Issue 5

Jan 2014

APSIPA NEWSLETTER

Asia-Pacific Signal and Information Processing Association Newsletter

APSIPA ASC 2013

APSIPA Annual Summit and Conference (ASC) 2013 was held in Kaohsiung, Taiwan, on October 29 to November 1, 2013. It was the 5th APSIPA annual conference. This year, we had the conference in the 85 sky tower hotel in Kaohsiung, Taiwan, and got a lot of excellent and dedicated scholars. The conference had high turnouts, which were over 410 attendees.

The Papers submission and presentations were organized in 6 separate tracks, corresponding to the technical areas covered by six APSIPA technical committees. Among them, the Speech, Language and Audio (SLA) track and the Image, Video and Multimedia (IVM) track attracted the largest numbers of submissions. For this conference, we also got the technical sponsorship from the IEEE Signal Processing Society, and all accepted papers are accessible via IEEE Xplore.

The technical program included 8 tutorial sessions, 3 keynote speeches, 2 plenary sessions (one on the **Speech, Language, and Audio (SLA)** and the other on **Image, video and multimedia processing(IVM)**), one forum discussion session (on the interaction of academia and industry), together with 48 oral sessions. Three best papers awards were announced in the conference banquet.

They are:

The Best Paper Award (IVM Track): Masaki Aono, Hitoshi Koyanagi and Atsushi Tatsuma, "3D Shape Retrieval focused on Holes and Surface Roughness"

The Best Paper Award (SLA Track): Ryoichi Miyazaki, Hiroshi Saruwatari, Satoshi Nakamura, Kiyohiro Shikano, Kazunobu Kondo, Jonathan Blanchette and Martin Bouchard , "Toward Musical-Noise-Free Blind Speech Extraction: Concept and Its Applications"

In this issue

APSIPA in Quick!

APSIPA Mission: To promote broad spectrum of research and education activities in signal and information processing in Asia Pacific.

APSIPA Conferences: ASPIPA Annual Summit and Conference.

APSIPA Publications: Transactions on Signal and Information Processing in partnership with Cambridge Journals since 2012; APSIPA Newsletters.

APSIPA Social Network: To link members together and to disseminate valuable information more effectively.

APSIPA Distinguished Lectures: An APSIPA educational initiative to reach out to the community.

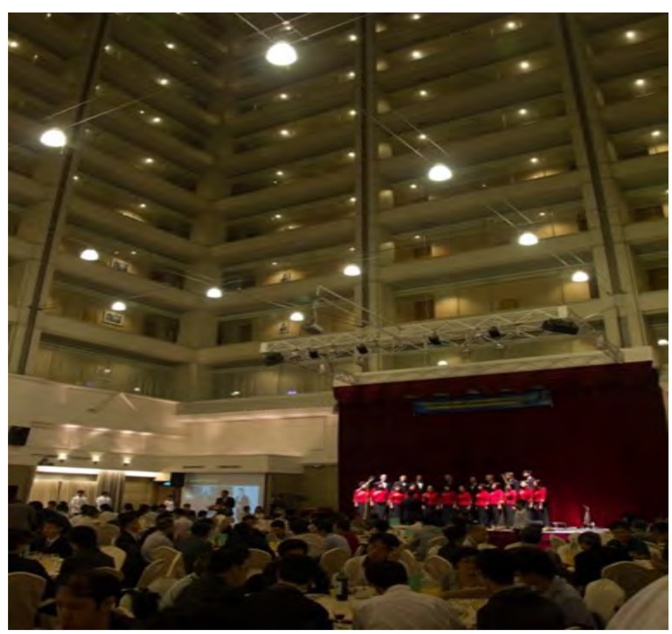
Registration Desks

Introducing Keynote Speaker

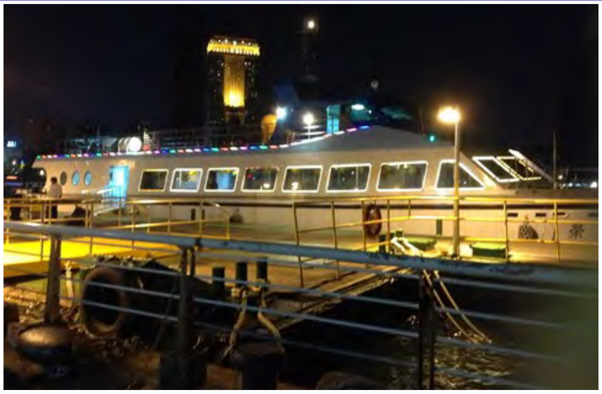
The Best Paper Award (WCN+SIPTM Tracks): Yung-Shun Wang, Y.-W. Peter Hong and Wen-Tsuen Chen, "Sum-Rate Maximization and Energy-Cost Minimization for Renewable Energy Empowered Base-Stations using Zero-Forcing Beamforming"

During the conference, we prepared a series of social programs, including a welcome reception, a conference banquet with a unique cultural performance for all registrants. There were also a cruise banquet and buffet banquet.

The night for the conference banquet was amazing and amusing for all participants. There are the wonderful musical numbers as well as the stunning and inspiring drum dancing arts. The cruise banquet went with the touring around Kaohsiung harbour, which is the largest international Commercial port in Taiwan and is one of the prominent container ports in the world.



Conference Banquet



Cruise Dinner and Tour

During the tour, participants saw various scenic spots and came to know the historical and natural facts of Kaohsiung harbor and its surrounding area with our dedicated English touring guide.

APSIPA ASC 2013 was successfully concluded at the afternoon of November 1st. It is your participation that makes APSIPA ASC 2013 successful. We welcome your contributions to ASIPA ASC 2014, which will be held in Chiang Mai, Thailand, December 9 - 12, 2014. We are looking forward to seeing old as well as new friends there.

