THE JOURNAL OF THE ACOUSTICAL SOCIETY OF KOREA Vol.28, No.3E 2009. 9. pp. 79~87

Comparison of Sound Source Localization Methods Based on Zero Crossings

Yong-Jin Park*, Soo-Yeon Lee*, Hyung-Min Park*
*Department of Electronic Engineering, Sogang University
(Received September 1 2009; accepted September 5 2009)

Abstract

This paper reviews several multi-source localization methods which estimate ITDs based on zero crossings (ZCs). Employing signal-to-noise ratio (SNR) estimation from ITD variances, these ZC-based source localization methods are more robust to diffuse noise than the cross-correlation (CC)-based one with less computational complexity. In order to take reverberant environments into account, two approaches detect intervals which dominantly contain direct-path components from sources to sensors because they may effectively provide reliable ITDs corresponding to source directions. One accomplishes the detection by comparing the original and cepstral-prefiltering-processed envelopes, and the other searches sudden increase of acoustic energy by considering typical characteristics of acoustic reverberation. Experiments for comparison of these methods demonstrate that the approach with energy-based detection efficiently achieves multi-source localization in reverberant environments.

Keywords: *Source localization, Zero crossings, Interaural time difference, Binaural processing, Onset*

I. Introduction

Human beings have an outstanding ability in sound source localization, and it plays an important role to select a particular sound source and track the sound originating from that source. It is known that the ability can be achieved by exploiting the differences between signals obtained from both ears. The two primary cues are inter-aural time differences (ITDs) and inter-aural intensity difference (IIDs). ITDs are the main cues to deal with sound components at frequencies below 1.5 kHz while IIDs can be used for higher-frequency components. The ITDs can also be useful to localize low-frequency envelopes of higher-frequency components [1].

In 1948, Jeffress suggested a simple and intuitive hypothesis to estimate ITDs based on interaural coincidences in the human auditory system, and then many researchers developed various computational models for sound source localization [2-5]. Most of these included frequency analysis and a mechanism for computing the interaural cross-correlation (CC) function to estimate ITDs in every frequency bands. Unfortunately, the CC-based ITD estimation methods require high computational complexity involved in the computation of CC, and they may suffer from inaccuracies in estimating the ITDs especially in noisy multi-source environments since some spurious peaks are usually generated from the computation of CC.

In order to overcome these disadvantages of the CC-based ITD estimation methods, Kim et al. presented a method which compares zero crossings (ZCs) from band-pass signals for an accurate and efficient estimation of ITDs in noisy multi-source environments [6]. Originally, ZCs have been used to find noise-robust speech features. Especially, the ZCs with peak

Corresponding author: Hyung-Min Park (hpark@sogang.ac.kr)
Department of Electronic Engineering, Sogang University
1 Shinsu-dong, Mapo-gu, Seoul 121-742, Korea

amplitudes (ZCPA) modelled the neural transduction of sound based on two parallel mechanisms of auditory nerve fibers: rate and temporal representations. Experimental results showed that the ZCPA may provide a more robust performance than some conventional speech features [7]. This is mainly due to the dominant frequency principle which states that the number of ZCs per unit time is close to two times the frequency of a dominant signal component when one exists [8][9]. In this paper, we review and compare several multi-source localization methods which estimate ITDs based on ZCs [6][10][11].

II. Sound source localization methods using ZC-based ITD estimation

Basically, source localization methods require three fundamental steps: frequency analysis, ITD estimation, and histogram construction to localize sound sources. The frequency analysis can be achieved by cochlear filtering which is simulated by passing input signals through a bank of Gammatone filters [4]. We considered a 25-channel filterbank with center frequencies spaced linearly in the equivalent rectangular bandwidth from 100 Hz to 4 kHz, and it was implemented by the approach as described in [5].

ITDs are estimated from band-pass outputs of Gammatone filters. As mentioned above, we consider ITD estimation which is based on ZCs. A ZC time is detected by identifying a sample point at which a filtered output signal changed from a negative value to a positive value or vice versa, and then elaborately estimating by linear interpolation between the two adjacent sample points that straddle the ZC. An estimated ITD is defined to be the relative time difference between ZCs observed from the two sensors.

