Unoki Laboratory

Associate Professor : Masashi Unoki

Auditory-motivated sound signal processing

URL http://www.jaist.ac.jp/~unoki/ E-mail unoki@jaist.ac.jp

Research Overview

Humans can easily listen to target sounds that they want to hear in real environments, such as those that are noisy and reverberant one. In addition, hearing abilities can be improved by using attention. However, it is very difficult for machines (i.e., computers) to do the same thing. Implementing auditory signal processing with the same functions as those of human hearing systems onto computers would enable us to accomplish human-like speech signal processing. Such a processing system would be highly suitable for a range of applications, such as speech recognition processing and hearing aids. Achieving this is the ultimate goal of our research team.

Auditory filterbank

The main function of the human auditory system is to decompose sound signals into frequency components (i.e., frequency selectivity), as shown in Fig. 1. It is well known that this frequency selectivity involves nonlinear signal processing. We have been correcting the masking data of various masking situations to find nonlinear frequency selectivity. We have been constructing a nonlinear auditory filterbank whose function is equivalent that of the human hearing system. In addition, we have been investigating active hearing by using attention.

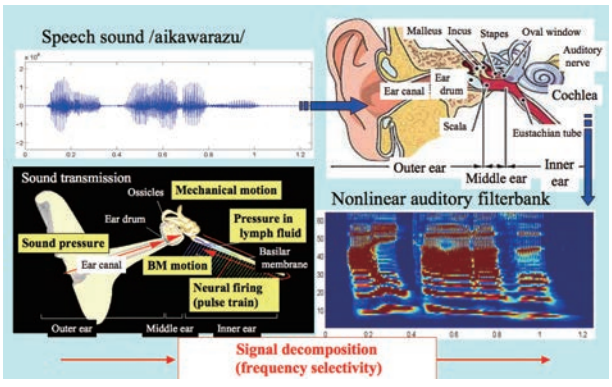

Fig. 1 Signal representation via auditory filterbank

Auditory-motivated speech signal processing

The following research projects have been used in an auditory filterbank to process speech signals: a selective sound segregation model and speech enhancement method based on the concept of the modulation transfer function. Our main purpose was to model the 'cocktail party effect' and to apply this model to solving challenging problems by developing our research projects into a nonlinear auditory filterbank and attain auditory signal processing.

A research project on multimedia information hiding, based on human auditory characteristics, is also currently being carried out for Internet security. There are, for example, digital rights management (DRM) problems with CDs, movies, and Internet speech communications, as shown in Fig. 2. We have been developing a digital audio watermarking technique based on human cochlear delay, which enables us to embed inaudible information into sound and to detect it from this. This technique has three main advantages of inaudibility, robustness against attacks, and confidentiality.

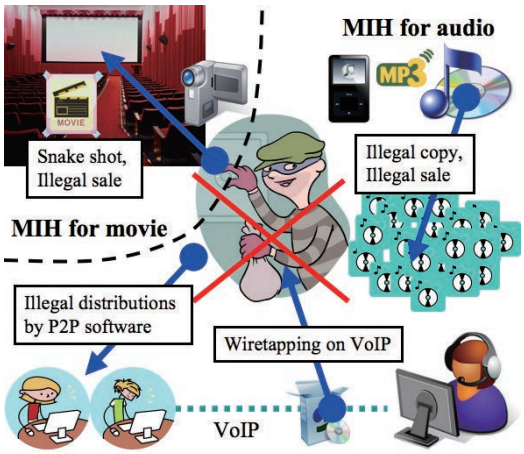

Fig. 2 Multimedia information hiding and its applications

Publications

- Nhut Mihn Ngo, Masashi Unoki, Ryota Miyauchi, Yôichi Suzuki,"Data Hiding Scheme for Amplitude Modulation Radio Broadcasting Systems," Journal of Information Hiding and Multimedia Signal Processing, Vol. 5, No. 3, pp. 307-318, July 2014.
- Masashi Unoki and Ryota Miyauchi, "Method of Digital-Audio Watermarking Based on Cochlear Delay Characteristics," Multimedia Information Hiding Technologies and Methodologies for Controlling Data, Ed. Kazuhiko Kondo, Chapter 2, pp. 42-70, IGI Global, 2012.
- Shunsuke Kidani and Masashi Unoki, "Effect of Presence of Cue Tone on Tuning of Auditory Filter Derived from Simultaneous Masking," Neurophysiological Bases of Auditory Perception, Edited by Enrique Lopez-Poveda, Alan R. Palmer, and Ray Meddis, Chapter 12, pp. 121–130, Springer, New York, 2010.
- Masashi Unoki, Toshio Irino, Brian Glasberg, Brian C. J. Moore, and Roy D. Patterson, "Comparison of the roex and gammachirp filters as representations of the auditory filter", J. Acoust. Soc. Am., vol. 120, no. 3, pp. 1474–1492, 2006.