HK
Hing Cheung So, City University of Hong Kong, HK
T. Blu, Chinese University of Hong Kong, HK

Finance Chair:
Chris, Y.H. Chan, Hong Kong Polytechnic University

Publicity Chairs:
Thanos Gkelias, Imperial College London, UK
Y. Andreopoulos, UCL, UK
Tan Lee, Chinese University of Hong Kong, HK

Publications Chairs:
Lap-Pui Chau, Nanyang Technological Univ. of Singapore

Local Arrangements Chair:
Frank Leung, Hong Kong Polytechnic University
Edward Cheung, Hong Kong Polytechnic University, HK

Steering Committee
Tieniu Tan (Chair) Chinese Academy of Sciences, China
F. Harris, San Diego State U, USA
Kin K Leung, Imperial College London, UK
Danilo Mandic, Imperial College London, UK
Vince Poor, Princeton University, USA
Wan-Chi Siu, Hong Kong Polytechnic University, Hong Kong

International Program Committee

D. Androutsos, Canada; J. Apostolopoulos, USA; G. Bebis, USA; K. Berberidis, Greece; A. Bharath, UK; N. Bourbakis, USA; V. Cappellini, Italy; F. Castanie, France; J. Chambers, UK; J. Cornelis, Belgium; C. F. N. Cowan, UK; J. Dittmann, Germany; M. Domanski, Poland; F. Dufaux, France; T. Ebrahimi, Switzerland; A. Ercil, Turkey; A. Floros, Greece; M. Gabbouj, Finland; M. Ghanbari, UK; G. Giannakis, USA; D. Hatzinakos, Canada; A. T. S. Ho, UK; A. K. Katsaggelos, USA; E. Kofidis, Greece; S. Kollias, Greece; C. Kotropoulos, Greece; G. Lampropoulos, Canada; I. Langendijk, The Netherlands; S. Z. Li, China; Y. C. Lim, Singapore; S. Loncaric, Croatia; K.Ma, Hong Kong; H. Lu, Hong Kong; B. Macq, Belgium; D. Mandic, UK; A. Manikas, UK; P. Maragos, Greece; D. Maroulis, Greece; R. Molina, Spain; J. Mourjopoulos, Greece; A. Munteanu, Belgium; P. Nasiopoulos, Canada; K. Nikita, Greece; B. Ottersten, Sweden; C. Papadias, Greece; P. Papamichalis, USA; T. Pappas, USA; C. Pattichis, Cyprus; F. Pereira, Portugal; B. Pesquet-Popescu, France; I. Pitas, Greece; K. N. Plataniotis, Canada; G. Ramponi, Italy; Y. M. Ro, Korea; A. Rontogiannis, Greece; M. Rupp, Austria; S. Sanei, UK; A. H. Sayed, USA; P. Schelkens, Belgium; N. Sidiropoulos, Greece; W. C. Siu, Hong Kong; A. Spanias, USA; T. Stathaki, UK; T. Stouraitis, Greece; L. Torres, Spain; T. Trump, Estonia; A. Venetsanopoulos, Canada; L. Verdoliva, Italy; M. Vetterli, Switzerland; T. Vlachos, Greece; H. R. Wu, Australia; L. Yaroslavsky, Israel; D. Zhang, Hong Kong

APSIPA Who's Who

President: C.-C. Jay Kuo, University of Southern California, USA

President-Elect: Haizhou Li, Institute for Infocomm Research, A*STAR, Singapore

Past President: Sadaoki Furui, Tokyo Institute of Technology, Japan.

VP - Conferences: Susanto Rahardja, Institute for Infocomm Research, Singapore

VP - Industrial Relations and Development: Haohong Wang, TCL Corporation, USA

VP - Institutional Relations and Education Program:
Thomas Zheng, Tsinghua University, China

VP - Member Relations and Development: Kenneth Lam, The Hong Kong Polytechnic University, Hong Kong

VP - Publications: Tatsuya Kawahara, Kyoto University, Japan

VP - Technical Activities: Oscar Au, Hong Kong University of Science and technology,
Hong Kong

Members-at-Large:

Waleed H. Abdulla, The University of Auckland, New Zealand

Kiyoharu Aizawa, The University of Tokyo, Japan

Mrityunjoy Chakraborty, Indian Institute of Technology, India

Kosin Chamnong, King Mongkut's University of Technology, Thailand

Homer Chen, National Taiwan University, Taiwan

Woon-Seng Gan, Nanyang Technological University, Singapore

Hsueh-Ming Hang, National Chiao-Tung University, Taiwan

Yo-Sung Ho, Gwangju Institute of Science and Technology, Korea

Jiwu Huang, Sun Yat-Sen University, China

Anthony Kuh, University of Hawaii at Manoa, USA

Tomoaki Ohtsuki, Keio University, Japan

Wan-Chi Siu, The Hong Kong Polytechnic University, Hong Kong

Ming-Ting Sun, University of Washington, USA

Headquarters

Address:

Asia Pacific Signal and Information Processing Association,
Centre for Signal Processing,
Department of Electronic and Information Engineering,
The Hong Kong Polytechnic University,
Hung Hom, Kowloon, Hong Kong.

Officers:

Director: Wan-Chi Siu, email: enwcsiu@polyu.edu.hk
Manager: Kin-Man Lam, Kenneth,
email: enkmlam@polyu.edu.hk
Secretary: Ngai-Fong Law, Bonnie,
email: ennflaw@polyu.edu.hk
Treasurer: Yuk-Hee Chan, Chris,
email: enyhchan@polyu.edu.hk

APSIPA NEWSLETTER
Asia-Pacific Signal and Information Processing Association Newsletter

APSIPA Signal and Information Processing