Especially, Kim *et al.* described a method to measure the reliability of ITDs by estimating the signal-to-noise ratio (SNR) and achieved robust localization in noisy multi-source environments [6]. Let us describe ZC times of noisy inputs from the two sensors in a channel by the following equations:

$$t_i^L(j) = \bar{t}_i^L(j) + r_i^L(j), \tag{1}$$

$$t_i^R(j) = \bar{t}_i^R(j) + r_i^R(j), \tag{2}$$

where $\bar{t}_i^L(j)$ and $\bar{t}_i^R(j)$ represent the ZC times of clean band-pass signals without noise from the left and right sensors, respectively. i and j denote the channel index of a Gammatone filterbank and the ITD sample index, respectively. $r_i^L(j)$ and $r_i^R(j)$ denote the perturbation of the ZC times due to noise from the sensors, and they are assumed to be identically and independently distributed with zero mean. The true ITD between the two sensors is represented by $\Delta_i(j)$ to get

$$\bar{t}_i^L(j) = \bar{t}_i^R(j) + \Delta_i(j). \tag{3}$$

Then, the mean of the estimated ITD $\Delta t_i(j)$ is given by

$$E[\Delta t_i(j)] = E[\Delta_i(j) + r_i^L(j) - r_i^R(j)] = \Delta_i(j). \tag{4}$$

If we assume that $r_i^L(j)$ and $r_i^R(j)$ are independent of each other and have zero mean, its variance can be derived as

$$\begin{aligned} Var[\Delta t_i(j)] &= E[\{\Delta t_i(j) - \Delta_i(j)\}^2] \\ &= Var[r_i^L(j)] + Var[r_i^R(j)]. \end{aligned} \tag{5}$$

Kim et al. analyzed the variances $Var[r_i^L(j)]$ and $Var[r_i^R(j)]$ to show

$$Var[\Delta t_i(j)] \approx \frac{1}{2\omega_i^2}\left(\frac{1}{10^{SNR_i^L(j)/10}} + \frac{1}{10^{SNR_i^R(j)/10}}\right). \tag{6}$$

where $SNR_i^L(j)$ and $SNR_i^R(j)$ denote the SNRs of band-pass inputs from the left and right sensors, respectively [6]. ω_i represent the i-th channel band-pass input frequency. If intensity difference between the two sensors is negligible, the common

SNR can be obtained by

$$SNR_i(j) \approx 10\log_{10}\frac{1}{\omega_i^2 Var[\Delta t_i(j)]}. \quad (7)$$

Eq. (7) implies that the SNR may be approximately estimated from the variance of ITD samples and the center frequency of the channel. The estimated SNR can be effectively used for identifying reliable ITD samples to achieve noise-robust sound source localization.

Finally, to get a weighted histogram, each estimated ITD is converted into the corresponding azimuth angle, and its estimated SNR value is added to a bin including the azimuth angle. We accumulate the SNR values for all the ITDs across channels, and then search the peaks in the histogram to identify azimuth angles corresponding to source directions.

The ITD estimation has an ambiguity in the temporal analysis because of periodicity of band-pass signals, and the ambiguity disturbs accurate ITD estimation more seriously as a higher frequency band is considered. Several approaches exploited IIDs of the signal to disambiguate it [12][13]. However, IIDs can be easily biased especially in reverberant environments [14]. Since SNR estimation can identify reliable ITD samples from observations contaminated by diffuse noise, we employ only the ITDs to localize sound sources. In order to avoid the ambiguity in ITD estimation without exploiting IIDs, we make use of closely spaced sensors so that the largest possible ITDs between the sensors are always less than half a period over all considered frequencies. Therefore, the closest zero crossings across sensors provide the desired ITD value. As a result, the closely spaced sensors can localize sound sources with a simple implementation because they do not need disambiguation in ITD estimation. Also, they are appropriate for compact implementation, and the estimated ITDs might be more reliable because the distortion between sensor inputs is small.

III. Localization methods in reverberant environments

Although ZC-based ITDs with the SNR estimation can provide desired source directions from observations corrupted by diffuse noise, many practical applications involve acoustic reverberation, and sound sources should be successfully localized in these reverberant environments. Unfortunately, it is known that the reverberation significantly degrades localization accuracy [15]. Signal components through direct paths from sources to sensors generate ITDs which correspond to sound source directions whereas reflection components interfere with the desired ITDs. Here, we consider two noticeable ZC-based approaches to sound source localization in reverberant environments [10][11].