About APSIPA

APSIPA Development Prevue

The Asia-Pacific Signal and Information Processing Association (APSIPA) is a non-profit organization with the following objectives as its mission:

- Providing education, research and development exchange platforms for both academia and industry
- Organizing common-interest activities for researchers and practitioners
- Facilitating collaboration with region-specific focuses and promoting leadership for worldwide events
- Disseminating research results and educational material via publications, presentations, and electronic media
- Offering personal and professional career opportunities with development information and networking

The idea of APSIPA was born in Hawaii in 2007. It was formally incorporated as "Asia Pacific Signal and Information Processing Association Limited" in Hong Kong on July 23, 2009. Dr. Sadaoki Furui was elected as the first President and all other BoG members were elected at the BoG meeting held on the first day of APSIPA ASC 2009, October 4, 2009. All BoG members were approved at the first Annual General Meeting held on October 5, 2009, and APSIPA was officially established.

Major Events in APSIPA Preparation Stage

April 19, 2007: the initialization meeting was held in Honolulu, Hawaii, USA

December 8, 2007: the 1st APSIPA Steering Committee Meeting was held in Tokyo Institute of Technology, Tokyo, Japan

April 3, 2008: the 2nd APSIPA Steering Committee Meeting was held in Las Vegas, Nevada, USA

December 13, 2008: the 3rd APSIPA Steering Committee Meeting was held in the Hong Kong Polytechnic University, Kowloon, Hong Kong

April 23, 2009: the 4th APSIPA Steering Committee Meeting was held in Taipei, Taiwan

APSIPA Milestones

October 4-7, 2009

2009 APSIPA Annual Summit and Conference was held in Sapporo, Japan

Dr. Sadaoki Furui was elected as the first President of APSIPA

The first APSIPA Board of Governors was installed

December 14-17, 2010: 2010 APSIPA Annual Summit and Conference was held in Biopolis, Singapore

October 18-21, 2011: 2011 APSIPA Annual Summit and Conference was held in Xi'an, China

December, 2011: The APSIPA Social Network Program was launched

January, 2012: The APSIPA Distinguished Lecturer Program was launched

April, 2012: The 1st issue of APSIPA Newsletters was published

May, 2012: The open-access journal "APSIPA Transactions on Signal and Information Processing", published by the Cambridge University Press, was launched

December 3-6, 2012: 2012 APSIPA Annual Summit and Conference was held in Los Angeles, California, USA

Dr. C.-C. Jay Kuo succeeded Dr. Furui to become the 2nd President of APSIPA

September 1, 2013: The APSIPA Friend Labs Program was launched

October 29-November 1, 2013: 2013 APSIPA Annual Summit and Conference was held in Kaohsiung, Taiwan

November, 2013: the first APSIPA Advisory Board was established

An Insight into Dynamic Aspects of Speech Emotional Expression

Chung-Hsien Wu, Jen-Chun Lin, and Wen-Li Wei
Department of Computer Science and Information Engineering
National Cheng Kung University, Tainan, Taiwan
{chunghsienwu, jenchunlin, lilijinjin}@gmail.com

1. Introduction

Emotion recognition is the ability to identify what you are feeling from moment to moment and to understand the connection between your feelings and your verbal/non-verbal expressions. Intact perception of emotion is vital for communication in the social environment. Although various studies in emotion recognition from speech, text or facial expression have shown the benefits using different features and classifiers [1-5], toward high-performance emotion recognition, an important issue is the modeling of dynamic aspects of emotional expression. In a natural conversation, a complete emotional expression is typically composed of dynamic aspects representing temporal phases of onset, apex, and offset. In general, when the temporal course of emotional expression is complex, the temporal information could be lost owing to inappropriate model structures which may lead to inaccurate estimates of the statistical model parameters and therefore result in unsatisfactory classification [6-10]. To capture the temporal information, many design issues regarding the structure and the training process of the hidden Markov model (HMM) have been investigated.

In most HMM-based emotion recognition schemes, the left-to-right topology of the HMM structure was used [11-13], and has proven useful in modeling the speech signal for describing the temporal courses of emotional expressions. However, it may be invalid for utterance-based emotion recognition, especially in natural conversation. Typically, a complete emotional expression is expressed by more than one utterance in natural conversation, and in more detail, each utterance may contain several temporal phases of emotional expression. Figure 1 shows that when the emotional state (i.e., happiness) of Speaker 1 is evoked through conversation, Utterance 1 only covers the temporal phase of onset, while the apex and offset phases are covered in Utterance 2. Effective modeling of dynamic aspects in a real conversational environment is desirable to model the complex temporal structure in speech emotional expression.

2. Dynamic Aspect Modeling

For complex temporal structure modeling, in this article, a sequence of M temporal phases, each modeled by a sub-emotion HMM, is used to characterize an emotional state expressed in an isolated sentence. Besides temporal courses, emotional expression is further characterized by low/high intensity as shown in Figure 2. In order to better describe the temporal course of an emotional expression, a sub-emotion transition model similar to the language model widely and successfully used in speech recognition is employed. By integrating the sub-emotion transition model, the proposed temporal course modeling scheme can further provide a constraint on allowable temporal phase sequences to determine an optimal emotional state in an isolated utterance.

Given the observation sequence od $O = o_1^T = o_1, o_2, ..., o_T$, the probability of an emotional state with temporal phase sequence E can be estimated using (1), where $\hat{E}$ represents the emotion recognition result by maximizing the a posteriori probability $P(E|O)$.

$$\hat{E} = \arg\max_{E} P(E \mid O) \qquad (1)$$

The a posteriori probability $P(E|O)$ can be further decomposed and simplified using the Bayes' rule to obtain the emotion recognition result as follows:

$$\hat{E} = \arg\max_{E} P(O \mid E) P(E) \qquad (2)$$

Where $P(O|E)$, denoting the likelihood of the observation, is calculated using the corresponding sub-emotion HMM sequence. $P(E)=P(e_1,e_2,...,e_M)$ is the a priori probability of observing temporal phase sequence $E=e_1, e_2, ..., e_M$ and is estimated by the sub-emotion transition model. A bigram transition model is adopted and constructed to estimate the probability

$$P(E) = P(e_1, e_2, ..., e_M) = \prod_{k=2}^{M} P(e_k \mid e_{k-1})$$

A recognition network for the pre-defined grammar shown in Figure 3 is constructed based on the temporal phase definition [7-9].

