Session 1aID

Interdisciplinary: Opening Ceremonies, Plenary Lectures

Whitlow W. L. Au, Cochair
Hawaii Inst. of Marine Biology, P.O. Box 1106, Kailua, HI 96734

Sadaoki Furui, Cochair
Tokyo Inst. of Technology, Dept. of Computer Science, 2-12-2 O-akayama Meguro-ku, Tokyo 152-8552, Japan

Opening Ceremony—8:00

Chair's Introduction—8:10

Invited Paper

8:15

1aID1. How can auditory presence be generated and controlled? Masayuki Morimoto (Environ. Acoust. Lab., Faculty of Eng., Kobe Univ., Nada, Kobe 657-8501 Japan)

This paper reviews key results of many listening tests on auditory presence, especially auditory localization (AL) and auditory spatial impression (ASI) by the author. The author gave the first demonstration that AL in any direction can be simulated through two loudspeakers using head-related transfer functions (HRTFs). However, individual differences in HRTFs affect the accuracy of AL. It is basically possible to localize sound images in any direction using median-plane HRTFs combined with interaural differences. Furthermore, HRTFs can be simplified by the combination of only some spectral peaks and notches. Meanwhile, the author demonstrated that ASI comprises at least two perceptual components. One is auditory source width (ASW), defined as the width of a sound image fused temporally and spatially with the direct sound image. The other is listener envelopment (LEV), which is the degree of fullness of sound images around the listener, excluding the sound image composing ASW. A listener can perceive these two components separately. The perception of ASW and LEV has a close connection with the law of the first wavefront. Acoustic components under the upper limit of the law contribute to ASW, and acoustic components beyond the upper limit of the law contribute to LEV.

Chair's Introduction—9:05

Invited Paper

9:10

1aID2. Therapeutic ultrasound. Lawrence Crum (Ctr. for Industrial and Med. Ultrasound, Appl. Phys. Lab., 1013 NE 40th St., Seattle, WA 98105, lac@apl.washington.edu)

The use of ultrasound in medicine is now quite commonplace, especially with the recent introduction of small, portable, and relatively inexpensive, hand-held diagnostic imaging devices. Moreover, ultrasound has expanded beyond the imaging realm, with methods and applications extending to novel therapeutic and surgical uses. Among these applications are tissue ablation, acoustocautery, lipoplasty, site-specific and ultrasound mediated drug activity, extracorporeal lithotripsy, and the enhancement of natural physiological functions such as wound healing and tissue regeneration. A particularly attractive aspect of this technology is that diagnostic and therapeutic systems can be combined to produce totally noninvasive, image-guided, bloodless surgery. This general lecture will review a number of these exciting new applications of ultrasound and address some of the basic scientific questions and future challenges in developing these methods and technologies for general use in our society. We shall particularly emphasize the use of high-intensity focused ultrasound (HIFU) in the treatment of benign and malignant tumors as well as the induction of acoustic hemostasis, especially in organs that are difficult to treat using conventional medical and surgical techniques. [Work supported in part by the NIH, NSBRI, ONR, and DARPA.]

pression of the brake pedal was observed. The proposed lexicographical speech-feature-based method also detected 33 false alarms to detect 80% of these 11 scenes. As for the other 17 scenes, our method based on two-dimensional histograms of brake pressure and its dynamics achieved an 80% detection rate for 473 false alarms. Analyses of data recorded while drivers interacted with a machine and a Wizard of Oz system as well as a rank of the most commonly uttered words in dangerous situations are also presented.