3.1. A method using a cepstral prefiltering technique [10]

This method starts from reducing the effect of reverberation in observations. If diffuse noise can be ignored, the mixture signals of microphones can then be expressed as follows:

$$x(n) = h(n) * s(n), \quad (8)$$

where $h(n)$ and $s(n)$ represent the source signal and the transmission channel between the source and each of microphones, respectively. $*$ denotes the operation of convolution. The complex cepstrum of this signal is defined as [16]

$$\begin{aligned}\hat{x}(k) &= F^{-1}\{\log X(w)\} \\ &= F^{-1}\{\log[H(w)S(w)]\} = \hat{h}(k) + \hat{s}(k),\end{aligned} \quad (9)$$

where $X(w)$, $H(w)$ and $S(w)$ are the Fourier transforms of $x(n)$, $h(n)$ and $s(n)$, respectively. $F^{-1}\{\bullet\}$ represents the inverse Fourier transform, and $\hat{h}(k)$ and $\hat{s}(k)$ are the cepstrum value of $h(n)$ and $s(n)$, respectively.

In the frequency domain, the room impulse response $H(w)$ can be decomposed into a minimum–phase component (MPC) and an all–pass component (APC) [16].

$$H(w) = H_{ap}(w)H_{min}(w). \tag{10}$$

While $H_{ap}(w)$ corresponding to APC is related to time delay, $H_{min}(w)$ corresponding to MPC is the component which makes the distortion of the speech signal by reverberant terms. If we compute the cepstrum of Eq. (10), we can represent that

$$\hat{h}(k) = \hat{h}_{ap}(k) + \hat{h}_{min}(k), \tag{11}$$

where $\hat{h}_{ap}(k)$ and $\hat{h}_{min}(k)$ can be computed by their properties of symmetry as follows:

$$\hat{h}_{ap}(k) = \begin{cases} \hat{h}(k), & k<0, \\ 0, & k=0, \\ -\hat{h}(-k), & k>0, \end{cases} \tag{12}$$

and

$$\hat{h}_{min}(k) = \begin{cases} 0, & k<0, \\ \hat{h}(k), & k=0, \\ \hat{h}(k)+\hat{h}(-k), & k>0. \end{cases} \tag{13}$$

Using (9) and (11), we can represent the cepstrum of microphone signal as

$$\hat{x}(k) = \hat{h}_{ap}(k) + \hat{h}_{min}(k) + \hat{s}(k). \tag{14}$$

As Eq. (14) reveals, we can accurately estimate the time delay between a signal and a microphone to localize sound sources by preventing speech signal from the distortion by reverberation, which may result from subtracting and eliminating the MPC part of the channel cepstra from the whole microphone cepstra. Note that computing the cepstra is processed on a frame by frame basis. Usually because of the non–stationary property of speech signal, the MPC of speech cepstra varies in each frame but the MPC of channel cepstra does not vary a lot by the fixed location of a source and a microphone. Assuming that the MPC of the speech cepstra is zero mean, the MPC of the channel cepstra can be computed by averaging the MPC of microphone cepstra recursively for input frames. The MPC of the channel cepstra denoted by $\hat{h}_{min}(k)$ is subtracted from the microphone cepstra $\hat{x}(k)$. The obtained cepstra information is transformed back to the time domain. Finally we can get a new signal with less reverberant effect [17].

In reverberant environments, we should compute the ITDs corresponding to the accurate direction of sources which are determined by the signal components through the direct paths from sources to microphones. By comparing the envelopes of an original signal and cepstral–prefiltering–processed signal, we can detect intervals which are dominated by the direct paths. In the intervals affected by reverberation, the resulting envelope of the cepstral prefiltering is smaller than the original one since the cepstral pre–filtering technique reduces the effect of reverberation by subtracting the MPCs of the acoustic channel cepstra. Therefore, we can detect the intervals dominated by the direct–path components as follows:

$$\frac{E_i(n)}{E_{Ci}(n)} < \mathrm{Th}_1 * m\left(\frac{E_i(n)}{E_{Ci}(n)}\right). \tag{15}$$