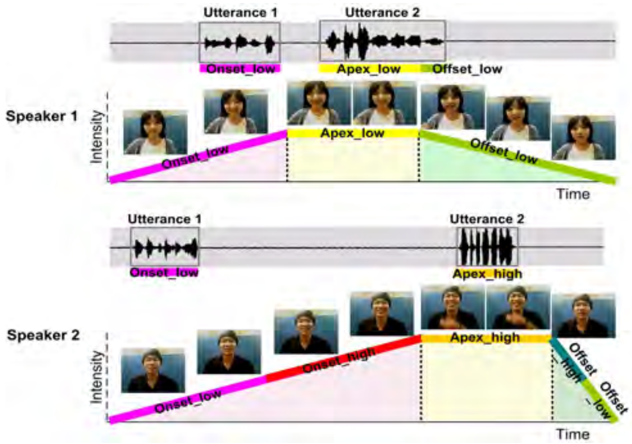


Figure 1: An example of various temporal phases of happy emotional expression occurred to different utterances in a real conversational environment.

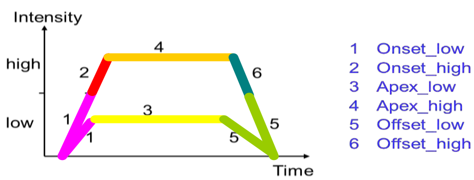


Figure 2: Emotional expression for temporal phases with different intensities.

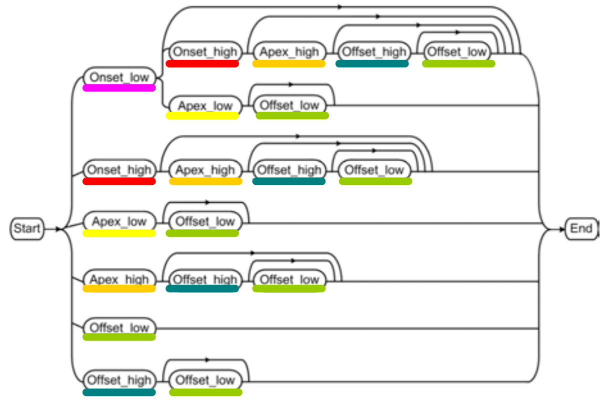


Figure 3: Recognition network based on the predefined grammar for characterizing an emotional state expressed in

3. Experimental Results

A conversation-based affective speech corpus [10] was collected from the Multimedia Human-Machine Communication (MHMC) Laboratory. The speech data were provided by 53 students of both genders in National Cheng Kung University, Taiwan. During the recording session, toward naturalistic conversation, a conversation topic was first selected by each paired participants, and for each topic, the participants spoke as they like instead of navigating a pre-designed script. For four emotional states (Happy, Angry, Sad and Neutral), a total of 2,120 utterances were collected to form the MHMC conversation-based affective speech corpus.

Subjective tests were performed to set the ground truth of emotional expression for the recorded data. Three annotators were recruited from the MHMC laboratory, and each of them was asked to give an opinion on the emotion label for the recorded data. During the labeling process, the annotators were allowed to check the emotion expression of the recorded data more than once to ensure that the labels can truly reflect their feelings. After the labeling process, each labeled data was then evaluated by checking the opinions of all annotators. If less than two annotators reached atn agreement, the data was not included in the experiment. Finally, a total of 1,114 data, which passed the evaluation (i.e., simultaneously passed the emotion and temporal phase labeling procedure), were regarded as the ground truth data for the ensuing experiments.

In the experiments, two popular classifiers were considered: Support Vector Machine (SVM) and HMM [5], [12], [13]. For performance comparison, the SVM with radial basis kernel function was used and the traditional HMM with left-to-right topology and eight hidden states was adopted to model the entire sentence for comparison. For SVM, the global features [5] were used in which the minimum, mean, and maximum of the extracted prosodic features were considered. In the proposed approach, a left-to-right HMM with three hidden states was applied for modeling each temporal phase. In the experiments, 80% of the ground truth utterances were selected from the MHMC conversation-based affective speech corpus for training, and the remaining utterances were selected for testing.

The average recognition accuracy for three approaches is shown in Table 1. The confusion matrix of the proposed approach is further shown in Table 2. The results in Table 1 show that comparing with SVM and the traditional HMM approaches, the proposed approach of the temporal course modeling achieved the best recognition accuracy. The results confirm that considering the temporal phases and combining the sub-emotion transition model is able to better describe the dynamic aspects of emotional expression in natural conversation. Compared to the previous approaches, with the property of expression intensity, the proposed temporal course modeling is helpful to alleviate the effect of diverse expression intensities and therefore to diminish the variations of model parameters and distinguish the expression styles.

4. Conclusion

This article presented an insight into the dynamic aspects of speech emotional expression. The complex temporal structure, characterizing the dynamic aspects of emotional expression, is helpful for improving the recognition accuracy. For effective emotion recognition, future research to explore the expression styles from different users is a viable direction, which may be further related to the expression manner and significantly associated to the personality trait.

Table 1 Average emotion recognition rates of four emotional states using different models.

Models	SVM	HMM	Temporal Course Modeling
Accuracy	50.22%	56.50%	79.82%

Table 2 Confusion matrix of the proposed temporal course modeling approach for four emotional states.

	Happy	Angry	Sad	Neutral	Accuracy
Happy	31	1	2	0	91.18%
Angry	18	25	1	1	55.56%
Sad	4	0	32	2	84.21%
Neutral	6	7	3	90	84.91%

References

[1] Z. Zeng, M. Pantic, G. I. Roisman and T. S. Huang, "A survey of affect recognition methods: audio, visual, and spontaneous expressions", IEEE Trans. Pattern Anal. Mach. Intell., 31(1):39-58, 2009.

[2] C. H. Wu and W. B. Liang, "Emotion recognition of affective speech based on multiple classifiers using acoustic-prosodic information and semantic labels", IEEE Trans. Affective Computing, 2(1):1-12, 2011.

[3] C. H. Wu, Z. J. Chuang and Lin, Y.C., "Emotion recognition from text using semantic label and separable mixture model", ACM Trans. on Asian Language Information Processing, 5(2):165-183, 2006.

[4] C. H. Wu, W. L. Wei, J. C. Lin and W. Y. Lee, "Speaking effect removal on emotion recognition from facial expressions based on eigenface conversion", IEEE Trans. Multimedia, DOI (identifier) 10.1109/TMM.2013.2272917, 2013.

