STATISTICAL CONFIDENCE MEASURE FOR DIRECTION-OF-ARRIVAL ESTIMATE

Mitsunori Mizumachi

Kyushu Institute of Technology
1-1 Sensui-cho, Tobata-ku, Kitakyushu, Fuuoka 804-8550, Japan
E-mail: mizumach@ecs.kyutech.ac.jp

ABSTRACT

This paper proposes to estimate the reliability of each direction-of-arrival (DOA) estimate explicitly. Observed spatial features, which give DOA estimates, are usually distorted by acoustical noise and reverberation. Then, distorted spatial features are filtered out by particle filters with a system model, which describes sound source dynamics on spatial state space. After particle filtering is carried out, set of the weighted particles can be considered as the discrete approximation of the true spatial feature. It is reasonable that particles with large weights concentrate around the true DOA on spatial state space. In this paper, crest factor, skewness, and kurtosis are calculated based on the second, third, and forth order moments from the particle set for representing how particles concentrate simply. It is confirmed that the skewness confidence measure is the most suitable for estimating DOA reliability under both directional and distributed noise conditions.

1. INTRODUCTION

Direction-of-arrival (DOA) plays an important role in multichannel signal processing [1]. In the field of acoustical signal processing, DOA contributes to achieve beamforming, sound source separation, dereverberation, and voice activity detection. DOA estimation has been a popular issue, but the almighty approach has not been established yet until now. Problems on background noise, reverberation, and multiple target sources make DOA estimation challenging in the real world. Observed spatial feature, which gives a DOA estimate, is distorted by interferences. Filtering the spatial feature is indispensable for achieving robust DOA estimation. Recently, particle filters are widely used for such purpose as flexible non-linear filters [2]. Particle filters can deal with non-linear models with non-Gaussian noises as system models and observation models. Particles are distributed time by time according to a system model, and their weights are updated by likelihood. We obtain better spatial features through state estimation by particle filtering than observed spatial features. Robust DOA estimation is achieved with particle filters even under noisy and reverberant environments [3]-[6].

In this paper, we focus on estimating reliability of each DOA estimate under noisy environments. Once we obtain a DOA estimate with its reliability, advanced signal processing can be applied for many applications. For example, in case of noise reduction, spatial filtering is not suitable when a DOA estimate is not reliable, and then spectral subtraction might be substituted for beamforming. It is also possible to understand complexity of acoustic scene and estimate scene change based on DOA reliability roughly. In this paper, discrete approximation of a spatial feature is prepared in the framework of particle filtering. Particles with weights are diverted from DOA estimation to reliability estimation. It is the key to measure how convergent the particles are around the true direction of a target sound source on spatial state space. Descriptive statistics, which are calculated from a set of weighted particles, are investigated as confidence measures for DOA estimates. Crest factor as the 2nd order statistics, skewness as the 3rd order statistics, and kurtosis as the 4th order statistics are candidates as the confidence measures. Feasibility of the higher-order statistical measures is examined for estimating the reliability of each DOA estimate under both diffused and directional noise conditions.

Finally, advanced DOA estimation is proposed using the higher-order statistical confidence measure. The author has previously proposed a robust DOA finder with a sequentially-updated noise model under noisy environments[7]. In this paper, the noise model is adaptively switched over depending on reliability of each DOA estimate automatically. Robustness of DOA estimation with the adaptive noise model is examined under time-variant noise conditions.

This study was partially supported by the Industrial Technology Research Grant Program from the New Energy and Industrial Technology Development Organization (NEDO) of Japan, and the Grant-in-Aid for Young Scientists (B) from the Ministry of Education, Science, Sports and Culture of Japan.