Here, $m(\cdot)$ denotes the median value. $E_i(n)$ denotes the i–th channel smoothened envelope which is the moving average values given by

$$E_i(n) = \frac{1}{2N+1}\sum_{k=-N}^{N} e_i(n+k), \tag{16}$$

where $e_i(n)$ is an instantaneous envelope and N determines the number of envelopes to be averaged. The envelope can be obtained by the magnitude of a Gammatone filter output after replacing the cosine term with a complex sinusoid [5,18]. Also, $E_{C_i}(n)$ represent the envelope for a processed signal. The second condition is for avoiding intervals with too small energy. Accordingly, we can approximately estimate intervals dominated by direct–path components and then find the accurate starting and ending points of the intervals. Comparing to reverberant–path signals, a direct path is short, so the signal through this path reaches the microphone faster. Therefore, it is effective to find onset intervals which contain fewer reflection components. The accurate starting point is chosen by the time which has the minimum envelope value in a period of duration T before the estimated duration by Eq. (15). On the other hand, the ending point is selected by the time at the maximum envelope value in a period of duration T after the estimated duration by Eq. (15).

It is possible to estimate accurate source direction by selecting ITDs only in the intervals detected above. Hence, we first obtain output signal through a Gammatone filterbank and then construct a weighted histogram from the ITDs based on zero crossings only in the intervals detected by the cepstral prefiltering technique.

Even though this approach detects the intervals elaborately, it may still contain reflection components primarily due to late reflections. That is because lots of impulse responses of acoustic reverberation have so long tails that signal components from late reflections are frequently overlapped with the onset parts. Therefore, we need to discriminate and neglect ITDs affected by late reflection components even though they are in onset parts. Usually, late reflections are very close to one another, and these reflections generate different ITDs corresponding to different paths. Therefore, an ITD from input signals reverberated by the late reflections is usually different even from adjacent ones, and the variance of ITDs may estimate how much the current input signal is contaminated by late reflection components. Kim *et al.* [6] estimated the SNR from the variance of ITD samples in a window which is given by Eq. (7), but here the SNR estimation method is employed to measure how much late reflection components interfere with ITDs.

3.2. A method using energy-based detection of onset intervals [11]

Although the previous approach can improve estimation of source directions by reducing the effect of reverberant components and comparing the original and cepstral-prefiltering-processed envelopes. It requires a great deal of additional computations to estimate the reverberant components. In addition, remaining reverberant components still significantly disturb source localization, which may result in a limit in the performance improvement.

In order to accomplish efficient sound source localization in reverberant environments, we have to elaborately select ITDs from signal components through direct paths from sources to sensors, which correspond to the directions of sources. Figure 1 illustrates the overall procedure of the approach. Intervals which may dominantly contain the direct components generally correspond to the onset parts of input signal.

Here, we also detect the intervals by employing the envelope of a Gammatone filter output. Since a large ratio between current and previous smoothened envelopes generally corresponds to an onset of a signal, it can be determined by

$$\frac{E_i(n)}{E_i(n-\Delta n)} > \mathrm{Th}_2, \qquad (16)$$

where Δn represents the time difference between the current and previous envelopes.

Once an onset is detected, we need to refine an elaborated segment which may provide ITDs corresponding to a source direction. The exact starting point is chosen by the time which has the minimum envelope value in a period of duration T before a

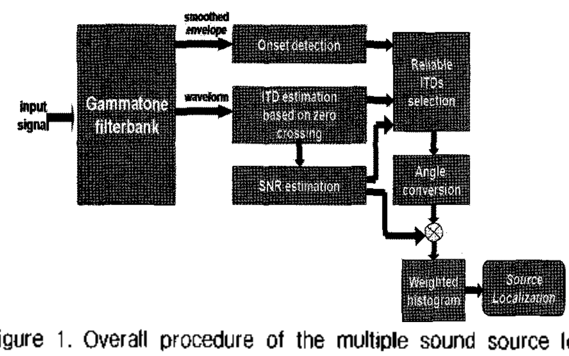

Figure 1. Overall procedure of the multiple sound source localization based on zero-crossings with detecting onset durations.

detected onset interval because direct components of a signal suddenly increase acoustic energy in contrast with reflection components. On the other hand, the exact ending point is selected by the time at the maximum envelope value in a period of duration T after a detected onset interval, due to the fact that a signal after the maximum envelope value may contain many reflection components.

Similar to the previous approach, this method also employs the SNR estimation to discriminate ITDs which are affected by late reflection components even though the onset interval is carefully refined.

