ROBUST TIME DELAY ESTIMATION EXPLOITING SPATIAL CORRELATION

Jingdong Chen, Jacob Benesty, and Yiteng (Arden) Huang

Bell Laboratories, Lucent Technologies
600 Mountain Avenue
Murray Hill, NJ 07974, USA
E-mail: {jingdong, jb, arden}@research.bell-labs.com

ABSTRACT

To find the position of an acoustic source in a room, a set of relative delays among different microphone pairs has to be determined. The generalized cross-correlation method is the most popular to do so and is well explained in a landmark paper by Knapp and Carter. In this paper, we show how we can take advantage of the redundancy when more than two microphones are available. It is believed that the redundancy will help to better cope with noise and reverberation. The idea of cross-correlation coefficient between two signals is generalized to the multichannel case by using the notion of spatial prediction. The multichannel spatial correlation matrix is then deduced and it is shown how it can be used for time delay estimation.

1. INTRODUCTION

Traditionally, time delay estimation (TDE), from measurements provided by an array of sensors, has played an important role in radar, sonar, and seismology for localizing radiating sources. Nowadays, with the increased development of communications among humans and human-machine interfaces, the need for localizing and tracking acoustic sources in a room has become essential. Two specific examples are automatic camera tracking for video-conferencing and microphone array beam steering for suppressing reverberation in all types of communication and voice processing systems. Thus, the time delay estimation-based locator has become the technique of choice in these applications, especially in recent digital systems.

The generalized cross-correlation (GCC) method, proposed by Knapp and Carter in 1976 [1], is the most popular technique for TDE. The delay estimate is obtained as the time-lag that maximizes the cross-correlation between filtered versions of the received signals. Since then, many new ideas have been proposed to deal better with noise and reverberation; see [2], [3], [4], [5], [6], [7], [8], [9]. In this paper, we develop some ideas around the spatial correlation matrix of multiple microphones and show how to apply this to TDE. As it will be shown, our approach is a generalization of the GCC to the multichannel (more than 2 microphones) case.

2. SIGNAL MODEL

Suppose that we have $L+1$ microphone signals $x_l[n]$, $l = 0, 1, ..., L$. Without loss of generality, we assume that the wave is in-phase at microphone 0. We consider the following propagation model:

$$x_l[n] = \alpha_l s[n - t - f_l(\tau)] + w_l[n], \qquad (1)$$

where α_l, $l = 0, 1, 2, ..., L$, are the attenuation factors due to propagation effects, t is the propagation time from the unknown source $s[n]$ to microphone 0, $w_l[n]$ is an additive noise signal at the lth microphone, τ is the relative delay between microphones 0 and 1, and $f_l(\tau)$ is the relative delay between microphones 0 and l. The function f_l depends of τ but also of the microphone array geometry. For example, in the far-field case (plane wave propagation), for a linear equispaced array, we have:

$$f_l(\tau) = l\tau, \qquad (2)$$

and for a linear non-equispaced array, we have:

$$f_l(\tau) = \frac{\sum_{i=0}^{l-1} d_i}{d_0}\tau, \qquad (3)$$

where d_i is the distance between microphones i and $i+1$, $i = 0, 1, 2, ..., L-1$. In the near-field case, f_l depends also on the position of the source. In general τ is not known, but the geometry of the antenna is known such that the exact mathematical relation of the relative delay between microphones 0 and l is well defined and given. It is further assumed that $s[n]$ and $w_l[n]$, $l = 0, 1, 2, ..., L$, are zero-mean, mutually uncorrelated, stationary Gaussian random processes.

3. SPATIAL PREDICTION AND INTERPOLATION

The notion of spatial prediction was presented in [10] but in the simple case that makes the spatial prediction equivalent to the classical linear prediction. In this section, we generalize this idea in a way that the geometry of the array is taken into account as well as the relative delay among the elements of this array. As a result, the spatial correlation matrix has a much more general form.