2. CONVENTIONAL DOA ESTIMATION

DOA of an acoustic signal can be estimated using stereo observations, $\mathbf{x}(t) \equiv (x_1(t), x_2(t))$, which are signals observed by spatially-separated two microphones.

$$x_1(t) = h_1(t) * s(t - \tau_1) + n_1(t), \tag{1}$$

$$x_2(t) = h_2(t) * s(t - \tau_2) + n_2(t), \tag{2}$$

where $h_1(t)$ and $h_2(t)$ are acoustical impulse responses between a target sound source and each of two microphones, $s(t)$ is the signal presented by the target sound source, and $n_1(t)$ and $n_2(t)$ are background noises, respectively. Time difference between τ_1 and τ_2 is the time difference of arrival (TDOA). TDOA estimate gives a DOA estimate straightforwardly on single-dimensional state space in [-90 degrees, $+90$ degrees].

Conventional DOA finders employ cross-correlation-based functions [8] or MUSIC spectra [9] as spatial features. In this paper, a cross-correlation function is employed as a non-parametric spatial feature, and the spatial feature can be regarded as a probability distribution for DOA existence. DOA estimate $\hat{\theta}$ is given as the direction with the maximum in the spatial feature $p(\theta|\mathbf{x})$.

$$\hat{\theta} = \arg\max_{\theta}\{p(\theta|\mathbf{x})\}. \quad (3)$$

Difficulty in DOA estimation is caused by distortion on the spatial feature $p(\theta|\mathbf{x})$ due to noise $n(t)$ and room reverberation $h(t)$. In this paper, we focus on the problem on noise.

3. DOA ESTIMATION WITH SOURCE DYNAMICS MODEL

Knowledge on dynamics of a target sound source contributes to obtain a robust spatial feature in adverse conditions. Provided that DOA estimation is performed frame by frame independently, it is hard to obtain accurate DOA estimates even under moderate noise conditions. Temporal trajectory of DOA can be modeled by a state space model, and it is estimated through state estimation procedure using particle filters. Markov modeling of temporal DOA trajectory contributes to yielding accurate and stable DOA estimates, assuming that the target sound source moves smoothly in between short-term frames. Here, random walk process is applied to model stochastic source movement as follows:

$$\theta_k = \theta_{k-1} + \nu_k,\ \nu_k \sim N(0, \sigma^2), \quad (4)$$

where θ_k represents the true DOA at the k-th time frame, and ν means the zero-mean Gaussian noise with the variance σ^2.

The true DOA trajectory and the sampled observations up to the k-th frame is noted as follows.

$$\theta_{1:k} = \{\theta_1, \theta_2, \cdots, \theta_k\}, \quad (5)$$
$$\mathbf{x}_{1:k} = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k\}. \quad (6)$$

Spatial feature is prepared from a cross correlation function, $r_{\mathbf{x}_k}(\tau)$, and is normalized by half-wave rectification. The half-wave rectified cross-correlation function ranges from 0 to 1. Now, the spatial feature, $p(\mathbf{x}_k|\theta_k)$, can be regarded as likelihood.

$$p(\mathbf{x}_k|\theta_k) \propto \max\{r_{\mathbf{x}_k}(\tau),\ 0\}. \quad (7)$$

State estimation is formally done in a recursive form of the posterior distribution, $p(\theta_{1:k}|\mathbf{x}_{1:k})$, as follows [2].

$$p(\theta_{1:k}|\mathbf{x}_{1:k}) \propto p(\theta_{1:k-1}|\mathbf{x}_{1:k-1})p(\mathbf{x}_k|\theta_k)p(\theta_k|\theta_{k-1}). \quad (8)$$

Sequential state estimation is carried out by particle filtering in the Bayesian framework [2]. We employ a bootstrap filter, which uses the system model as proposal [2]. DOA estimation is performed with a filtered spatial feature by particle filtering, where weighted particles are sequentially updated according to Eq. (8) as below.

STEP 0: (initial distribution)
To represent a distribution probability of the initial DOA θ_0, particles $\{z_0^{(l)}\}_{l=1}^{M}$, where $l = \{1, 2, \cdots, M\}$ means the particle index, are drawn in the single-dimentional spatial domain according to an uniform distribution. Initially, each particle has the same weight $w_0 = 1/M$.