Note that the detection of direct components based on onset and SNR estimation is much simpler and more efficient than the dereverberation or channel estimation methods including the previous approach based on the cepstral prefiltering technique [17][19-21]. In addition, this method using energy-based detection of onset intervals should be combined with ZC-based ITD estimation because the onset detection requires ITD estimations at specific time indices so as to select ITDs strictly in onset parts and the SNR estimation needs a number of ITDs in a short interval so as to estimate reliable sample variances.

IV. Experimental results

We have compared the presented methods in the following experimental setup. The reverberated signal at a sensor from a source was obtained by convolving the source signal with an impulse response which simulates acoustics from the source to the sensor [22], and an observation was created by combining all reverberated signals at a sensor. Figure 2 represents a configuration of sources and sensors used in the experiment. The reflection coefficients were chosen to provide the reverberation time RT_{60} of 0.3 and 0.5 seconds. Each source signal was concatenated sentences uttered by a speaker from the TIMIT database [23], and its sampling rate was 16 kHz. Since we simulated signal measurement from the sensors which were nominally separated by

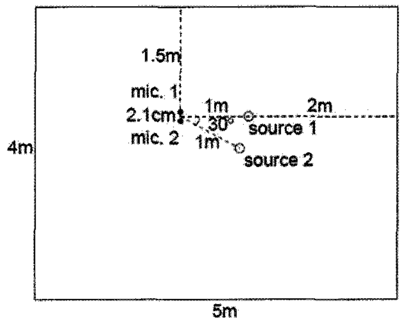

Figure 2. A configuration of sources and microphones in a rectangular room. The simulated height of the room was 3 m, and the height of all sources and microphones was 1.1 m.

Table 1. Parameter values in the experiments.

Parameter	Th_1	Th_2	Δn	N	T
Value	0.6	2.5	1 ms	1.5 ms	10 ms

43 mm to avoid spatial aliasing up to 4 kHz, the sampling period is too coarse to get a sufficient azimuthal resolution for the simulated reverberant filters. Thus, source signals were upsampled to 1,024 kHz and convolved with the reverberant filters which were originally generated at the 1,024-kHz sampling rate. After combining all the convolved signals at a sensor, the resulting signal was downsampled back to 16 kHz.

To evaluate these methods for sound source localization, we have used the parameter values in Table 1. Also, we estimated an SNR for each ITD by calculating the sampling variance in a window of 5 ZCs since an onset interval is usually quite short. Also, histograms were composed of ITDs whose estimated SNRs were greater than 30 dB, and their azimuthal resolution was 1°.

Figure 3 presents the histograms for sound source localization using observations as described above when the reverberation time is 0.3 sec. Figure 3(a) shows the result for the basic ZC-based method without any further procedure to handle reverberated signals. Figures 3(b) and (c) displays the results for the ZC-based methods using the cepstral prefiltering technique and the energy-based onset detection, respectively. Although the cepstral prefiltering technique improved the localization performance by reducing

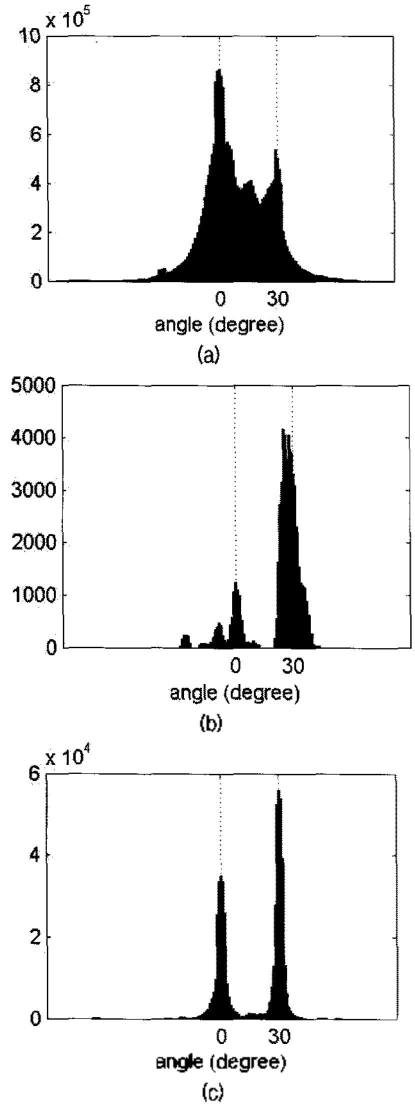

Figure 3. Weighted histograms for sound source localization in a reverberant environment with $RT_{60} = 0.3$ sec,
(a) Basic method,
(b) Using the cepstral prefiltering technique,
(c) Using the energy-based onset detection.

