Hindawi Publishing Corporation
EURASIP Journal on Applied Signal Processing
Volume 2006, Article ID 46357, Pages 1–3
DOI 10.1155/ASP/2006/46357

Editorial
Advances in Multimicrophone Speech Processing

Sharon Gannot,[1] Jacob Benesty,[2] Jörg Bitzer,[3] Israel Cohen,[4] Simon Doclo,[5] Rainer Martin,[6] and Sven Nordholm[7]

[1] *School of Engineering, Bar-Ilan University, Ramat-Gan, 52900, Israel*
[2] *INRS-EMT, University of Quebec, 800 de la Gauchetiere Ouest, Montreal, QC, Canada H5A 1K6*
[3] *Institute of Audiology and Hearing Science, University of Applied Sciences, Oldenburg/Ostfriesland/Wilhelmshaven Ofener Street 16, 26121 Oldenburg, Germany*
[4] *Department of Electrical Engineering, Technion — Israel Institute of Technology, Technion City, Haifa 32000, Israel*
[5] *Department of Electrical Engineering (ESAT-SCD), Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, 3001 Leuven, Belgium*
[6] *Institute of Communication Acoustics, Ruhr-Universitaet Bochum, 44780 Bochum, Germany*
[7] *Western Australian Telecommunications Research Institute, The University of Western Australia, 35 Stirling Hwy, Crawley, 6009, Australia*

Received 18 January 2006; Accepted 18 January 2006

Copyright © 2006 Sharon Gannot et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Speech quality may significantly deteriorate in the presence of interference, especially when the speech signal is also subject to reverberation. Consequently, modern communication systems, such as cellular phones, employ some speech enhancement procedure at the preprocessing stage, prior to further processing (e.g., speech coding).

Generally, the performance of single-microphone techniques is limited, since these techniques can utilize only spectral information. Especially for the dereverberation problem, no adequate single-microphone enhancement techniques are presently available. Hence, in many applications, such as hands-free mobile telephony, voice-controlled systems, teleconferencing, and hearing instruments, a growing tendency exists to move from single-microphone systems to multimicrophone systems. Although multimicrophone systems come at an increased cost, they exhibit the advantage of incorporating both spatial and spectral information.

The use of multimicrophone systems raises many practical considerations such as tracking the desired speech source, and robustness to unknown microphone positions. Furthermore, due to the increased computational load, real-time algorithms are more difficult to obtain and hence the efficiency of the algorithms becomes a major issue.

The main focus of this special issue is on emerging methods for speech processing using multimicrophone arrays. In the following, the specific contributions are summarized and grouped according to their topic. It is interesting to note that none of the papers deal with the important and difficult problem of dereverberation.

Speaker separation

In the paper "Speaker separation and tracking system," Anliker et al. propose a two-stage integrated speaker separation and tracking system. This is an important problem with several potential applications. The authors also propose quantitative criteria to measure the performance of such a system, and present experimental evaluation of their method. In the paper "Speech source separation in convolutive environments using space-time-frequency analysis" Dubnov et al. present a new method for blind separation of convolutive mixtures based on the assumption that the signals in the time-frequency (TF) domain are partially disjoint. The method involves detection of single-source TF cells using eigenvalue decomposition of the TF-cells correlation matrices, clustering of the detected cells with expectation-maximization (EM) algorithm based on Gaussian mixture model (GMM), and estimation of smoothed transfer functions between microphones and sources via extended Kalman filtering (EKF). Serviere and Pham propose in their paper "Permutation correction in the frequency-domain in blind separation of speech mixtures" a method for blind separation of convolutive mixtures of speech signals, based on the joint diagonalization of the time-varying spectral matrices of the observation records. This paper proposes

a two-step method. First, the frequency continuity of the unmixing filters is used in the initialization of the diagonalization algorithm. Then, the continuity of the time variation of the source energy is exploited on a sliding frequency bandwidth to detect the remaining frequency permutation jumps. In their paper "Geometrical interpretation of the PCA subspace approach for overdetermined blind source separation" Winter et al. discuss approaches for blind source separation where the number of sources can exceed the number of users. Two methods are compared. The first is based on principal component analysis (PCA). The second is based on geometric considerations.

Echo cancellation

In their paper "Efficient fast stereo acoustic echo cancellation based on pairwise optimal weight realization technique," Yukawa et al. propose a class of efficient fast acoustic echo cancellation algorithms with linear computational complexity. These algorithms are based on pairwise optimal weight realization power technique. Numerical examples demonstrate that the proposed schemes significantly improve the convergence behavior compared with conventional methods in terms of system mismatch as well as echo return loss enhancement (ERLE).