1pSC47. Computational complexity of a distance-based active search algorithm. Masahide Sugiyama (The Univ. of AIZU, Tsuruga, Ikki-machi, Aizu-Wakamatsu-shi, Fukushima, 965-8580 Japan)

An efficient pattern (segment) search in a large-scale database is a very important technology in current widespreaded Internet society. To achieve efficient searches, many efficient pattern search algorithms have been proposed. Active search (AS) is one of them; it is based on skipping of search calculation. Initially, AS was formulated between two histograms using a similarity measure. In fact, AS can be formulated using L_*p* distance, where *p* is greater than 1. The search efficiency can be evaluated using a number of distance calculations. The first result in this paper shows that the L_1 distance in AS gives the minimum number of distance calculations among L_*p* distances. This property can be derived from a newly derived inequality between the L_*p* and L_1 norm on output probability space. A new distance can be derived from positive weighting of any distance. The second result of this paper is that the average skip width for any weighed L_*p* distance cannot be greater than the most efficient distance. Therefore, no positive combination of L_*p* can be more efficient than L_1. The two results indicate that L_1 distance is the most efficient for active searching.

TUESDAY AFTERNOON, 28 NOVEMBER 2006 HONOLULU ROOM, 1:00 TO 4:50 P.M.

Session 1pSP

Signal Processing in Acoustics: Blind Signal Processing

Leon H. Sibul, Cochair
Pennsylvania State Univ., Applied Research Lab., P.O. Box 30, State College, PA 16804

Shoji Makino, Cochair
NTT Communication Science Lab., 2-4 Hikaridau, Seika-cho, Sorakugun, Kyoto, 619-0237, Japan

Chair's Introduction—1:00

Invited Papers

1:05

1pSP1. Frequency domain blind source separation in a noisy environment. Ryo Mukai, Hiroshi Sawada, Shoko Araki, and Shoji Makino (NTT Commun. Sci. Labs., NTT Corp., 2-4 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0237, Japan)

A prototype system for blind source separation (BSS) in a reverberant and noisy environment is presented. Our system uses a small three-dimensional array with eight microphones and realizes BSS in various situations such as the separation of many speech signals located in a three-dimensional space, and the extraction of primary sound sources surrounded by many background interferences. The mixed signals observed by the microphone array are processed by independent component analysis (ICA) in the frequency domain. The permutation problem is solved by normalized basis vector clustering, which is a generalized version of directions of arrival (DOA) clustering and has several advantages to the DOA based method. The system estimates the DOA of the source signals as a by-product of the separation process. Moreover, the system has the ability to distinguish primary target signals and ambient noise. Our system performs an on-the-spot BSS of live recorded signals. Live demonstration will be performed at the meeting.

1:25

1pSP2. Beyond the independent component analysis: New blind acoustic sound separation in real world via single-input multiple-output-based independent component analysis. Hiroshi Saruwatari (Nara Inst. of Sci. and Technol., 8916-5 Takayama-cho, Ikoma, Nara, 630-0192 Japan)

This review discusses a new generation of blind acoustic-signal separation in the real world, focusing on a recently proposed ICA method, Single-Input Multiple-Output (SIMO)-signal-based ICA (SIMO-ICA). The term ''SIMO'' is referred to as a specific transmission system in which the input is a single source signal and the outputs are its transmitted signals observed at multiple microphones. The SIMO-ICA consists of multiple ICA parts and a fidelity controller, and each ICA runs in parallel under the fidelity control of the entire separation system. In the SIMO-ICA scenario, mixed multiple sources detected at the microphones can be separated, not into monaural source signals but into SIMO-model-based contributions from independent sources as they are at the microphones. Thus, the separated signals of the SIMO-ICA can maintain the spatial qualities of each sound source. This attractive feature of the SIMO-ICA shows the promise of applicability and connectability with the other multichannel signal processing for achieving the higher performance beyond the conventional ICA methods. As the successful examples of SIMO-ICA's application, three kinds of systems are described: (*i*) combination with inverse filtering for MIMO decomvolution, *ii*) combination with adaptive beamformer for higher-performance BSS, and *iii* combination with binary masking for real-time processing.