STEP 1: (filtering by CC likelihood)
Particles at the k-th frame are drawn from the system model in Eq. (4), and weight for each particle is updated by the likelihood in Eq (7) as follows:

$$z_k^{(l)} \sim p(\theta_k|\theta_{k-1}^{(l)}), \quad (9)$$
$$w_k^{(l)} = p(\mathbf{x}_k|\theta_k^{(l)}), \quad (10)$$

where $l = \{1, 2, \cdots, M\}$.

STEP 2: (resampling)
The particles $\{z_k^{(l)}\}_{l=1}^{M}$ are sampled with replacement in proportion to the weight $\{w_k^{(l)}\}_{l=1}^{M}$. The resampled particles are used as the proposal particle distribution $\{z_{k+1}^{(l)}\}_{l=1}^{M}$ in the next (k+1)-th frame.

STEP 3: (finding DOA)
DOA is estimated by finding the peak of the filtered spatial feature, which is obtained from the filtered, weighted particles $\{z_k^{(l)}\}_{l=1}^{M}$ convolved with Gaussian kernels.

STEP 4:
go to STEP 1.

4. STATISTICAL CONFIDENCE MEASURES FOR DOA ESTIMATES

The author has previously proposed a confidence measure for estimating reliability of each DOA estimate [10]. It is important for estimating DOA reliability to measure how convergent the particles with large weights are around the true direction of a target sound source on the spatial state space. Effective sample size (ESS),

$$ESS_k = \frac{1}{\sum_{l=1}^{M}\{w_k^{(l)}\}^2}, \quad (11)$$

which had been originally proposed for deciding necessity of the resampling process in particle filtering [11], has been diverted to a DOA reliability measure. It represents how much particles are concentrated in the spatial domain. Feasibility of the ESS-based confidence measure has been confirmed under diffused noise conditions [10].

In this paper, availabilities of well-known higher-order descriptive statistics are also investigated for measuring reliabilities of DOA estimates. To measure the degree of convergence of weighted particles, crest factor as the 2nd order statistics, skewness as the 3rd order statistics, and kurtosis as the 4th order statistics are calculated from the weighted particles in Eq. (10). They are defined using the mean, $w_{\text{mean}} = \frac{1}{M}\sum_{l=1}^{M} w_k^{(l)}$, and the root mean square, $w_{\text{rms}} = \sqrt{\frac{1}{M}\sum_{l=1}^{M}\{w_k^{(l)}\}^2}$, of the weighted particles as follows.

Crest factor:

$$CrestFactor_k = \frac{w^{(l_{\max})}}{w_{\text{rms}}}, \quad (12)$$

where $l_{\max} = \underset{l}{\text{argmax}}\{w^{(l)}\}$ for $l = \{1, 2, \cdots, M\}$.

Skewness:

$$Skewness_k = \frac{1}{M}\sum_{l=1}^{M}\frac{\{w^{(l)} - w_{\text{mean}}\}^3}{w_{\text{rms}}^3}. \quad (13)$$

Kurtosis:

$$Kurtosis_k = \frac{1}{M}\sum_{l=1}^{M}\frac{\{w^{(l)} - w_{\text{mean}}\}^4}{w_{\text{rms}}^4}. \tag{14}$$