3.1. Linear Forward Spatial Prediction

Considering the microphone 0, we would like to align successive time samples of this microphone signal with spatial samples from the L other microphone signals. It is clear that $x_0[n - f_L(\tau)]$ is in-phase with the signals $x_l[n - f_L(\tau) + f_l(\tau)]$, $l = 1, 2, ..., L$. From these observations, we define the following forward spatial prediction error signal:

$$e_0[n - f_L(m)] = x_0[n - f_L(m)] - \mathbf{x}_{1:L}^T[n - f_L(m)]\mathbf{a}_m, \qquad (4)$$

where m is any guessed relative delay, superscript ^{T} denotes transpose of a vector or a matrix,

$$\mathbf{x}_{1:L}[n - f_L(m)] = [x_1[n - f_L(m) + f_1(m)] \cdots x_L[n]]^T$$

and

$$\mathbf{a}_m = \begin{bmatrix} a_{m,1} & a_{m,2} & \cdots & a_{m,L} \end{bmatrix}^T$$

is the linear forward spatial predictor. Consider the criterion

$$J_{m,0} = E\{e_0^2[n - f_L(m)]\}, \qquad (5)$$

where $E\{\cdot\}$ denotes mathematical expectation.

Minimization of (5) leads to the equation:

$$\mathbf{R}_{m,1:L}\mathbf{a}_m = \mathbf{r}_{m,1:L}, \qquad (6)$$

where

$$\mathbf{R}_{m,1:L} = E\{\mathbf{x}_{1:L}[n - f_L(m)]\mathbf{x}_{1:L}^T[n - f_L(m)]\}$$

$$= \begin{bmatrix} E\{x_1^2[n]\} & \cdots & E\{x_1[n-f_L]x_L[n-f_1]\} \\ \vdots & \vdots & \vdots \\ E\{x_L[n-f_1]x_1[n-f_L]\} & \cdots & E\{x_L^2[n]\} \end{bmatrix}$$

is the spatial correlation matrix, and

$$\begin{aligned}\mathbf{r}_{m,1:L} &= E\{\mathbf{x}_{1:L}[n - f_L(m)]x_0[n - f_L(m)]\} \\ &= \begin{bmatrix} E\{x_1[n - f_L(m) + f_1(m)]x_0[n - f_L(m)]\} \\ \vdots \\ E\{x_L[n]x_0[n - f_L(m)]\} \end{bmatrix} \\ &= \begin{bmatrix} E\{x_1[n]x_0[n - f_1(m)]\} \\ \vdots \\ E\{x_L[n]x_0[n - f_L(m)]\} \end{bmatrix}\end{aligned}$$

is the spatial correlation vector.

Note that the spatial correlation matrix is not Toeplitz in general, except for some particular cases.

For $m = \tau$ and for the noise free case where $w_l[n] = 0$, $l = 1, 2, ..., L$, it can easily be checked that with our signal model, the rank of matrix $\mathbf{R}_{\tau,1:L}$ is equal to 1. This means that the samples $x_0[n-\tau]$ can be perfectly predicted from any of one other microphone samples. However, the noise is never zero in practice and is in general isotropic. The energy of the different noises at the microphones will be added at the main diagonal of the correlation matrix $\mathbf{R}_{\tau,1:L}$, will regularize it, and this matrix will become positive definite (which we suppose in the rest of this paper). A unique solution to (6) is then guaranteed whatever the number of microphones. This solution is optimal from a Wiener theory point of view.

3.2. Linear Backward Spatial Prediction

Considering the microphone L, we would like to align successive time samples of this microphone signal with spatial samples from the L other microphone signals. It is clear that $x_L[n]$ is in-phase with the signals $x_l[n - f_L(\tau) + f_l(\tau)]$, $l = 0, 1, ..., L-1$. From these observations, we define the following backward spatial prediction error signal:

$$e_L[n - f_L(m)] \quad = \quad x_L[n] - \mathbf{x}_{0:L-1}^T[n - f_L(m)]\mathbf{b}_m, \quad (7)$$