[5] M. E. Ayadi, M. S. Kamel and F. Karray, "Survey on speech emotion recognition: features, classification, schemes, and databases", Pattern Recognition, 44 (3):572-587, 2011.

[6] R. W. Picard, Affective Computing. MIT Press, 1997.

[7] P. Ekman, Handbook of Cognition and Emotion. Wiley, 1999.

[8] M. F. Valstar, and M. Pantic, "Fully automatic recognition of the temporal phases of facial actions", IEEE Trans. Systems, Man and Cybernetics–Part B, 42(1):28-43, 2012.

[9] C. H. Wu, J. C. Lin and W. L. Wei, "Two-level hierarchical alignment for semi-coupled HMM-based audiovisual emotion recognition with temporal course", IEEE Trans. Multimedia, DOI (identifier) 10.1109/TMM.2013.2269314, 2013.

[10] W. L. Wei, C. H. Wu, J. C. Lin and H. Li, "Interaction style detection based on cross-correlation model in spoken conversation", to appear in Proc. Int'l Conf. Acoustics, Speech, and Signal Processing (ICASSP), 2013.

[11] J. C. Lin, C. H. Wu and W. L. Wei, "Error weighted semi-coupled hidden Markov model for audio-visual emotion recognition", IEEE Trans. Multimedia, 14 (1):142-156, 2012.

[12] B. Schuller, G. Rigoll and M. Lang, "Hidden Markov model-based speech emotion recognition", Proc. Int'l Conf. Acoustics, Speech, and Signal Processing (ICASSP), II 1-4, 2003.

[13] S. Ntalampiras and N. Fakotakis, "Modeling the temporal evolution of acoustic parameters for speech emotion recognition", IEEE Trans. Affective Computing, 3(1):116-125, 2012.

APSIPA in Quick!

APSIPA Mission: To promote broad spectrum of research and education activities in signal and information processing in Asia Pacific.

APSIPA Conferences: ASPIPA Annual Summit and Conference.

APSIPA Publications: Transactions on Signal and Information Processing in partnership with Cambridge Journals since 2012; APSIPA Newsletters.

APSIPA Social Network: To link members together and to disseminate valuable information more effectively.

Some Advances in Adaptive Source Separation

Jen-Tzung Chien*, Hiroshi Sawada† and Shoji Makino‡
* Department of Electrical and Computer Engineering, National Chiao Tung University, Hsinchu, Taiwan
† NTT Service Evolution Laboratories, NTT Corporation, Yokosuka-shi, Kanagawa, Japan
‡ Graduate School of Systems and Information Engineering, University of Tsukuba, Ibaraki, Japan
jtchien@nctu.edu.tw, sawada.hiroshi@lab.ntt.co.jp, maki@tara.tsukuba.ac.jp

I. Introduction

We are surrounded by sounds and noises in presence of room reverberation [15]. The observed mixed signals are usually less than source signals. The mixing condition is prone to be varied by the moving sources or in case of source replacement. It becomes challenging to estimate the desired audio and speech signals and develop a comfortable acoustic communication channel between humans and machines. Audio source separation in realistic conditions has been a fascinating avenue for research which is crucial for broad extensions and applications ranging from speech enhancement, speech recognition, music retrieval, sound classification, human-machine communication and many others. How to extract and separate a target audio or speech signal from noisy and nonstationary observations is now impacting the communities of signal processing and machine learning.

The traditional blind source separation (BSS) approaches based on independent component analysis (ICA) were designed to resolve the instantaneous mixtures by optimizing a contrast function based on the measure of independence or non-Gaussianity. In previous BSS methods, the frequency characteristics and location of individual sources and how these sources were mixed were not sophisticatedly investigated. Solving the instantaneous mixtures did not truly reflect the real reverberant environment which structurally mixed the sources as the convolutive mixtures [11]. The underdetermined problem in presence of more sources than sensors may not have been carefully treated [14]. The contrast functions may not flexibly and honestly measure the independence for an optimization with convergence [3]. The static mixing system could not catch the underlying dynamics in source signals and sensor networks. The uncertainty of system parameters may not be precisely characterized so that the robustness against adverse conditions was not guaranteed [5].

Generally, signal processing and machine learning provide fundamental knowledge and algorithm to resolve different issues in audio source separation. The goal of this article is to overview a series of recent advances in adaptive processing and learning algorithms for BSS in presence of speech and music signals. We survey the recent solutions to overdetermined/underdetermined convolutive separation [12], sparse source separation [1], nonnegative matrix factorization (NMF) [4][13], information-theoretic learning [2][5], online learning [5] and Bayesian inference.

In general, these algorithms are classified into front-end processing and back-end learning as shown in Figure 1 [6]. In front-end processing, we highlight on the adaptive signal processing to analyse the information on each source, such as its frequency characteristics and location, or identifying how the sources are mixed. We review the works on frequency-domain audio source separation which could align the permutation ambiguities [12], separate the convolutive mixtures, identify the number of sources [1], resolve the overdetermined/underdetermined problem [14]. The back-end learning is devoted to recover the source signals by using only the information about their mixtures observed in each microphone without frequency and location information on each source. We build a statistical model and infer the model by using the mixtures. We introduce the estimation of demixing parameters through construction and optimization of information-theoretic contrast function [2][3]. The solutions to music source separation based on NMF [13] and sparse learning [4] are addressed. Next, we focus on the uncertainty modeling for the regularized signal separation in accordance with Bayesian perspective. The nonstationary and temporally-correlated source separation [5] is presented.

II. Front-End Processing

Considering the issue of unknown number of sources, a Gaussian mixture model with Dirichlet prior for mixture weight parameter was proposed to identify the direction-of-arrival (DOA) of source speech signal from individual time-frequency units. This model was applied to estimate the number of sources and deal with the sparse source separation [1].

For the determined or the overdetermined problem, the number of microphones is enough for the number of sources. The complex-valued ICA was proposed to separate the frequency bin-wise mixtures. For each frequency bin, the ICA demixing matrix is optimized so that the distribution of the demixed elements is far from a Gaussian [11].