5. PERFORMANCE EVALUATION OF THE PROPOSED CONFIDENCE MEASURES

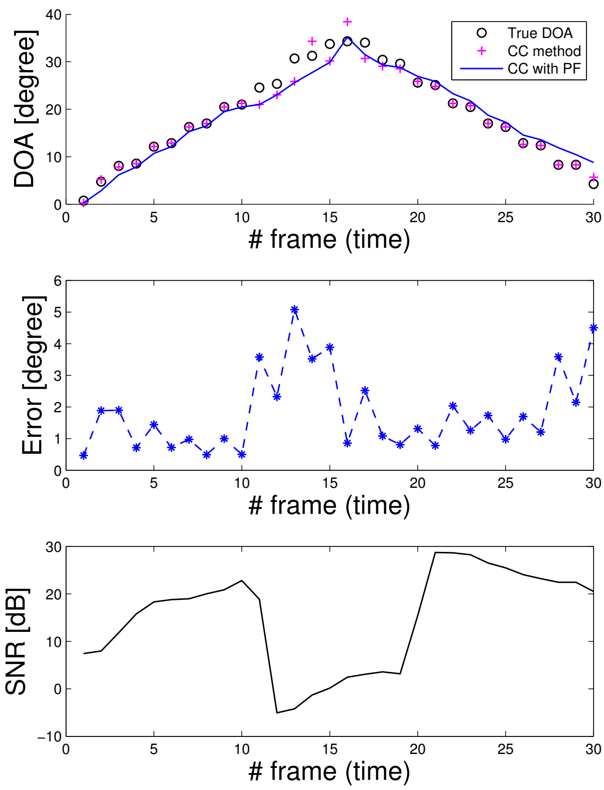

Fig. 1. True DOAs and two kinds of DOA estimates, errors in DOA estimation with particle filtering, and SNR in each frame are given in upper, middle, and lower panels under diffused noise condition.

Feasibility of the crest factor, skewness, kurtosis, and ESS are examined as DOA confidence measures. The target signal, which was a female speech from the TI-digit speech database, was re-recorded using a paired-microphone with the spacing of 0.10 m in a sound-proofed room. The target sound source moved smoothly and continuously. In this paper, feasibility of each confidence measure is investigated individually under either diffused or directional noise condition. The diffused noises were channel-independent, white Gaussian noises. In the directional noise condition, another white Gaussian noise was generated at the direction of -15 degrees. 500 particles were used for particle filtering, that is, $M = 500$. The system noise in Eq. (4) is optimized in advance using the above data.

Figures 1 and 3 show the true DOAs, DOA estimates by the conventional cross-correlation method in Section 2 and the robust method with the source dynamics model in Section 3 in the upper panels, errors in DOA estimation by the robust method in the middle panels, and signal-to-noise ratios in the lower panels, for diffused and directional noise conditions, respectively. Both noises appeared during from the 11th frame to the 20-th frame, and were added into the recorded target signal later in a computer. Figures 2 and 4 give ESS, crest factor, skewness, and kurtosis in each frame under diffused and directional noise conditions, respectively.