where

$$\mathbf{x}_{0:L-1}[n - f_L(m)] = [x_0[n - f_L(m) + f_0(m)] \\ \cdots \; x_{L-1}[n - f_L(m) + f_{L-1}(m)]]^T$$

and

$$\mathbf{b}_m \quad = \quad [\; b_{m,1} \quad b_{m,2} \quad \cdots \quad b_{m,L} \;]^T$$

is the linear backward spatial predictor. Minimization of the criterion

$$J_{m,L} \quad = \quad E\{e_L^2[n - f_L(m)]\} \quad (8)$$

leads to the equation:

$$\mathbf{R}_{m,0:L-1}\mathbf{b}_m \quad = \quad \mathbf{r}_{m,0:L-1}, \quad (9)$$

where

$$\mathbf{R}_{m,0:L-1} = E\{\mathbf{x}_{0:L-1}[n - f_L(m)]\mathbf{x}_{0:L-1}^T[n - f_L(m)]\}$$

and

$$\mathbf{r}_{m,0:L-1} \quad = \quad E\{\mathbf{x}_{0:L-1}[n - f_L(m)]x_L[n]\}.$$

3.3. Linear Spatial Interpolation

The ideas presented for spatial prediction can easily be extended to spatial interpolation, where we consider any microphone element l, $l = 0, 1, 2, ..., L$. The spatial interpolation error signal is defined as

$$e_l[n - f_L(m)] \quad = \quad -\mathbf{x}_{0:L}^T[n - f_L(m)]\mathbf{c}_{m,l}, \quad (10)$$

where

$$\mathbf{x}_{0:L}[n - f_L(m)] = [x_0[n - f_L(m) + f_0(m)] \\ x_1[n - f_L(m) + f_1(m)] \cdots x_L[n]]^T$$

and

$$\mathbf{c}_{m,l} \quad = \quad [\; c_{m,l,0} \quad c_{m,l,1} \quad \cdots \quad c_{m,l,L} \;]^T$$

with $c_{m,l,l} = -1$, is the spatial interpolator. The criterion associated with (10) is:

$$J_{m,l} \quad = \quad E\{e_l^2[n - f_L(m)]\}. \quad (11)$$

The rest flows immediately from the previous sections on prediction.

4. APPLICATION TO TIME DELAY ESTIMATION

In this section, we only use the forward spatial prediction idea but of course spatial interpolation can also be used. So we consider the minimization of criterion $J_{m,0}$ for different m.

Let $J_{m,0;\min}$ denote the minimum mean-squared error, for the value m, defined by

$$J_{m,0;\min} \quad = \quad E\{e_{0;\min}^2[n - f_L(m)]\}. \quad (12)$$

If we replace $\mathbf{a}_m$ by $\mathbf{R}_{m,1:L}^{-1}\mathbf{r}_{m,1:L}$ in (4), we get:

$$e_{0;\min}[n - f_L(m)] = x_0[n - f_L(m)] - \\ \mathbf{x}_{1:L}^T[n - f_L(m)]\mathbf{R}_{m,1:L}^{-1}\mathbf{r}_{m,1:L}. \quad (13)$$

We deduce that:

$$J_{m,0;\min} = E\{x_0^2[n - f_L(m)]\} - \mathbf{r}_{m,1:L}^T\mathbf{R}_{m,1:L}^{-1}\mathbf{r}_{m,1:L}. \quad (14)$$

The value of m that gives the minimum $J_{m,0;\min}$, for different m, corresponds to the time delay between microphone 0 and 1. Mathematically, the solution to our problem is then given by

$$\hat{\tau} \quad = \quad \arg\min_m J_{m,0;\min}, \quad (15)$$

where $\hat{\tau}$ is an estimate of τ.