Acoustic source localization

Time-delay estimation is a first stage that feeds into subsequent processing blocks for identifying, localizing, and tracking radiating sources. The paper "Time-delay estimation in room acoustic environments: an overview" by Chen et al. presents a systematic overview of the state of the art of time-delay-estimation algorithms ranging from the simple cross-correlation method to the advanced blind channel identification based techniques. In their work "Kalman filters for time-delay of arrival-based source localization," Klee et al. propose an algorithm for acoustic source localization based on time-delay-of-arrival (TDOA) estimation. In their approach, they use a Kalman filter to directly update the speaker position estimate based on the observed TDOAs. In their contribution, "Microphone array speaker localizers using spatial-temporal information," Gannot and Dvorkind propose to exploit the speaker's smooth trajectory for improving the position estimate. Based on TDOA readings, three localization schemes, which use the temporal information, are presented. The first is a recursive form of the Gauss method. The other two are extensions of the Kalman filter to the nonlinear problem at hand, namely, the extended Kalman filter and the unscented Kalman filter. In their paper, "Particle filter design using importance sampling for acoustic source localization and tracking in reverberant environments," Lehmann and Williamson develop a new particle filter for acoustic source localization using importance sampling, and compare its tracking ability with that of a bootstrap algorithm proposed previously in the literature. A real-time implementation of the algorithm also shows that the proposed particle filter can reliably track a person talking in real reverberant rooms.

Speech enhancement and speech detection

The paper "Dual channel speech enhancement by superdirective beamforming" by Lotter and Vary presents a dual channel input-output speech enhancement system. The proposed algorithm is an adaptation of the well-known superdirective beamformer including postfiltering to the binaural application. In contrast to conventional beamformer processing, the proposed system outputs enhanced stereo signals while preserving the important interaural amplitude and phase differences of the original signal. In their paper "Sector-based detection for hands-free speech enhancement in cars" Lathoud et al. investigate an adaptation control of beamforming interference cancellation techniques for in-car speech acquisition. Two efficient adaptation control methods are proposed that avoid target cancellation. Experiments on real in-car data validate both methods, including a case with 100 km/h background road noise. In their paper "Using intermicrophone correlation to detect speech in spatially-separated noise," Koul and Greenberg provide a theoretical analysis of a system for determining intervals of high and low signal-to-noise ratio when the desired signal and interfering noise arise from distinct spatial regions. The system uses the correlation coefficient between two microphone signals configured in a broadside array as the decision variable in a hypothesis test, and can, for example, be used as an adaptation control method for an adaptive beamformer.

Sharon Gannot
Jacob Benesty
Jörg Bitzer
Israel Cohen
Simon Doclo
Rainer Martin
Sven Nordholm



Sharon Gannot received his B.S. degree, *(summa cum laude)* from the Technion – Israeli Institute of Technology, Israel, in 1986, and the M.S. *(cum laude)* and Ph.D. degrees from Tel-Aviv University, Tel-Aviv, Israel, in 1995 and 2000, respectively, all in electrical engineering. From 1986 to 1993, he was the head of a research and development section, in an R&D center of the Israeli Defense Forces. In the year 2001, he held a postdoctoral position at the Department of Electrical Engineering (SISTA) at K. U. Leuven, Belgium. From 2002 to 2003, he held a Research and Teaching Position at the Signal and Image Processing Lab (SIPL), Faculty of Electrical Engineering, Technion-Israeli Institute of Technology, Israel. Currently, he is a Lecturer in the School of Engineering, Bar-Ilan University, Israel. He is also an Associate Editor of the EURASIP Journal of Applied signal Processing, an Editor of a special issue on advances in multimicrophone speech processing of the same journal, a Guest Editor of Elsevier Speech Communication Journal, and a Reviewer of many IEEE journals. His research interests include parameter estimation, statistical signal processing, and speech processing using either single- or multimicrophone arrays.



Jacob Benesty was born in 1963. He received the Masters degree in microwaves from Pierre and Marie Curie University, France, in 1987, and the Ph.D. degree in control and signal processing from Orsay University, France, in April 1991. During his Ph.D. program (from November 1989 to April 1991), he worked on adaptive filters and fast algorithms at the Centre National d'Etudes des Telecommunications (CNET), Paris, France. From January 1994 to July 1995, he worked at Telecom Paris University on multichannel adaptive filters and acoustic echo cancellation. From October 1995 to May 2003, he was first a Consultant and then a Member of the Technical Staff at Bell Laboratories, Murray Hill, NJ, USA. In May 2003, he joined the University of Quebec, INRS-EMT, in Montreal, Quebec, Canada, as an Associate Professor. His research interests are in acoustic signal processing and multimedia communications. He received the 2001 Best Paper Award from the IEEE Signal Processing Society. He was a Member of the editorial board of the EURASIP Journal on Applied Signal Processing and was the Cochair of the 1999 International Workshop on Acoustic Echo and Noise Control. He coauthored the books *Acoustic MIMO Signal Processing* (Springer, Boston, Mass, 2006) and *Advances in Network and Acoustic Echo Cancellation* (Springer, Berlin, 2001). He is also a coeditor/coauthor of the books *Speech Enhancement* (Spinger, Berlin, 2005), *Audio Signal Processing for Next Generation Multimedia Communication Systems* (Kluwer Academic Publishers, Boston, 2004), *Adaptive Signal Processing: Applications to Real-World Problems* (Springer, Berlin, 2003), and *Acoustic Signal Processing for Telecommunication* (Kluwer Academic Publishers, Boston, 2000).