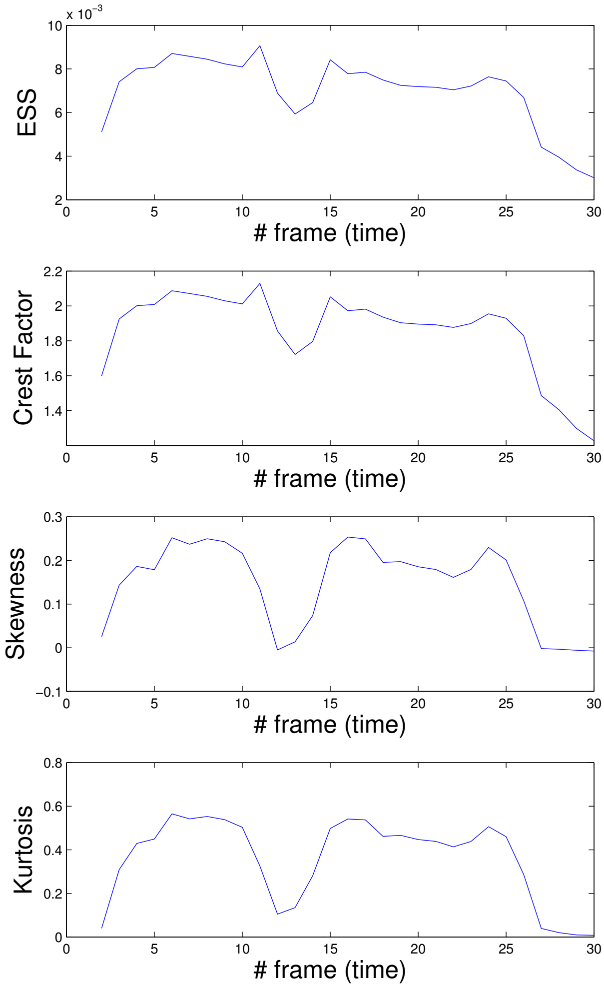

Fig. 2. ESS, crest factor, skewness, and kurtosis are given for each DOA estimate under diffused noise condition.

From Figs. 2 and 4, it is considered that the ESS confidence measure is not suitable under directional noise conditions. Behavior of the crest factor is similar to that of the ESS. The skewness value bears a close resemblance to the kurtosis value in both noise cases. Comparing the behavior at the 11th frame in the directional noise condition in Fig. 4, the skewness has slight advantage over the kurtosis as a DOA confidence measure. Those trends are preserved in real noise conditions: stationary car interior noise and non-stationary factory floor noise.

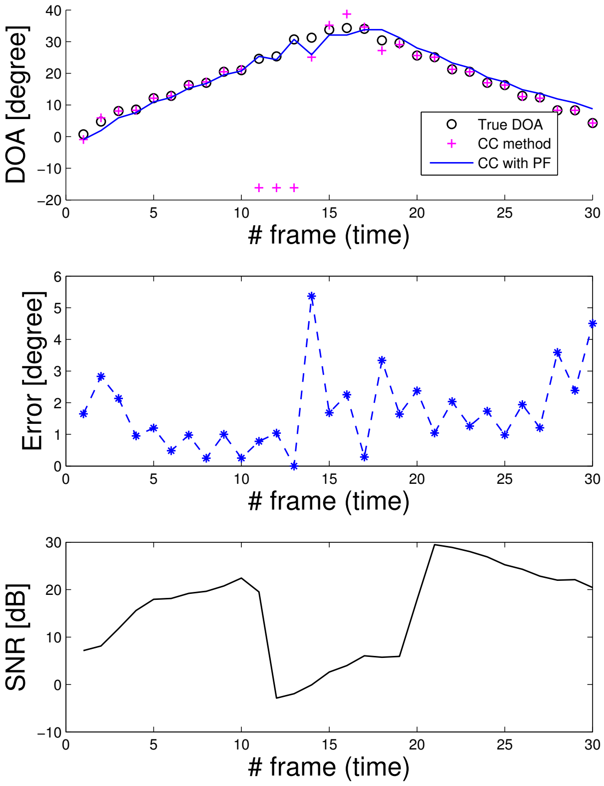

Fig. 3. True DOAs and two kinds of DOA estimates, errors in DOA estimation with particle filtering, and SNR in each frame are given in upper, middle, and lower panels under directional noise condition.