Particular case: Two microphones ($L = 1$). In this case, the solution is:

$$\begin{aligned} \hat{\tau} &= \arg\min_m \left\{ E\{x_0^2[n]\} \left[1 - \frac{E^2\{x_0[n-m]x_1[n]\}}{E\{x_0^2[n]\}E\{x_1^2[n]\}}\right] \right\} \\ &= \arg\min_m \left\{1 - \rho_{m,01}^2\right\} \\ &= \arg\max_m \left(\rho_{m,01}^2\right), \end{aligned} \quad (16)$$

where $\rho_{m,01}$ ($\rho_{m,01}^2 \leq 1$) is the cross-correlation coefficient between $x_0[n-m]$ and $x_1[m]$. When the cross-correlation coefficient is close to 1, this means that the two signals that we compare are highly correlated which happens when the signals are in-phase, i.e. $m \approx \tau$ and this implies that $J_{\tau,0;\min} \approx 0$. This approach is similar to the generalized cross-correlation method proposed by Knapp and Carter [1]. Note that in the general case with any number of microphones, the proposed approach can be seen as a cross-correlation method, but we take advantage of the knowledge of the microphone array to estimate only one time delay (instead of estimating multiple time delays independently) in an optimal way in a least mean square sense.

5. OTHER INFORMATION FROM THE SPATIAL CORRELATION MATRIX

Consider the $L+1$ microphone signals x_l, $l = 0, 1, ..., L$, the corresponding spatial correlation matrix is:

$$\begin{aligned} \mathbf{R}_{m,0:L} &= \mathbf{R}_m \\ &= E\{\mathbf{x}_{0:L}[n - f_L(m)]\mathbf{x}_{0:L}^T[n - f_L(m)]\}. \end{aligned} \quad (17)$$

It can be shown that $\mathbf{R}_m$ can be factored as:

$$\mathbf{R}_m \quad = \quad \mathbf{D}\widetilde{\mathbf{R}}_m\mathbf{D}, \quad (18)$$

where

$$\mathbf{D} = \begin{bmatrix} \sqrt{E\{x_0^2[n]\}} & 0 & \cdots & 0 \\ 0 & \sqrt{E\{x_1^2[n]\}} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \sqrt{E\{x_L^2[n]\}} \end{bmatrix} \quad (19)$$

is a diagonal matrix,

$$\widetilde{\mathbf{R}}_m = \begin{bmatrix} 1 & \rho_{m,01} & \cdots & \rho_{m,0L} \\ \rho_{m,01} & 1 & \cdots & \rho_{m,1L} \\ \vdots & \ddots & \ddots & \vdots \\ \rho_{m,0L} & \cdots & \rho_{m,L-1L} & 1 \end{bmatrix} \quad (20)$$

is a symmetric matrix, and

$$\rho_{m,kl} = \frac{E\{x_k[n - f_l(m)]x_l[n - f_k(m)]\}}{\sqrt{E\{x_k^2[n]\}E\{x_l^2[n]\}}}, \quad k, l = 0, 1, ..., L, \tag{21}$$

is the cross-correlation coefficient between $x_k[n - f_l(m)]$ and $x_l[n - f_k(m)]$.

We now give two propositions that will be useful for TDE.

Proposition 1. We have:

$$0 < \det\left(\widetilde{\mathbf{R}}_m\right) \leq 1, \tag{22}$$

where "det" stands for *determinant*.

Proof. This proposition can be shown by induction, *i.e.*,

$$\det\left(\widetilde{\mathbf{R}}_m\right) \leq \det\left(\widetilde{\mathbf{R}}_{m,1:L}\right) \leq \cdots \leq 1. \tag{23}$$

Proposition 2. We have:

$$\det\left(\widetilde{\mathbf{R}}_m\right) \leq \frac{J_{m,0;\min}}{E\{x_0^2[n]\}} \leq 1. \tag{24}$$

Proof. It can be shown, by using the Lagrange multiplier, that:

$$J_{m,0;\min} = \frac{1}{\boldsymbol{\delta}^T \mathbf{R}_m^{-1} \boldsymbol{\delta}}, \tag{25}$$

where $\boldsymbol{\delta} = [1\ 0\ \cdots\ 0]^T$. In this case, using (18), (25) becomes:

$$\begin{aligned} J_{m,0;\min} &= \frac{E\{x_0^2[n]\}}{\boldsymbol{\delta}^T \widetilde{\mathbf{R}}_m^{-1} \boldsymbol{\delta}} \\ &= E\{x_0^2[n]\} \frac{\det\left(\widetilde{\mathbf{R}}_m\right)}{\det\left(\widetilde{\mathbf{R}}_{m,1:L}\right)}. \end{aligned} \tag{26}$$

Using (23), it is clear that proposition 2 is verified.