There is scaling ambiguity in an ICA solution. For an audio source separation task, the scaling ambiguity is resolved by representing the observed signals at microphones with the scaled separated signals [11].

In an underdetermined system, the number of microphones N is insufficient for the number of sources M, we typically employ the method based on time-frequency masking, where we need to estimate which source has the largest amplitude for

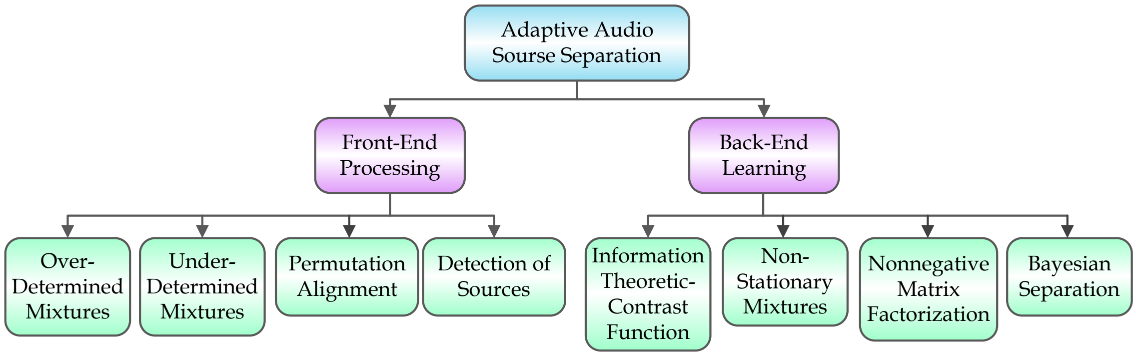


Fig. 1. Some issues in adaptive audio source separation.

each time frequency slot (f, t). For this purpose, we apply a clustering method to M-dimensional observation vectors $\mathbf{x}_{ft}$ and to calculate the posterior probability $p(C_m|\mathbf{x}_{ft})$ that a vector $\mathbf{x}_{ft}$ belongs to a cluster or a source C_m. Then, the time frequency masks $\mathcal{M}_{ftm}$ are made and used to find the separated signals $\hat{\mathbf{s}}_{ft}^{(m)} = \mathcal{M}_{ftm}\mathbf{x}_{ft}$. The posterior probability $p(C_m|\mathbf{x}_{ft})$ is calculated by using a likelihood function $p(\mathbf{x}_{ft}|\boldsymbol{\Theta})$ based on a Gaussian mixture model (GMM) with parameters $\boldsymbol{\Theta}$ [12].

The method based on ICA or GMM performs a source separation task in a frequency bin-wise manner. Therefore, we need to align the permutation ambiguity of the ICA or GMM results in each frequency bin so that a separated signal in the time domain contains frequency components from the same source signal. This problem is well known as the permutation problem of frequency-domain BSS [9]. The *dominance measures* [10][12] performs very well for this problem. When using ICA, we employ the power ratio of the scaled separated signals as a dominance measure $r_f^{(m)}(t)$ [10]. On the other hand, when using a GMM for time-frequency masking, we employ the posterior probability $r_f^{(m)}(t) = p(C_m|\mathbf{x}_{ft})$ as a dominance measure [12]. After calculating the dominance measure, we basically interchange the indices m of the separated signals so that the correlation coefficient $\rho(r_f^{(m)}, r_{f'}^{(m)})$ between the dominance measures at different frequency bins f and f' is maximized for the same source.

III. Back-End Learning

In this section, we focus on the *machine learning* solutions to audio source separation. We consider blind speech or music separation as a learning problem without special treatment on convolutive mixtures or extraction of frequency features and location information on each source signal. Let the observation vector $\mathbf{x}_t = [x_{t1}, \ldots, x_{tN}]^{\mathrm{T}}$ from N microphones at time frame t be mixed by $\mathbf{x}_t = \mathbf{A}\mathbf{s}_t$ where $\mathbf{A}$ is an unknown $N \times M$ mixing matrix and $\mathbf{s}_t = [s_{t1}, \ldots, s_{tM}]^{\mathrm{T}}$ denotes a vector of M mutually-independent source signals. For the case of $N = M$, BSS problem is resolved by ICA method which optimizes a contrast function $\mathcal{D}(\mathbf{X}, \mathbf{W})$ measuring the independence or non-Gaussianity of the demixed signals $\hat{\mathbf{s}}_t$ based on a demixed matrix or separation matrix $\mathbf{W}$, i.e. $\hat{\mathbf{s}}_t = \mathbf{W}\mathbf{x}_t$. The demixing matrix can be estimated in accordance with the gradient descent algorithm or the natural gradient algorithm from a set of audio signals $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_T\}$. The metrics of likelihood function, negentropy and kurtosis are popular to serve as ICA contrast functions. More meaningfully, the information-theoretical contrast function is adopted to measure the independence between the demixed signals.

The *statistical hypothesis test* was recently proposed to carry out an information measure of confidence towards independence by investigating the null hypothesis $\mathcal{H}_0$ where the demixed signals $\hat{\mathbf{S}} = \{\hat{\mathbf{s}}_1, \ldots, \hat{\mathbf{s}}_T\}$ are independent against the alternative hypothesis $\mathcal{H}_1$ where the demixed signals are dependent [2]. The contrast function was formed as a log likelihood ratio given by $\mathcal{D}(\mathbf{X}, \mathbf{W}) = \log p(\hat{\mathbf{S}}|\mathcal{H}_0) - \log p(\hat{\mathbf{S}}|\mathcal{H}_1)$. More generally, the measure of independence is calculated as a divergence between the joint distribution of the demixed signals and the product of marginal distributions of individual demixed signals. This divergence measure equals to zero in case that the condition of independence is met. A general convex divergence measure was derived by substituting a general convex function $f(t) = \frac{4}{1-\alpha^2}\left[\frac{1-\alpha}{2} + \frac{1+\alpha}{2}t - t^{(1+\alpha)/2}\right]$ into Jensen's inequality to construct a contrast function for ICA optimization. This convex divergence $\mathcal{D}(\mathbf{X}, \mathbf{W}, \alpha)$ is developed with an adjustable convexity parameter. In cases of $\alpha = 1$ and $\alpha = -1$, the general convex divergence is realized to the *convex-Shannon* divergence and the *convex-logarithm* divergence where the convex functions based on Shannon's entropy and negative logarithm are adopted, respectively.