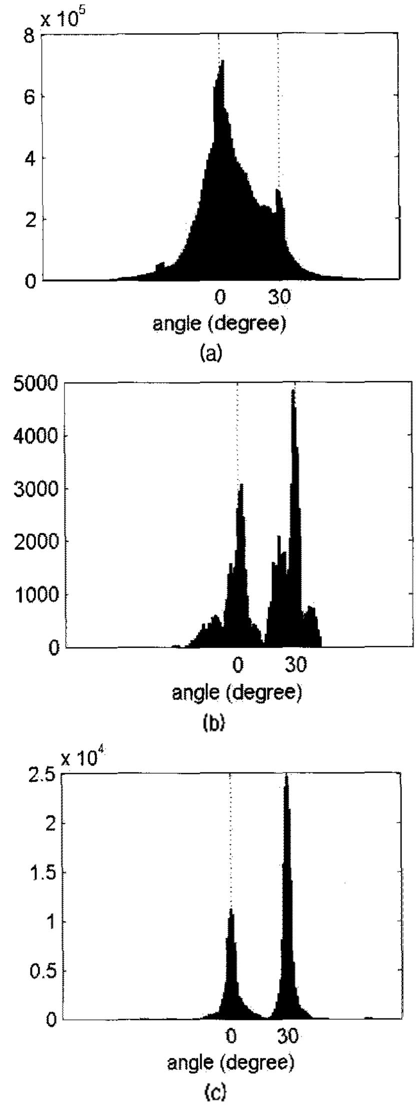

Figure 4. Weighted histograms for sound source localization in a reverberant environment with $RT_{60} = 0.5$ sec,
(a) Basic method,
(b) Using the cepstral prefiltering technique,
(c) Using the energy-based onset detection.

reverberant components, histograms obtained by the method using the energy−based onset detection dis− played two distinct peaks corresponding to the azimuths of sources. This was due to the fact that this method effectively identified intervals which were not conta− minated by reverberation and constructed a histogram from the intervals to estimate desired source directions. We repeated the experiments for a 0.5−sec reverberation time and the method using the energy−based onset detection could localize all the sources successfully as shown in Fig. 4.

V. Concluding remarks

In this paper, we have reviewed and compared three ZC−based methods to localize multiple sound sources. To achieve robustness for observations con− taminated by diffuse noises, these methods could construct weighted histograms by employing SNR estimation based on ITD variances. The cepstral pre− filtering technique tried to remove reverberant com− ponents directly so that the histograms were not affected by reverberation. Especially, the method using the energy-based onset detection could provide distinct peaks corresponding to source directions by successfully selecting intervals where signal components through direct paths are dominant. Comparison of these methods was confirmed by simulations.

Acknowledgements

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2009-0070250).

