External Awards

TAF Telecom System Technology Award
Winner: Hiroshi Sawada, NTT Service Evolution Laboratories; Shoko Araki, NTT Communication Science Laboratories; Shoji Makino, Tsukuba University
Date: March 23, 2015
Organization: The Telecommunications Advancement Foundation (TAF)

For "Underdetermined Convolutive Blind Source Separation via Frequency Bin-wise Clustering and Permutation Alignment."

This paper presents a blind source separation method for convolutive mixtures of speech/audio sources. The method can even be applied to an underdetermined case where there are fewer microphones than sources. The separation operation is performed in the frequency domain and consists of two stages. In the first stage, frequency-domain mixture samples are clustered into each source by an expectation-maximization (EM) algorithm. Since the clustering is performed in a frequency bin-wise manner, the permutation ambiguities of the bin-wise clustered samples should be aligned. This is solved in the second stage by using the probability on how likely each sample belongs to the assigned class. This two-stage structure makes it possible to attain a good separation even under reverberant conditions. Experimental results for separating four speech signals with three microphones under reverberant conditions show the superiority of the new method over existing methods. We also report separation results for a benchmark data set and live recordings of speech mixtures.
Published as: H. Sawada, S. Araki, and S. Makino, "Underdetermined Convolutive Blind Source Separation via Frequency Bin-wise Clustering and Permutation Alignment," IEEE Trans. Audio, Speech, and Language Processing, Vol. 19, No. 3, pp. 516–527, Mar. 2011.

The Commendation for Science and Technology by the Minister of Education, Culture, Sports, Science and Technology, the Prize for Science and Technology (Research Category)
Winner: Shin'ya Nishida, NTT Communication Science Laboratories
Date: April 15, 2015
Organization: Ministry of Education, Culture, Sports, Science and Technology

For research on the mechanism of human perception that is achieved through the mutual interaction of different kinds of sensory attributes such as an object's color, shape, and movement, as well as sensory modalities such as visual, auditory, and tactile perception.
Published as: S. Nishida and A. Johnston, "Influence of Motion Signals on the Perceived Position of Spatial Pattern," Nature, Vol. 397, pp. 610–612, Feb. 1999, and as S. Nishida and A. Johnston, "Marker Correspondence, Not Processing Latency, Determines Temporal Binding of Visual Attributes," Current Biology, Vol. 12, No. 5, pp. 359–368, Mar. 2002.

Papers Published in Technical Journals and Conference Proceedings

Uncompressed 8K-video System Using High-speed Video Server System over IP Network
H. Kimiyama, M. Maruyama, M. Kobayashi, and M. Sakai
Proc. of APMediaCast2015 (1st Asia Pacific Conference on Multimedia and Broadcasting), pp. 99–105, Bali, Indonesia, April 2015.

We developed a scalable high-speed video server system to deliver high-quality video, including uncompressed HD and uncompressed 4K video, streams over IP networks and a synchronization method for it that combines multiple 4K transmission systems and video servers. We demonstrated the method's validity by implementing and testing an uncompressed 8K video real-time transmission system and an uncompressed 8K video on-demand system. Through experiments performed on both systems via commercial 100-Gbit/s Ethernet, we succeeded in achieving the world's first bidirectional uncompressed 8K video transmission and 8K video on-demand systems.

Multi-resolution Signal Decomposition with Time-domain Spectrogram Factorization
H. Kameoka
Proc. of ICASSP 2015 (40th International Conference on Acoustics, Speech and Signal Processing), pp. 86–90, Brisbane, Australia, April 2015.

This paper proposes a novel framework that makes it possible to realize non-negative matrix factorization (NMF)-like signal decompositions in the time domain. This new formulation also allows for an extension to multi-resolution signal decomposition, which was not possible with the conventional NMF framework.

L_P-norm Non-negative Matrix Factorization and Its Application to Singing Voice Enhancement

T. Nakamura and H. Kameoka

Proc. of ICASSP 2015, pp. 2115–2119, Brisbane, Australia, April 2015.

Measures of sparsity are useful in many aspects of audio signal processing including speech enhancement, audio coding, and singing voice enhancement, and the well-known method for these applications is non-negative matrix factorization (NMF), which decomposes a non-negative data matrix into two non-negative matrices. Although previous studies on NMF have focused on the sparsity of the two matrices, the sparsity of reconstruction errors between a data matrix and the two matrices is also important, since designing the sparsity is equivalent to assuming the nature of the errors. We propose a new NMF technique, which we called L_p-norm NMF, that minimizes the L_p norm of the reconstruction errors, and we derive a computationally efficient algorithm for L_p-norm NMF according to an auxiliary function principle. This algorithm can be generalized for the factorization of a real-valued matrix into the product of two real-valued matrices. We apply the algorithm to singing voice enhancement and show that adequately selecting p improves the enhancement.

Generative Modeling of Voice Fundamental Frequency Contours

H. Kameoka, K. Yoshizato, T. Ishihara, K. Kadowaki, Y. Ohishi, and K. Kashino

IEEE/ACM Transactions on Audio, Speech, and Language Processing, Vol. 23, No. 6, pp. 1042–1053, June 2015.

This paper introduces a generative model of voice fundamental frequency (F_0) contours that allows us to extract prosodic features from raw speech data. The present contour model is formulated by translating the Fujisaki model, a well-founded mathematical model representing the control mechanism of vocal fold vibration, into a probabilistic model described as a discrete-time stochastic process.

Motion Estimation for Dynamic Texture Videos based on Locally and Globally Varying Models

H. Sakaino

IEEE Transactions on Image Processing, June 2015 (published on web; print version will be issued in November 2015).

Motion estimation, i.e., optical flow, of fluid-like and dynamic texture (DT) images/videos is an important challenge, particularly for understanding outdoor scene changes created by objects and/or natural phenomena. Most optical flow models use smoothness-based constraints using terms such as fluidity from the fluid dynamics framework, with constraints typically being incompressibility and low Reynolds numbers (Re). Such constraints are assumed to impede the clear capture of locally abrupt image intensity and motion changes, i.e., discontinuities and/or high Re over time. This paper exploits novel physics based optical flow models/constraints for both smooth and discontinuous changes using a wave generation theory that imposes no constraint on Re or compressibility of an image sequence. Iterated two-step optimization between local and global optimization is also used; first, an objective function with varying multiple sine/cosine bases with new local image properties, i.e., orientation and frequency, and with a novel transformed dispersion relationship equation are used. Second, the statistical property of image features is used to globally optimize model parameters. Experiments on synthetic and real DT image sequences with smooth and discontinuous motions demonstrate that the proposed locally and globally varying models outperform previous optical flow models.

Development of Wireless Systems for Disaster Recovery Operations

T. Hirose, F. Nuno, and M. Nakatsugawa

IEICE Transactions on Electronics, Vol. E98-C, No. 7, pp. 630–635, July 2015.

This paper presents wireless systems for use in disaster recovery operations. The Great East Japan Earthquake of March 11, 2011 reinforced the importance of communications in, to, and between disaster areas as lifelines. It also revealed that conventional wireless systems used for disaster recovery need to be renovated to cope with technological changes and to provide their services with easier operations. To address this need, we have developed new systems, which include a relay wireless system, subscriber wireless systems, business radio systems, and satellite communication systems. They will be chosen and used depending on the situations in disaster areas as well as on the required services.