The dictionary learning based on the nonnegative matrix factorization (NMF) is recently hot issue in audio source separation [7]. NMF attempts to decompose the nonnegative mixed samples $\mathbf{X} \in \mathbb{R}^{N \times T}$ into a product of nonnegative mixing matrix $\mathbf{A} \in \mathbb{R}^{N \times M}$ and nonnegative source signals $\mathbf{S} \in \mathbb{R}^{M \times T}$ by minimizing a divergence measure $\mathcal{D}(\mathbf{X}, \mathbf{A}, \mathbf{S})$ between $\mathbf{X}$ and $\mathbf{AS}$. NMF is a parts-based representation which only allows additive combination and can be directly applied to decompose the nonnegative mixed audio signals. The absolute values of short-time Fourier transform (STFT) are calculated to form $\mathbf{X}$. The standard NMF is fulfilled according to a regularized least square criterion with sparsity constraint.

More recently, the convex divergence [3] and Itakura-Saito (IS) divergence [13] were treated as the objective function to derive solution to NMF. For example, IS divergence is written by $\mathcal{D}_{\text{IS}}(\mathbf{X}, \mathbf{A}, \mathbf{S}) = \sum_{n,t}(\frac{X_{nt}}{[\mathbf{AS}]_{nt}} - \log\frac{X_{nt}}{[\mathbf{AS}]_{nt}} - 1)$ which depends only on the ratio $\frac{X_{nt}}{[\mathbf{AS}]_{nt}}$. In [8][13], minimizing IS divergence was shown to be equivalent to maximizing the log-likelihood $\log p(\tilde{\mathbf{X}}|\mathbf{A}, \mathbf{S})$ based on the multivariate complex-valued Gaussian distributions where $\tilde{\mathbf{X}}$ denotes a matrix of STFT complex-valued coefficients.

In [4], Bayesian NMF was proposed for monaural music source separation which decomposed a single-channel mixed signal $\mathbf{X}$ into a rhythmic signal $\mathbf{X}_r$ and a harmonic signal $\mathbf{X}_h$. Let the nonnegative monaural matrix $\mathbf{X} \in \mathbb{R}^{F\times T}$ in time-frequency domain be chunked into L segments $\{\mathbf{X}^{(l)}\}$. Each segment is represented by $\mathbf{X}^{(l)} = \mathbf{X}_r^{(l)} + \mathbf{X}_h^{(l)} = \mathbf{A}_r\mathbf{S}_r^{(l)} + \mathbf{A}_h^{(l)}\mathbf{S}_h^{(l)}$ where $\{\mathbf{S}_r^{(l)}, \mathbf{S}_h^{(l)}\}$ are two groups of segment-dependent encoding coefficients, $\mathbf{A}_h^{(l)}$ denotes the bases for harmonic source which are individual for different segments l, and $\mathbf{A}_r$ denotes the bases for rhythmic source which are shared across segments. Assuming the basis components are Gamma distributed and the encoding coefficients are Laplacian distributed, *Bayesian group sparse learning* for NMF was performed to resolve the underdetermined source separation through a Gibbs sampling procedure.

Further, we face the challenges of changing sources or moving speakers, namely the source signals may abruptly appear or disappear, the speakers may be replaced by new ones or even moving from one location to the other. The mixing conditions and source signals are accordingly nonstationary and should be traced to assure robustness in nonstationary source separation [5]. A meaningful approach to deal with the robustness issue in audio source separation is constructed from *Bayesian perspective*. Some prior information is introduced for uncertainty modeling and knowledge integration. Let $\mathbf{X}^{(l)} = \{\mathbf{x}_t^{(l)}\}$ denote a set of mixed signals at segment l. The signals are mixed by a linear combination of M unknown source signals $\mathbf{S}^{(l)} = \{\mathbf{s}_t^{(l)}\}$ using a mixing matrix $\mathbf{A}^{(l)}$, i.e. considering a noisy ICA model $\mathbf{x}_t^{(l)} = \mathbf{A}^{(l)}\mathbf{s}_t^{(l)} + \varepsilon_t^{(l)}$ where $\mathbf{E}^{(l)} = \{\varepsilon_t^{(l)}\}$ denotes the noise signals. We assume that $\mathbf{A}^{(l)}$ and $\mathbf{S}^{(l)}$ are unchanged within a segment l but varied across segments. To tackle the nonstationary source separation, we attempt to incrementally characterize the variations of $\mathbf{A}^{(l)}$ and $\mathbf{S}^{(l)}$ from the observed segments $\mathcal{X}^{(l)} = \{\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \ldots, \mathbf{X}^{(l)}\}$. *Online learning* is conducted to compensate for nonstationary conditions of mixing coefficients and source signals segment by segment. We also present the solution to nonstationary and *temporally correlated* source separation where the mixing condition is changed continuously and the temporal correlation in time-series signals, e.g. mixing coefficients and source signals, is taken into account. Online learning and *Gaussian process* are merged into a separation system which compensates for the nonstationary and temporally correlated mixing environments and source signals, respectively.

IV. Conclusions

We have presented a series of adaptive methods which were developed for different issues in BSS. In front-end processing, we addressed high-performance solutions to overdetermined and underdetermined problems which are based on the processing of complex-valued time-frequency signals and the noise-masking method using Gaussian mixture model. The permutation problem was solved according to the correlation coefficient between dominance measures at different frequency bins. In back-end learning, we addressed the importance of information-theoretical learning for ICA optimization. The recent methods of sparse learning and dictionary learning based on NMF were presented for speech/music source separation. The online learning and Bayesian learning designed for nonstationary source separation were also presented for improving the robustness for audio source separation.

References

[1] S. Araki, T. Nakatani, H. Sawada, and S. Makino, "Blind sparse source separation for unknown number of sources using Gaussian mixture model fitting with Dirichlet prior," in *Proc. of ICASSP*, 2009, pp. 33–36.