In the general case, for any interpolator, we have:

$$\det\left(\widetilde{\mathbf{R}}_m\right) \leq \frac{J_{m,l;\min}}{E\{x_l^2[n]\}} \leq 1,\ l = 0, 1, ..., L. \tag{27}$$

As we can see, the determinant of the spatial correlation matrix is related to the minimum mean-squared error and to the correlation of the signals. Let's take the two-channel case. It is obvious that the cross-correlation coefficient between the two signals x_0 and x_1 is linked to the determinant of the corresponding spatial correlation matrix:

$$\rho_{m,01}^2 = 1 - \det\left(\widetilde{\mathbf{R}}_{m,0:1}\right). \tag{28}$$

By analogy to the cross-correlation coefficient definition between two signals, we define the multichannel correlation coefficient among the signals x_l, $l = 0, 1, ..., L$, as:

$$\rho_{m,0:L}^2 = 1 - \det\left(\widetilde{\mathbf{R}}_{m,0:L}\right). \tag{29}$$

From proposition 2, we give a new bound for $\rho_{m,0:L}^2$:

$$1 - \frac{J_{m,0;\min}}{E\{x_0^2[n]\}} \leq \rho_{m,0:L}^2 \leq 1. \tag{30}$$

Basically, the coefficient $\rho_{m,0:L}$ will measure the amount of correlation among all the channels. This coefficient has some interesting properties. For example, if one of the signals, say x_0, is completely decorrelated from the others because the microphone is defective, or it picks up only noise, or the signal is saturated, this signal will not affect $\rho_{m,0:L}$ since $\rho_{m,0l} = 0$, $\forall l$. In this case:

$$\rho_{m,0:L}^2 = \rho_{m,1:L}^2. \tag{31}$$

In other words, the measure "drops" the signals who have no correlation with the others. This makes sense from a correlation point of view, since we want to measure the degree of correlation only from the channels who have something in common. In the extreme cases where all the signals are uncorrelated, we have $\rho_{m,0:L}^2 = 0$, and where any two signals (or more) are perfectly correlated, we have $\rho_{m,0:L}^2 = 1$.

Obviously, the multichannel coefficient $\rho_{m,0:L}^2$ can be used for time delay estimation in the following way:

$$\begin{aligned} \hat{\tau} &= \arg\max_m \left(\rho_{m,0:L}^2\right) \\ &= \arg\min_m \left[\det\left(\widetilde{\mathbf{R}}_{m,0:L}\right)\right]. \end{aligned} \tag{32}$$

This method can be seen as a multichannel correlation approach for the estimation of time delay and it is clear that (32) is equivalent to (15).

6. PERFORMANCE EVALUATION

6.1. Experiment Setup

Experiments were carried out in the Varechoic Chamber which is a unique facility at Bell Laboratories. The chamber is a 6.7 × 6.1 × 2.9 m room whose surfaces are covered by a total of 369 active panels which can be controlled digitally. Each panel consists of two perforated sheets. When the holes in the sheets are aligned, absorbing material behind the sheets will be exposed to the sound field, whereas a highly reflective surface can be formed if the holes are shifted to misalignment. Combination of open and closed panels can produce 2^{369} different acoustic environments where the 60-dB reverberation time T_{60} can change from 0.2 to almost 1 second. See [11] for more details.

A linear microphone array which consists of 22 omnidirectional Panasonic WM-61A microphones was mounted at the distance of 0.5 m from the north wall of the chamber and approximately at the center of the wall. The 22 microphones are uniformly distributed along an aluminum rod whose diameter is 1 cm. The spacing between adjacent microphones is 10 cm. The source signal is played by a Cabasse Baltic Murale loudspeaker in 46 different positions. An illustration of this setup is shown in Fig. 1.