1:45

1pSP3. How blind are we in blind signal processing? Walter Kellermann (Univ. Erlangen-Nuremberg, Cauer-str. 7, 91058 Erlangen, Germany, wk@lnt.de)

Blind signal processing algorithms aim at blindly solving classical signal processing problems such as system identification, signal separation, or source localization. Here it is investigated to which extent some of the more popular blind signal processing concepts are really blind when they are used for signal acquisition in the acoustic domain. Blindness regarding the source configuration and the source signal properties, blindness with respect to channel properties, and, finally, blindness regarding the microphone configuration are investigated. It turns out that many algorithms are explicitly nonblind with respect to important properties of the source configuration, such as the number of sources and their pointlike nature. The source signals typically have to fulfill certain statistical properties or have to meet sparseness constraints. Some algorithms rely on a dominant direct acoustic path between sources and sensors, and linearity of the transmission channel model is implied with all convolutive mixture models. Finally, some popular algorithms are actually aiming at direction of arrival estimation, which always requires knowledge on the microphone array geometry. In essence, blind algorithms are wellinformed in many respects and are blind only with regard to a few, although decisive, properties.

2:05

1pSP4. Informed acoustic source separation and localization. Kevin H. Knuth (Dept. of Phys., Univ. at Albany, Albany, NY 12222)

Advances in Bayesian computational technology in the last decade have enabled the development of new source separation and source localization algorithms. These algorithms are greatly improved by the encoding of prior information about a specific problem in the form of the chosen relevant model parameters, the assignment of the likelihood functions, and the assignment of the prior probabilities of the model parameter values. I refer to such source separation algorithms as informed source separation for the reason that they are endowed with specific and often vital information about the problem. Furthermore, the Bayesian methodology allows source separation to be united with source localization simply by including the model parameters that are of interest to the researcher. Here, I will discuss the union of source separation and source localization under the Bayesian methodology, the incorporation of prior information, and the construction of an informed algorithm using the new computational technologies that allow us to estimate the values of the parameters that define these high-dimensional problems.

2:25

1pSP5. Recovering the quality of speech degraded by reverberations in a room. Masato Miyoshi (NTT Commun. Sci. Labs, 2-4 Hikaridai Seika-cho, Keihanna Sci. City, Kyoto 6190237 Japan, miyo@cslab.kecl.ntt.co.jp)

Speech signals captured by distant microphones in a room are usually reverberated due to wall reflections. Reverberation may seriously deteriorate the signal characteristics, thus damaging the quality of such applications as hands-free telephony or automatic speech recognition. To eliminate such room reverberation effects, the inverse-filtering of the room transfer functions (RTFs) between a speaker and the microphones appears to be a very promising approach. When applied to a microphone system, this processing should work blindly because no explicit information on RTFs or source signals is available. The blind design of an inverse-filter set may be roughly classified into three approaches: (1) calculating an inverse-filter set of RTF estimates obtained from observed reverberant signals based on the ‘‘subspace method,’’ (2) calculating the filter set from observed signals and replicas of direct-sound signals using statistical properties inherent in speech signals, and (3) calculating the filter set directly from received signals, by whitening these signals without excessive degradation of the original signal characteristics. This presentation will provide an overview of these three approaches and introduce our speech dereverberation trials.

2:45

1pSP6. Blind deconvolution for noisy dynamic channels. Michael J. Roan (Dept. of Mech. Eng., Virginia Tech., Blacksburg, VA 24060)

It is common in acoustics to measure a signal that has been degraded by propagation through an unknown, noisy channel prior to measurement. While only the degraded measured signal is available, the data of interest are the original signal and the channel parameters. Often, it is desirable to reverse the filtering process by application of an inverse filter to recover the original signal. When neither the input signal properties nor the channel properties are known, this process is known as blind deconvolution (BDC). Typically, BDC algorithms assume noiseless, stationary propagation channels and input sources. These assumptions are usually violated in practical applications (e.g., noisy multipath propagation environments with moving source and receiver). To model these effects, predictive techniques are applied to incorporate *a priori* information about the system into the existing blind processing framework. The original contributions of this work follow. First, a novel formulation of the extended Kalman filter (EKF) is proposed. This allows incorporation of *a priori* information into gradient-based blind processing algorithms. This formulation is then applied to the natural gradient (NG) BDC algorithm. Finally, results are presented that suggest significant improvement in signal recovery performance over the NG BDC algorithm for dynamic noisy channels.