[2] J.-T. Chien and B.-C. Chen, "A new independent component analysis for speech recognition and separation," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 14, no. 4, pp. 1245–1254, 2006.

[3] J.-T. Chien and H.-L. Hsieh, "Convex divergence ICA for blind source separation," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 20, no. 1, pp. 290–301, 2012.

[4] J.-T. Chien and H.-L. Hsieh, "Bayesian group sparse learning for music source separation," *EURASIP Journal on Audio, Speech, and Music Processing*, 5 July, 2013. (doi: 10.1186/1687-4722-2013-18)

[5] J.-T. Chien and H.-L. Hsieh, "Nonstationary source separation using sequential and variational Bayesian learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 5, pp. 681–694, 2013.

[6] J.-T. Chien, H. Sawada and S. Makino, "Adaptive processing and learning for audio source separation," in *Proc. of APSIPA ASC*, 2013.

[7] A. Cichocki, R. Zdunek, and S. Amari, "New algorithms for non-negative matrix factorization in applications to blind source separation", in *Proc. of ICASSP*, 2006, pp. 621-624.

[8] C. Fevotte, N. Bertin, and J.-L. Durrieu, "Nonnegative matrix factorization with the Itakura-Saito divergence: with application to music analysis," *Neural Computation*, vol. 21, no. 3, pp. 793–830, 2009.

[9] H. Sawada, R. Mukai, S. Araki, and S. Makino, "A robust and precise method for solving the permutation problem of frequency-domain blind source separation," *IEEE Transactions on Speech and Audio Processing*, vol. 12, pp. 530–538, 2004.

[10] H. Sawada, S. Araki, and S. Makino, "Measuring dependence of bin-wise separated signals for permutation alignment in frequency-domain BSS," in *Proc. of ISCAS*, pp. 3247–3250, 2007.

[11] H. Sawada, S. Araki, and S. Makino, "Frequency-domain blind source separation," in *Blind Speech Separation*, S. Makino, T.-W. Lee, and H. Sawada, Eds. Springer, 2007, pp. 47–78.

[12] H. Sawada, S. Araki, and S. Makino, "Underdetermined convolutive blind source separation via frequency bin-wise clustering and permutation alignment," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 19, no. 3, pp. 516–527, 2011.

[13] H. Sawada, H. Kameoka, S. Araki, and N. Ueda, "Multichannel extensions of non-negative matrix factorization with complex-valued data," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 21, no. 5, pp. 971–982, 2013.

[14] S. Winter, W. Kellermann, H. Sawada, and S. Makino, "MAP-based underdetermined blind source separation of convolutive mixtures by hierarchical clustering and ℓ_1-norm minimization," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, pp. 1–12, 2007.

[15] T. Yoshioka, A. Sehr, M. Delcroix, K. Kinoshita, R. Maas, T. Nakatani and W. Kellermann, "Making machines understand us in reverberant rooms - robustness against reverberation for automatic speech recognition," *IEEE Signal Processing Magazine*, vol. 29, pp. 114-126, 2012.

Congratulations on Prof. Li Haizhou and team winning President's Technology Award 2013, Singapore

The **President's Technology Award (PTA)** gives recognition to research scientists and engineers in Singapore who have made outstanding contributions to research & development resulting in significant new technology or innovative use of established technology. On 25 September 2013, Prof. Li Haizhou received the PTA award from His Excellency Dr Tony Tan Keng Yam, the President of the Republic of Singapore, ***"For their outstanding contributions to human language technology that have empowered the industry and benefited the Asian society".*** Prof. Li's team was the only PTA winner in Singapore in 2013.

Prof. Li Haizhou received the PTA Award from President Tony Tan on 25 September 2013.

Professor Li Haizhou, an internationally-renowned scientist, and his team Dr Ma Bin, Ms Aw Ai Ti, and Dr Su Jian have made a remarkable breakthrough in human language technology that transforms the interface of mobile applications and breaks down the language barriers for Asian society.

Professor Li Haizhou (2nd from the left) and his team (Ms Aw Ai Ti, Dr Ma Bin, Dr Su Jian)

Among the 7,105 living spoken languages that *Ethnologue* documented in 2013, 2,304 are spoken by Asians, representing more than half of the world's population. However, traditional human language technologies were developed using English and other major languages as the workbench, which cannot be applied to many Asian languages. Over the past nine years, Professor Li and his team pioneered new approaches to speaker recognition, multilingual speech recognition, tonal language processing, as well as lexical, syntactic, semantic and discourse analysis. These novel inventions now serve as the foundation of the Abacus language engine, a commercial grade technology solution for the Bahasa Indonesian, English, Malay, Mandarin Chinese, Thai, and Vietnamese languages. Abacus accurately converts continuous speech into text, identifies the accents, dialects, and languages being spoken, establishes a speaker's identity by his/her voice, and translates languages between one another.

The technological breakthrough is significant. The Abacus engine achieved a leading performance in US National Institute of Standards and Technology (NIST) international benchmarking competitions, including NIST Language Recognition Evaluation 2007, NIST Speaker Recognition Evaluation 2008 and 2012, and NIST Text Analysis Conference 2011, representing the state-of-the-art in academia and industry. In developing the Abacus engine, the team addressed the unique research problems that Asian languages face, such as multilingual speech and tonal language processing, and translation between Asian languages. The team also formulated a novel industry process for rapid technology deployment that has been adopted widely by the industry.

The team's recent achievements have put Singapore on the world map. In 2009, Professor Li was elected as a Board Member of the International Speech Communication Association (ISCA) and named one of the two Nokia Visiting Professors by the Nokia Foundation. In 2012, Dr Su Jian was elected as an Executive Committee Member of the Association for Computational Linguistics (ACL). Their work has also been published as an 'Invited Paper' in the Proceedings of the IEEE in 2013, the most highly-cited general interest journal in electrical engineering and computer science, and honored as 'The Most Cited Article' in Elsevier Speech Communication during 2008-2013. One major outcome of the team's research is the establishment of the Baidu-I^{2}R Research Centre (BIRC) in Singapore. In 2012, the internet giant Baidu and I^{2}R set up BIRC as Baidu's first overseas joint laboratory to further the research of speech information processing and Asian language processing. The establishment of BIRC is an endorsement of the team's technological achievements. The Abacus language engine has become one of the most sought after solutions internationally for text input, question and answering, spoken dialogue, and voice biometrics in mobile applications. For the past three years, the Abacus engine has been licensed to more than 15 leading international companies to enable many innovative products. In particular, the Abacus engine was adopted in 2012 to power the Lenovo A586, the world's first voiceprint smartphone. The team also contributed to ITU-T F.745 and H.625 international standards for network-based speech to speech translation in 2010.