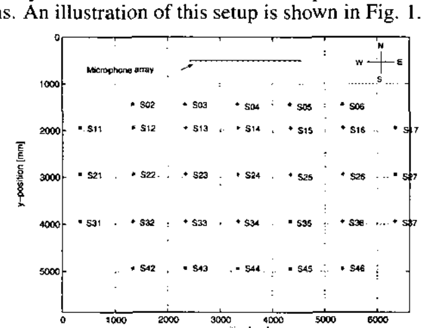

Fig. 1. Layout of the microphone array and source positions in the Varechoic Chamber.

For the purpose of data reusability, the impulse response from each source location to each microphone was measured [12]. The observed signal is then obtained by convolution of a recorded speech signal with the measured impulse responses. The measurement of the impulse responses were performed using the built-in measurement tool of the Huron Lake system. A 65536-point long logarithmic sweep signal digitized at a sampling rate of 48 kHz was used as the excitation signal. From each source location to each microphone, the excitation is played and recorded. An estimate of the transfer function is obtained by spectral division between the original source excitation and the recorded microphone signal.

6.2. Performance Criteria

Following [13], [14], we distinguish an estimate as either an *anomaly* or a *nonanomaly* according to its absolute error. If the absolute error $|\hat{\tau}_i - \tau_i| > T_c/2$, the estimate is identified as an anomaly; otherwise it is declared as a nonanomaly, where T_c is the signal correlation time. In our experiment, T_c is computed as the 3 dB width of the main lobe of the source signal autocorrelation function. The TDE performance is evaluated in terms of the percentage of anomalous estimates over the total estimates, and the bias and standard deviation of the nonanomalous estimates.

6.3. TDE performance versus the Number of Microphones

Several experiments were conducted to study the TDE performance of the proposed approach in different reverberation and noise conditions. For brevity, we report one set of experimental results here. The sound source is in S31. 89% of the 369 panels are open and reverberation time T_{60} is approximately 0.24 s (moderate reverberation). The observed signal is obtained by convolution of 4-minute speech from a female speaker with the measured impulse response. Computer-generated white Gaussian noise is then added to the signal to control the signal-to-noise ratio (SNR) to be 0 dB. The signal sequence is segmented into non-overlapping frames with a frame width of 128 ms. A short-time energy based voice activity detector (ACT) is applied to the signal at microphone 0 to distinguish each frame as speech or noise-only. For each speech frame, a time delay is obtained by estimator described in (32).

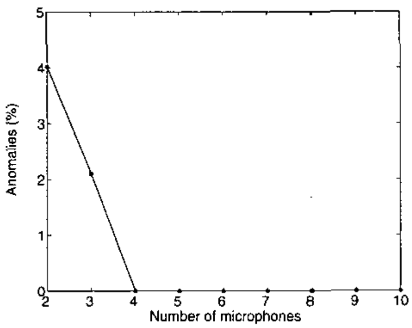

Fig. 2. Percentage of anomalous time delay estimates versus number of microphones.

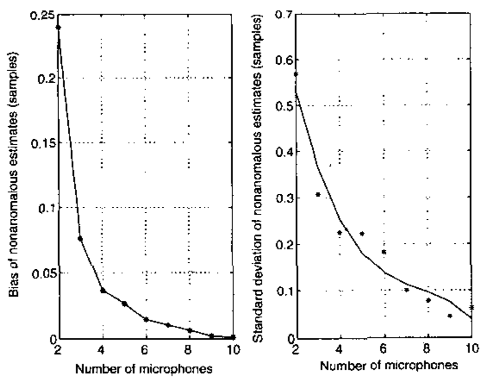

Fig. 3. Bias and standard deviation of nonanomalous time delay estimates versus number of microphones.