Contributed Papers

3:20

1pSP7. Blind dereverberation based on auto-correlation functions of framewise time sequences of frequency components. Kenko Ota and Masuzo Yanagida (Doshisha Univ., 1-3, Tatara-Miyakodani, Kyotanabe, Kyoto, 610-0321, Japan, etf1704@mail4.doshisha.ac.jp)

Proposed is a new blind dereverberation method based on auto-correlation function of framewise time series of each frequency component. Inverse filtering of source-microphone transfer functions is widely employed for suppressing the effects of reflected waves, but this method cannot be employed for cases where source-microphone transfer functions are not available or for time-variant cases. Several methods of spectral subtraction have been proposed to cope with these cases. However, most of them require transfer functions among sources and microphones. The proposed method, however, does not require transfer function inverse filters. Moreover, the method can handle frequency characteristics of sound paths. To realize dereverberation, it is necessary to estimate the delay time and frequency characteristics of reflection, though most conventional methods assume flat frequency characteristics. The proposed method estimates these parameters based on auto-correlation function of each time series of spectral components in frequency spectra obtained every frameshift interval. The proposed method improves dereverberation performance, comparing with methods that assume flat frequency characteristics. The proposed method raises average segmental SNR by 3.4 dB and reduces reverberation time from 390 to 109 ms for a room employed for experiments. [Work supported by Knowledge Cluster Project, MEXT, and by Academic Frontier Project Doshisha University.]

3:35

1pSP8. Sparseness-based 2ch blind source separation in reverberant environments based on direction-of-arrival estimation with a reliability measure. Yosuke Izumi, Nobutaka Ono, and Shigeki Sagayama (Dept. of Information Sci. and Technol., The Univ. of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan, izumi@hil.t.u-tokyo.ac.jp)

We present a novel sparseness-based 2ch blind source separation (BSS) method based on robust direction-of-arrival (DOA) estimation that enables reasonable separation of source signals even in a highly reverberant environment. DOA estimation for BSS is not easy in reverberant environments since reflected sound waves behave as coherent interferences in the observed signals and make ambiguous the sound source direction. Our key idea to solve this problem is to find and integrate clean, reliable time-frequency fragments through the observation with a reliability measure. For that, (1) based on a diffused sound field model of reverberation, a theoretical correlation matrix is introduced for whitening the coherent interferences, (2) the DOA of each time-frequency bin is estimated with reliability by utilizing information of neighboring bins, and (3) the reliability-weighted DOAs are integrated over the whole time-frequency bins. After that, separated signals are obtained by time-frequency masking or inverse filtering. Experimental results including comparison with other methods are also reported.

3:50

1pSP9. Multilayered spatio-temporal gradient analysis for acoustic blind source separation. Kenbu Teramoto, Md. Tawhidul Islam Khan, Seiichirou Torisu, and Akito Uekihara (Saga Univ., 1-Honjo, Saga-shi, Japan, 8408502, tera@me.saga-u.ac.jp)

A novel blind source separation of a mixture of two or more voice signals has been proposed in the present paper. The separation system has been focused based on the spatio-temporal gradient analysis. The proposed algorithm utilizes the linearity among the signals: sound pressure of source signals, the three-dimensional (x, y, and z directional) particle velocity vector, and its gradient of the observed signals, all of which are governed by the equation of motion. Principally, as the mechanism of blind source separation uses no- *priori* information about the parameters of convolution, filtering as well as mixing of source signals, some simple assumptions such as the statistical independency of the linearly combined (mixed) observed signals containing zero mean as well as unit variance have been implied in the present separation algorithm. Therefore, the proposed method has successfully simplified the convoluted blind source separation problem into an instantaneous blind source separation problem over the spatio-temporal gradient spaces. An acoustic experiment with two female voices has been carried out to compare the simulated data as well. A eight-microphone array system has been adopted to evaluate the voice signals efficiently.