For their outstanding contributions to human language technology that have empowered the industry and benefited the Asian society, Professor Li Haizhou, Dr Ma Bin, Ms Aw Ai Ti and Dr Su Jian from the Institute for Infocomm Research are awarded the 2013 President's Technology Award.

APSIPA Advisory Board (AAB)

APSIPA has recently elected 10 prominent figures to act as members of the Advisory Board as follows:

Antonio Ortega, *University of Southern California, USA* (2015)

Eliathamby Ambikairajah, *University of New South Wales, Australia* (2015)

Hideaki Sakai, *Kyoto University, Japan* (2015)

Hitoshi Kiya, *Tokyo Metropolitan University, Japan* (2015)

Jiwu HUANG, *Sun Yat-Sen University, China* (2015)

K. J. Ray Liu, *University of Maryland, USA* (2015)

Li Deng, *Microsoft Reserch, USA* (2015)

Lin-Shan Lee, *National Taiwan University, Taiwan* (2015)

Mark Liao, *Institute for Information Science, Taiwan* (2015)

Yoshikazu Miyanaga, *Hokkaido University, Japan* (2015)

The Charter for APSIPA Advisory Board (AAB) can be found on: http://apsipa.org/members.htm

APSIPA Transactions on Signal and Information Processing

APSIPA TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING, VOLUME 2
Rate-dependent seam carving and its application to content-aware image coding
Yuichi Tanaka and Taichi Yoshida and Madoka Hasegawa and Shigeo Kato and Masaaki Ikehara

APSIPA Transactions on Signal and Information Processing, Volume 2, January 2013, e1
doi: 10.1017/ATSIP.2013.3, Published online by Cambridge University Press 27 Feb 2013
Behavior signal processing for vehicle applications
Chiyomi Miyajima and Pongtep Angkititrakul and Kazuya Takeda

APSIPA Transactions on Signal and Information Processing, Volume 2, January 2013, e2
doi: 10.1017/ATSIP.2013.2, Published online by Cambridge University Press 04 Mar 2013
Constant frame quality control for H.264/AVC
Ching-Yu Wu and Po-Chyi Su and Long-Wang Huang and Chia-Yang Chiou

APSIPA Transactions on Signal and Information Processing, Volume 2, January 2013, e3
doi: 10.1017/ATSIP.2013.4, Published online by Cambridge University Press 01 May 2013
Visual quality assessment: recent developments, coding applications and future trends
Tsung-Jung Liu and Yu-Chieh Lin and Weisi Lin and C.-C. Jay Kuo

APSIPA Transactions on Signal and Information Processing, Volume 2, January 2013, e4
doi: 10.1017/ATSIP.2013.5, Published online by Cambridge University Press 11 Jul 2013
Reversible color transform for Bayer color filter array images
Suvit Poomrittigul and Masanori Ogawa and Masahiro Iwahashi and Hitoshi Kiya

APSIPA Transactions on Signal and Information Processing, Volume 2, January 2013, e5
doi: 10.1017/ATSIP.2013.6, Published online by Cambridge University Press 27 Sep 2013
Dark and low-contrast image enhancement using dynamic stochastic resonance in discrete cosine transform domain
Rajib Kumar Jha and Rajlaxmi Chouhan and Kiyoharu Aizawa and Prabir Kumar Biswas

APSIPA Transactions on Signal and Information Processing, Volume 2, January 2013, e6
doi: 10.1017/ATSIP.2013.7, Published online by Cambridge University Press 12 Nov 2013

Are you an APSIPA member?
If not, then register online at
http://www.apsipa.org

Asia-Pacific Signal and Information Processing Association

APSIPA

APSIPA Distinguished Lecture

Speaker Verification –

The Present and Future of Voiceprint Based Security

presented by

Professor Eliathamby Ambikairajah

Head of School of Electrical Engineering & Telecommunications

at University of New South Wales (UNSW), Sydney, Australia

Summary: Speaker verification refers to a system that analyses and understands an individual's voice, but more specifically their voice print, which can be used for security. Specifically, the use of voice prints to verify if the speech utterance belongs to the claimed speaker. This talk will provide an overview of how current text independent speaker verification systems are implemented as well as pointing out some emerging trends for the future.

Date: Monday, 21st October 2013

Time: 11:00am

Venue: Main Conference Room, Faculty of Engineering, International Islamic University Malaysia (IIUM)

Host: Faculty of Engineering at IIUM, WCSP Research Group, and IEEE Intstrumentation and Measurement Society (IMS)

Albert Einstein Quotes

1. Insanity: doing the same thing over and over again and expecting different results.
2. Learn from yesterday, live for today, hope for tomorrow. The important thing is not to stop questioning.
3. We cannot solve our problems with the same thinking we used when we created them.
4. Try not to become a man of success, but rather try to become a man of value.
5. Anyone who doesn't take truth seriously in small matters cannot be trusted in large ones either.
6. Only a life lived for others is a life worthwhile.
7. If you can't explain it simply, you don't understand it well enough.
8. It's not that I'm so smart, it's just that I stay with problems longer.
9. Logic will get you from A to B. Imagination will take you everywhere.
10. Science without religion is lame, religion without science is blind.
11. It has become appallingly obvious that our technology has exceeded our humanity.
12. ***I know not with what weapons World War III will be fought, but World War IV will be fought with sticks and stones.***
13. The value of a man should be seen in what he gives and not in what he is able to receive.
14. Everyone should be respected as an individual, but no one idolized.
15. Never do anything against conscience even if the state demands it.
16. I think and think for months and years. Ninety-nine times, the conclusion is false. The hundredth time I am right.

Selected by Waleed H. Abdulla

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, I²R, 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

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