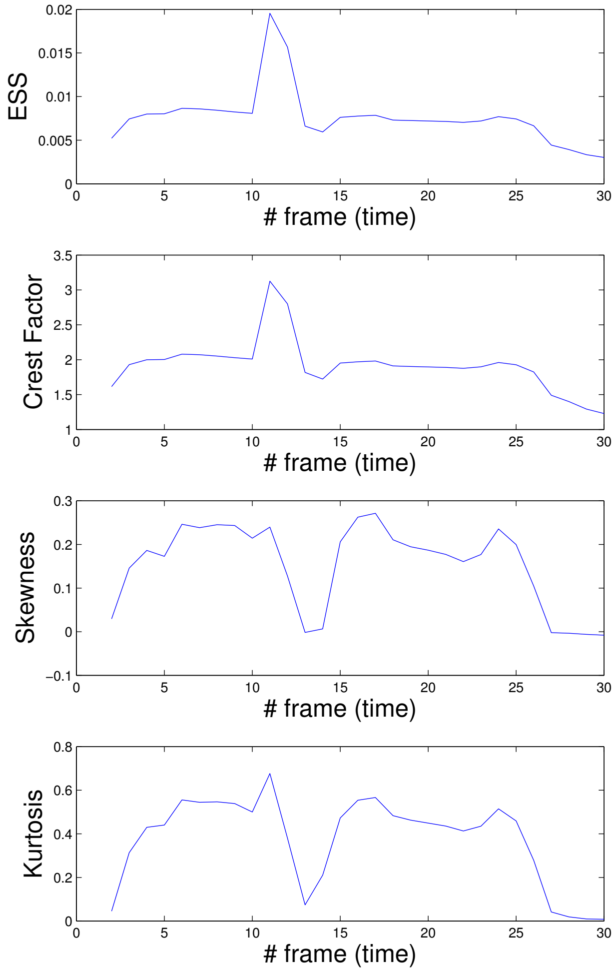

Fig. 4. ESS, crest factor, skewness, and kurtosis are given for each DOA estimate under directional noise condition.

We discuss why the confidence measure based on skewness of a spatial feature is the most suitable. First of all, the confidence measures based on the second order statistics do not work well under directional noise conditions as shown in Fig. 4. In the case of the kurtosis confidence measure, it evaluates the sharpness of the distribution of a spatial feature, and does not take the Gaussianity of the distribution into account. The skewness confidence measure can evaluate both the sharpness and the asymmetry of a spatial feature. Therefore, the skewness confidence measure is the most suitable for estimating the reliability of each DOA estimate theoretically.

6. APPLICATION TO ROBUST DOA ESTIMATION

Robust DOA estimation can be achieved relying on DOA reliability given by the skewness confidence measure, because the proposed method uses amplitude spectrum of background noise for estimating the dominant frequency of the target speech signal [7]. In the previous method [7], the noise model is updated compulsory by beamforming with a DOA estimate.

Noise model is adaptively updated based on both existence probability of a target signal and reliability of a DOA estimate as shown in Table 1. In other words, estimation of noise spectrum is performed by noise reduction. In the view point of noise reduction, either spectral subtraction or beamforming is properly employed based on reliability of a DOA estimate as shown in Table 2.

The adaptive noise reduction method is incorporated into DOA estimation. As the performance of the advanced DOA finder with the adaptive noise model, DOA estimates are given by a blue line in Fig. 5. In Fig. 5, the true DOA trajectory and DOA estimates by the previous method [7] are also plotted by black and red lines, respectively. The target speech signal is female utterances in the TI-digit speech database [12] . Noise scenarios is as follows. Factory floor noise [13] is presented in the whole section, and some sudden white noises appeared around 150th-250th and 450th-550th frames. It is obvious that the proposed adaptive noise model contributes to improve noise robustness in DOA estimation.

Signal	DOA reliability	**Noise model**
no	-	Update using observation
yes	low	Nothing to do
yes	high	Update by beamforming

Table 1. Adaptive scheme for updating noise model based on both signal existance probability and reliability of DOA estimate.

Signal	DOA reliability	**Noise reduction**
no	-	Nothing to do
yes	low	Spectral subtraction
yes	high	Beamforming

Table 2. Adaptive noise reduction for updating noise model.

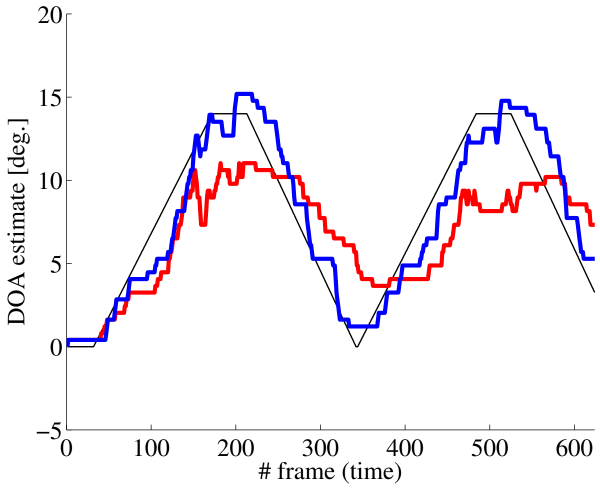

Fig. 5. DOA estimates with the proposed adaptive noise model in Table 1 and a static noise model are plotted in each time by blue and red lines, respectively. Black line shows a true DOA trajectory.