4:05

1pSP10. Study for blind source separation on time-frequency domain considering phase information. Fumio Sasaki, Masahito Yasuoka, and Osamu Tanaka (Dept. of Architecture, Tokyo Univ. of Sci., 1-3 Kagurazaka, Shinjuku-ku, Tokyo 1628601, Japan, fsasaki@rs.kagu.tus.ac.jp)

The methods that can be done for the estimation of the number of source signals and separation of source signals using time-frequency information are proposed as a method of blind source separation. In these methods, some independent conditions are assumed in a time-frequency domain, and observed signals are expanded to time-frequency domain using wavelet analysis. However, these methods are not considered phase differences (difference of distance) between source signals and observed signals. A new method is proposed. The method can be done not only for the source separation but also the specification of the locations of source signals considering phase differences. In this method, it is necessary to assume a rather stronger independent condition than former methods. But, on account of this assumption, a function that becomes a real value only when phase differences are coincidence in time-frequency domain is determined. Using this function, first of all, the number of source signals is calculated, then the locations of source signals and source signals are calculated. The effectiveness of the method is shown using actual synthetic signals.

4:20

1pSP11. Blind spatial subtraction array based on independent component analysis for speech enhancement and recognition. Yu Takahashi, Tomoya Takatani, Hiroshi Saruwatari, and Kiyohiro Shikano (Speech and Acoust. Processing Lab., Nara Inst. of Sci. and Technol., 8916-5 Takayama-cho, Ikoma-shi, Nara, 630-0192 Japan)

We propose a new blind spatial subtraction array (BSSA) that contains an accurate noise estimator based on independent component analysis (ICA) for the realization of noise-robust hands-free speech recognition. Many previous studies on ICA-based blind source separation often dealt with the special case of speech-speech mixing. However, such a sound mixing is not realistic under common acoustic conditions; the target speech can be approximated to a point source but real noises are often not point sources. Under the condition, our preliminary experiment suggests that the conventional ICA is proficient in the noise estimation rather than the direct speech estimation. Based on the above-mentioned findings, we propose a new noise reduction method that is implemented in subtracting the power spectrum of the estimated noise by ICA from the power spectrum of noise-contaminated observations. This architecture provides us a noise-estimation-error robust speech enhancement rather than a simple linear-filtering-based enhancement. Although nonlinear processing often generates an artificial distortion, the so-called musical noise, it is still applicable to the speech recognition system because the speech decoder is not so sensitive to such a distortion. Experimental results reveal that the proposed BSSA can improve the speech recognition rate by 20% compared with the conventional ICA.

4:35

1pSP12. Frequency-domain independent component analysis by overlap piecewise integration of separation processing. Tadashige Noguchi, Kenko Ota, Masuzo Yanagida (Doshisha Univ., 1-3 Tatara-Miyakodani, Kyotanabe, Kyoto, 610-0321 Japan, dtg0731@mail4.doshisha.ac.jp), and Leandro Di Persia (Universidad Nacional de Entre Ros Casa de la Universidad, Paran, Entre Ros, Argentina)

Conventional frequency-domain ICA yields the optimal separation for each frequency bin, but it suffers from the permutation problem. The authors have developed permutation-free ICA as a separation scheme by obtaining the separation matrix for a long vector consisting of temporal changes of all frequency components of the received signal. The permutation-free ICA, however, only yields a common separation matrix for all frequency bins. So, the separation matrix obtained in the permutation-free ICA has a common directivity pattern for all frequency bins, though the method can avoid the permutation problem. Proposed in this paper is a scheme of multibin ICA that deconvolves mixed signals into original source signals by shifting piecewise integration of a set of frequency bins consisting of the frequency bin in concern and neighboring frequency bins. The proposed method can yield nearly optimal directivity in the form of the separation matrix for each frequency bin, avoiding the permutation problem. Performance of multibin ICA is compared with those of conventional frequency-domain ICA and permutation-free ICA employing segmental SNR as an evaluation index. [Work supported by Knowledge Cluster Project, MEXT, and by Academic Frontier Project Doshisha University.]