References

1. J. W. Strutt, "On our perception of sound direction," *Philosophical Magazine*, vol. 13, pp. 214-232, 1907.
2. L. A. Jeffress, "A place theory of sound localization," *Journal of Computational Physiology and Psychology*, vol. 41, pp. 35-39, 1948.
3. R. M. Stern and H. Colburn, "Theory of binaural interaction based on auditory-nerve data. IV. A model of subjective lateral position," *The Journal of the Acoustical Society of America*, vol. 64, pp. 127-140, 1978.
4. G. J. Brown and M. P. Cooke, "Computational auditory scene analysis," *Computer Speech and Language*, vol. 8, pp. 297-336, 1994.
5. M. P. Cooke, *Modeling Auditory Processing and Organisation*, Cambridge University Press, 1993.
6. Y.-I. Kim and R. M. Kil, "Estimation of interaural time differences based on zero-crossings in noisy multisource environments," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, pp. 734-743, 2007.
7. D.-S. Kim, S.-Y. Lee, and R. M. Kil, "Auditory processing of speech signals for robust speech recognition in real-world noisy environment," *IEEE Transactions on Speech and Audio Processing*, vol. 7, no. 1, pp. 55-69, 1999.
8. N. Blachman, "Zero-crossing rate for the sum of two sinusoids or a signal plus noise," *IEEE Transactions on Information Theory*, vol. 21, no. 6, pp. 671-675, 1975.
9. B. Kedem, *Time Series Analysis by Higher Order Crossings*, IEEE Press, 1994.
10. Y.-J. Park, S.-Y. Lee, H.-M. Park, "Zero-crossing-based source direction estimation using a cepstral prefiltering technique," *Journal of the Korean Society of Phonetic Sciences and Speech Technology*, no. 67, pp. 121-133, 2008.
11. Y.-J. Park and H.-M. Park, "Non-stationary sound source localization based on zero crossings with the detection of onset intervals," *IEICE Electronics Express*, vol. 5, no. 24, pp. 1054-1060, 2008.
12. N. Roman, D. Wang, and G. Brown, "Speech segregation based on sound localization," *The Journal of Acoustic Society of America*, vol. 114, pp. 2236-2252, 2003.
13. H. Viste and G. Evangelista, "Binaural source localization," in *Proceedings of International Conference on Digital Audio Effects*, pp. 145-150, Oct. 2004.
14. B. Shinn-Cunningham, N. Kopco, and T. Martin, "Localizing nearby sound sources in a classroom: Binaural room impulse responses," *The Journal of Acoustic Society of America*, vol. 117, no. 5, pp. 3100-3115, 2005.
15. B. Champagne, S. Bedard, and A. Stephenne, "Performance of time-delay estimation in the presence of room reverberation," *IEEE Transaction on Speech and Audio Processing*, vol. 4, pp. 148-152, 1996.
16. A. V. Oppenheim, R. W. Schafer, *Discrete-Time Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1975.
17. A. Stephenne and B. Champagne, "A new cepstral prefiltering technique for estimating time delay under reverberant conditions," *Signal Processing*, vol. 59, pp. 253-266, 1997.
18. K. J. Palomaki, G. J. Brown, and D. Wang, "A binaural processor for missing data speech recognition in the presence of noise and small-room reverberation," *Speech Communication*, vol. 43, pp. 361-378, 2004.
19. J. Benesty, "Adaptive eigenvalue decomposition algorithm for passive acoustic source localization," *The Journal of the Acoustical Society of America*, vol. 107, pp. 384-391, 2000.
20. S. Doclo and M. Moonen, "Robust adaptive time delay estimation for speaker localization in noisy and reverberant acoustic environments," *EURASIP Journal on Applied Signal Processing*, vol. 11, pp. 1110-1124, 2003.
21. F. Antonacci, D. Lonoce, M. Motta, A. Sarti, and S. Tubaro, "Efficient source localization and tracking in reverberant environments using microphone arrays," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'05)*, pp. 1061-1064, 2005.
22. J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *The Journal of the Acoustical Society of America*, vol. 65, pp. 943-950, 1979.
23. J. S. Garofolo, L. F. Lamel, W. M. Fisher, J. G. Fiscus, D. S. Pallett, and N. L. Dahlgren, "DARPA TIMIT acoustic-phonetic continuous speech corpus," National Institute of Standards and Technology (NIST), Tech. Rep., 1993.

[Profile]

• **Yong-Jin Park**

Yong-Jin Park received the B.S. and M.S. degrees in Electronic Engineering from Sogang University, Seoul, Korea, in 2007 and 2009, respectively. His research interests are the theory and applications of acoustic signal processing such as sound source localization, independent component analysis, and acoustic echo cancellation.

• **Soo-Yeon Lee**

Soo-Yeon Lee received the B.S degree in Electronic Engineering from Sogang

University, Seoul, Korea in 2008 and has been an M.S candidate in same above. Her research interests are the theory and applications of acoustic signal processing such as sound source localization and independent component analysis.

• Hyung-Min Park

Hyung-Min Park received the B.S., M.S., and Ph.D. degrees in Electrical Engineering and Computer Science from Korea Advanced Institute of Science and Technology, Daejeon, Korea, in 1997, 1999, and 2003, respectively. From 2003 to early 2005, he was a Post-Doc. at the Department of Biosystems, Korea Advanced Institute of Science and Technology. From 2005 to 2006, he was with the Language Technologies Institute, Carnegie Mellon University. Since early 2007, he has been an Assistant Professor at the Department of Electronic Engineering, Sogang University, Seoul, Korea. His current research interests include the theory and applications of independent component analysis, blind source separation, adaptive noise canceling, and binaural or multi-microphone processing for noise-robust speech recognition.