7. CONCLUSIONS

This paper proposed some kinds of confidence measures for estimating reliability of a DOA estimate under noisy environments. Degree of sharpness of the main-lobe in a spatial feature was represented by each of descriptive statistics, which are crest factor, skewness, kurtosis as the second, third, and fourth order moments, for weighted particles in particle filtering. Feasibility of the crest factor, skewness, kurtosis besides the effective sample size are investigated as confidence measure for measuring reliabilities of DOA estimates under diffused and directional noise conditions. The skewness confidence measure was the most suitable for estimating the reliability of each DOA estimate. It is also confirmed that the DOA confidence measure contributes to improve noise robustness in DOA estimation. In future, the proposed DOA confidence measures will be tested under various kinds of noise conditions.

REFERENCES

[1] M. S. Brandstein and D. B.Ward (eds.), *Microphone Arrays: Signal Processing Techniques and Applications*, Springer-Verlag, 2001.

[2] A. Doucet, J. F. G. de Freitas, and N. J. Gordon, *Sequential Monte Carlo Methods in Practice*, Springer-Verlag, New York, 2001.

[3] J. Vermaark and A. Blake, "Nonlinear filtering for speaker tracking in noisy and reverberant environments," Proc. Intl. Conf. on Acoust., Speech, and Signal Processing (ICASSP '01), Vol. 5, pp. 3021–3024, 2001.

[4] D. B. Ward, E. A. Lehmann, and R. C. Williamson, "Particle filtering algorithms for tracking an acoustic source in a reverberant environment," IEEE Trans. Speech, and Audio Process., Vol. 11, No. 6, pp. 826–836, 2003.

[5] M. Mizumachi and K. Niyada, "DOA estimation using cross-correlation with particle filter," Proc. Joint Workshop on Hands-Free Speech Communication and Microphone Arrays (HSCMA2005), CD-ROM, 2005.

[6] M. Mizumachi and K. Niyada, "DOA estimation based on cross-correlation by two-step particle filtering," Proc. 14th European Signal Process. Conf. (EUSIPCO2006), CD-ROM, 2006.

[7] M. Mizumachi and K. Niyada, "Synergic integration of noise reduction and DOA estimation with stochastic target sound source and deterministic noise models," Proc. Intl. Congress of Acoustics (ICA2007), CD-ROM, 2007.

[8] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," IEEE Trans. Acoust., Speech, Signal Process., Vol. 24, pp. 320–327, 1976.

[9] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," IEEE Trans. Antennas Propagation, Vol. 34, No. 3, pp. 276–280, 1986.

[10] M. Mizumachi and K. Niyada, "Robust direction-of-arrival estimation by particle filtering with confidence measure based on effective sample size under noisy environments," Proc. SCIS&ISIS 2008, CD-ROM, 2008.

[11] J. S. Liu and R. Chen, "Blind deconvolution via sequential imputations," J. Amer. Stat. Assoc., vol. 90, pp. 567-576, 1995.

[12] R. G. Leonard, " A database for speaker independent digit recognition, "Proc. Intl. Conf. on Acoust., Speech, and Signal Process. (ICASSP '84), vol. 9, pp. 328-331, 1984.

[13] A. Varga, H.J.M. Steeneken, " Assessment for automatic speech recognition: II. NOISEX-92: A database and an experiment to study the effect of additive noise on speech recognition systems, "Speech Comm., vol. 12, no. 3, pp. 247-252, 1993.