TUESDAY AFTERNOON, 28 NOVEMBER 2006 — WAIALUA ROOM, 1:00 TO 5:35 P.M.

Session 1pUW

Underwater Acoustics: Array Processing, Sensors, and Technology

Paul Hursky, Cochair
Heat Light and Sound Research, 12730 High Bluff Dr., San Diego, CA 92130

Hiroyuki Hachiya, Cochair
Chiba Univ., Research Ctr. for Frontier Medical Engineering, 1-33 Yayoi-cho, Inake-ku, Chiba, 263-8522, Japan

Chair's Introduction—1:00

Contributed Papers

1:05

1pUW1. Acoustic seaglider. Bruce M. Howe (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105-6698, howe@apl.washington.edu)

Continually improving acoustics measurements in the ocean is a *sine qua non* for the advancement of ocean acoustics and related fields. Acoustic sensors on mobile platforms will enable basic research topics on temporal and spatial coherence and the description of ambient noise, with direct impact on the applicability of coherent processing with its associated gain to signal detection and acoustic navigation and communications within the context of distributed undersea sensor networks. We describe the integration of acoustic receiving and communication capability in gliders with results from several recent field tests. [Work supported by the ONR.]

1:20

1pUW2. Single element synthetic aperture using an ocean glider. Georges A. Dossot, James H. Miller, Kristy A. Moore (Dept. of Ocean Eng., Univ. of Rhode Island, 217 Sheets Bldg., Narragansett Bay Campus, Narragansett, RI 02882, georges@oce.uri.edu), Steven E. Crocker (Naval Undersea Warfare Ctr., Newport, RI 02871), Jason D. Holmes (Boston Univ., Boston, MA 02215), and Scott Glenn (Rutgers Univeristy, NB, NJ 08901-8521)

The feasibility of using a single transducer element on an ocean glider to create a synthetic aperture is discussed. Acoustic data were collected on two Webb Slocum gliders deployed by Rutgers University during the Shallow Water Experiment (SW06) on the continental shelf off New Jersey. These gliders periodically surfaced for GPS fixes and data transfer via satellite phone. A synthetic aperture is created through coherent processing of the acoustic data as the glider travels through the water. A number of issues including changes in depth, nonsteady motion of the glider, and clock drift can limit the performance of the processing. However, the glider provides a low-noise and low-speed platform, potentially improving the signal-to-noise ratio. The geometry of the experiment provided for near-broadside recording of the 224- and 400-Hz WHOI tomography sources. An acoustic normal mode representation of the field provides the basis for processing of the data similar to the Hankel transform approach of Frisk and Lynch [J. Acoust. Soc. Am. **76**, 205–216 (1984)] and Holmes *et al.* [J. Acoust. Soc. Am. **119**, 3346 (2006)]. Effects of spatial variations in sound speed are examined. [Work sponsored by the Office of Naval Research.]

1:35

1pUW3. Processing data from a low-drag array towed by a glider. Paul Hursky, Michael B. Porter, Martin Siderius (Heat, Light, and Sound Res. Inc., 12730 High Bluff Dr., Ste. 130, San Diego, CA 92130), Vincent K. McDonald, Mark Gillcrist, Brian Granger, Ryan Jones, Aaron Bratten, Andy Huizinga, Peter T. Sullivan, and Susan G. Briest (Space and Naval Warfare Systems Ctr., San Diego, CA 92152)

Small underwater robotic vehicles such as gliders and AUVs have limited capabilities in terms of propulsion, speed, and power consumption, so it takes great care to design suitable sensor systems such as towed arrays for these platforms. Two such arrays were deployed during the Makai experiment, off the coast of Kauai in Hawaii. A 15-element line array with an acoustic aperture of 21 m was towed by a Webb Research Slocum glider. A 40-element line array with an acoustic aperture of 60 m was towed by a small work boat. We will describe the design and construction of these arrays and present results of processing data recorded on these arrays. We will also discuss potential applications for these arrays deployed from gliders and AUVs.