Jörg Bitzer was born in Bremen in 1970. He received his Diploma and Doctorate in electrical engineering from the University of Bremen in 1996 and 2002, respectively. From 2000 to 2003, he was the Leading Researcher and the Head of the Algorithm Development Team at Houpert Digital Audio, a company specialized in audio signal processing. Since September 2003, he has been a Professor for audio signal processing at the University of Applied Science Oldenburg/Ostfriesland/Wilhelmshaven. His current research interests include beamforming, speech enhancement, audio restoration, audio effects for musical applications, and algorithms for hearing aids.



Israel Cohen received the B.S. (*Summa Cum Laude*), M.S., and Ph.D. degrees in electrical engineering in 1990, 1993, and 1998, respectively, all from the Technion–Israel Institute of Technology, Haifa, Israel. From 1990 to 1998, he was a Research Scientist at RAFAEL Research Laboratories, Haifa, Israel, Ministry of Defense. From 1998 to 2001, he was a Postdoctoral Research Associate at the Computer Science Department, Yale University, New Haven, Conn. Since 2001, he has been a Senior Lecturer with the Electrical Engineering Department, Technion, Israel. His research interests are statistical signal processing, analysis and modeling of acoustic signals, speech enhancement, noise estimation, microphone arrays, source localization, blind source separation, system identification, and adaptive filtering. He serves as an Associate Editor for the IEEE Transactions on Speech and Audio Processing and IEEE Signal Processing Letters, and as Guest Editor for a special issue of the Elsevier Speech Communication Journal on Speech Enhancement.



Simon Doclo was born in Wilrijk, Belgium, in 1974. He received the M.S. degree in electrical engineering and the Ph.D. degree in applied sciences from the Katholieke Universiteit Leuven, Belgium, in 1997 and 2003, respectively. Currently, he is a Postdoctoral Fellow of the Fund for Scientific Research-Flanders, affiliated with the Electrical Engineering Department of the Katholieke Universiteit Leuven. In 2005, he was a Visiting Postdoctoral Fellow at the Adaptive Systems Laboratory, McMaster University, Canada. His research interests are in microphone array processing for acoustic noise reduction, dereverberation and sound localisation, adaptive filtering, speech enhancement, and hearing aid technology. He received the first prize "KVIV-Studentenprijzen" (with E. De Clippel) for the best M.S. engineering thesis in Flanders in 1997, a Best Student Paper Award at the International Workshop on Acoustic Echo and Noise Control in 2001, and the EURASIP Signal Processing Best Paper Award 2003 (with M. Moonen). He was the Secretary of the IEEE Benelux Signal Processing Chapter (1998-2002) and serves as a Guest Editor for the EURASIP Journal on Applied Signal Processing.



Rainer Martin received the Dipl.-Ing. and Dr.-Ing. degrees from Aachen University of Technology, in 1988 and 1996, respectively, and the M.S.E.E. degree from Georgia Institute of Technology in 1989. From 1996 to 2002, he has been a Senior Research Engineer with the Institute of Communication Systems and Data Processing, Aachen University of Technology. From April 1998 to March 1999, he was on leave to the AT&T Speech and Image Processing Services Research Lab, Florham Park, NJ. From April 2002 until October 2003, he was a Professor of Digital Signal Processing at the Technical University of Braunschweig, Germany. Since October 2003, he is a Professor of information technology and communication acoustics at Ruhr-University Bochum, Germany. His research interests are signal processing for voice communication systems, hearing aids, acoustics, and human-machine interfaces.



Sven Nordholm was born in 1960. He got his Ph.D. in signal processing from Lund University in 1992, Licentiate of Engineering in 1989, and M.S.E.E. (Civilingenjör), 1983. He was one of the founders of the Department of Signal Processing, Blekinge Institute of Technology in Ronneby, in 1990 where he held positions as Lecturer, Senior Lecturer, Associate Professor, and Professor. Since 1999, he has been in Perth, Western Australia. From 1999 to 2002, he was the Director of the ATRI and Professor at Curtin University of Technology. Currently, he is a Professor and Director of Signal Processing Laboratories WATRI, Western Australian Telecommunication Research Institute, a joint institute between the University of Western Australia and Curtin University of Technology. He is also a Research Executive of the Wireless Program, ATcrc. His main research efforts have been spent in the fields of speech enhancement, adaptive and optimum microphone arrays, acoustic echo cancellation, adaptive signal processing, subband adaptive filtering, and filter design.