Figure 2 and 3 plot the percentage of anomalous time delay estimates, the bias and standard deviation of the nonanomalous estimates, all as a function of the number of microphones. It can be seen from Fig. 2 that the percentage of anomalous estimates decreases as more microphones are employed. For two microphones, the anomalies are approximately 4% over the total estimates. When more than four microphones are used, no anomalous estimates are observed. From Fig. 3, one can see that both the bias and standard deviation of the nonanomalous estimates reduces as the number of microphones is increased. For two microphones, the bias of the nonomalous is approximately 0.24 samples, while this bias reduces to almost 0 when ten microphones are used.

7. CONCLUSIONS

The spatial correlation matrix can be written in different ways. We have proposed a way which has included some a priori information of the microphone array geometry and the relation among the different time delays. Given the relative delay, τ, between microphones 0 and 1, we have supposed that the relative delay between microphones 0 and l is a function of τ. Thus, if τ is known, any microphone signal can be predicted from the others. This can be useful for multichannel coding. If τ is not known, it can be estimated by minimizing the spatial prediction error or, equivalently, by using the determinant of the spatial correlation matrix where all the redundancy is taken into account. Experimental results verified that this redundancy can make the estimation of τ more robust with respect to noise and reverberation.

ACKNOWLEDGMENTS

The authors would like to thank Dr. G. W. Elko for very helpful discussions during the development of this work. They are also grateful to Dr. A. Harma for providing the measured impulse responses.

8. REFERENCES

[1] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 24, pp. 320–327, Aug. 1976.

[2] D. R. Fischell and C. H. Coker, "A speech direction finder," in *Proc. IEEE ICASSP*, 1984, pp. 19.8.1–19.8.4.

[3] M. Omologo and P. Svaizer, "Acoustic event localization using a crosspower-spectrum phase based technique," in *Proc. IEEE ICASSP*, 1994, pp. II-273–II-276.

[4] M. Omologo and P. Svaizer, "Acoustic source location in noisy and reverberant environment using CSP analysis," in *Proc. IEEE ICASSP*, 1996, pp. 921–924.

[5] M. S. Brandstein, "A pitch-based approach to time-delay estimation of reverberant speech," in *Proc. IEEE ASSP Workshop Appls. Signal Processing Audio Acoustics*, 1997.

[6] H. Wang and P. Chu, "Voice source localization for automatic camera pointing system in videoconferencing," in *Proc. IEEE ASSP Workshop Appls. Signal Processing Audio Acoustics*, 1997.

[7] A. Stéphenne and B. Champagne, "A new cepstral prefiltering technique for time delay estimation under reverberant conditions," *Signal Processing*, vol. 59, pp. 253–266, 1997.

[8] J. Benesty, "Adaptive eigenvalue decomposition algorithm for passive acoustic source localization," *J. Acoust. Soc. Am.*, vol. 107, pp. 384–391, Jan. 2000.

[9] Y. Huang, J. Benesty, and G. W. Elko, "Microphone arrays for video camera steering," in *Acoustic Signal Processing for Telecommunication*, S. L. Gay and J. Benesty, eds., Kluwer Academic Publishers, chap. 11, pp. 239–259, 2000.

[10] S. Haykin, "Radar array processing for angle of arrival estimation," in *Array Signal Processing*, S. Haykin, ed., Prentice-Hall, chap. 4, pp. 194–292, 1985.

[11] W. C. Ward, G. W. Elko, R. A. Kubli, and W. C. McDougald, "The new varechoic chamber at AT&T Bell Labs," in *Proc. Wallance Clement Sabine Centennial Symposium*, 1994, pp. 343-346.

[12] A. Harma, "Acoustic measurement data from the varechoic chamber," *Technical Memorandum*, No. 110101, Agere Systems, 2001.

[13] J. P. Ianniello, "Time delay estimation via cross-correlation in the presence of large estimation errors, " *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, pp. 998-1003, Dec. 1982.

[14] B. Champagne, S. Bedard, and A. Stepheme, "Performance of time-delay estimation in presence of room reverberation," *IEEE Trans. Speech and Audio Processing*, vol. 4, pp. 148-152, Mar. 1996.