In Commemoration of the ...th Anniversary of ASK
KOREA-JAPAN JOINT SYMPOSIUM ON ACOUSTICS
"The Vision of Acoustics and Speech Communication toward the 21st Century"

Detection of Coherent Sound Source with Noise by Using Array Microphones

Jae Woong Choi, and Yang-Hann Kim
崔 在雄　　金 榮漢

Noise and Vibration Control Group
Dept. of Mechanical Engineering,
Korea Advanced Institute of Science and Technology

373-1 Guseong-dong, Yuseong-gu, Taejon-shi, Korea

1. INTRODUCTION

For simple case, a location of a sound source can be estimated using two sensors based on trigonometric relation between the source and microphones. It is therefore, quite natural to extend this basic understanding to estimate the locations, or in more general term, the information about the multiple sources. This can be done by employing multiple microphones as expected. The idea and concept of using as many microphones as required is straightforward but it does have interesting problems with respect to signal processing methods. First, they have a space aliasing problem due to the distance between microphones. Second, there exists only some small number of data in space due to the limitation of total number of sensors in space domain. This limits the spatial resolution. Third, there can be coherent sound sources and/or multipath of a signal. This prohibits the accurate detection of sound sources.

As far as the second problem is concerned, MEM(maximum entrophy method), ARMA(autoregressive moving average) methods[1] have been recognized as the most effective method, but there is no unique method to determine the number of poles and zeros, therefore, they are limited in practical applications.

By 1960, the classical beam forming technique[2] was introduced and Capon's minimum variance estimator[3] and linear prediction method[4] were followed. But these methods are highly sensitive to noise.

The MUSIC(MUltiple SIgnal Classification) algorithm was introduced by Schmidt[5] late 1970. It is proven to be very effective to estimate the bearing angle when the signal has the white noise in space domain. This algorithm utilizes the eigenstructure analysis of cross correlation matrix of the signals received by array microphones.

In this paper, this algorithm is introduced briefly and extended to the coherent source cases based on the spatial smoothing technique[6]. The spatial smoothing technique is derived in terms of doppler shift in space domain. To demonstrate the algorithm, various simulations and experiments were performed.

2. BASIC CONCEPTS OF BEARING ANGLE ESTIMATION

When the sound sources are located in the far field compared to wave length, the sound waves which are measured by array microphones can be regarded as plane waves(Fig. 1). It can be also assumed that the signals and noises are stationary, zero mean Gaussian random process, and further, the noises could be assumed to be independent and identical between themselves with common variance . Therefore, the i-th microphone have the signals,

$$x_i(t) = \sum_{k=1}^{K} u_k(t)\, e^{-j2\pi f \tilde{d} \cos\theta_k / c} + n_i(t) \tag{1}$$

where K is the total number of sound sources, $\tilde{d}$ is the distance of microphone from the reference point, $n_i(t)$ is noise signal received at i-th microphone, $u_k(t)$ is sound emitted from source, and $e^{-j2\pi f \tilde{d} \cos\theta_k / c}$ represents the plane wave propagation in space. For the all microphones, this becomes

$$x(t) = \begin{Bmatrix} x_1(t) \\ \vdots \\ x_M(t) \end{Bmatrix} = \begin{bmatrix} e^{-j2\pi f \tilde{d}_1 \cos\theta_1 / c} & \cdots & e^{-j2\pi f \tilde{d}_1 \cos\theta_K / c} \\ \vdots & \vdots & \vdots \\ e^{-j2\pi f \tilde{d}_M \cos\theta_1 / c} & \cdots & e^{-j2\pi f \tilde{d}_M \cos\theta_K / c} \end{bmatrix} \begin{Bmatrix} u_1(t) \\ \vdots \\ u_K(t) \end{Bmatrix} \tag{2}$$

The sound power in terms of bearing angle can be calculated as

$$P(\theta) = E[|y(t)|^2] = w_B^{+} E[xx^{+}] w_B \tag{3}$$

$$\text{where} \quad y(t) = \sum_{i=1}^{M} w_i^{*}(\theta) x_i(t) = w_B^{+} x \tag{4}$$

$$w_B^{+} = \frac{1}{\sqrt{M}} \left[1,\ e^{-j\omega d \cos\theta / c}, \ldots, e^{-j\omega d \cos\theta / c} \right]^T \tag{5}$$

That is, the sound power is obtained by weighting of looking direction w_B^{+}. Therefore, the looking direction function can be written as

$$F(\theta) = \sum_{i=1}^{M} e^{j2\pi d(i-1)\cos\theta / \lambda} \tag{6}$$

and looking direction gain $G(\theta)$ is

$$G(\theta) = \left| \frac{1}{M} F(\theta) \right|^2 = \left| \frac{\sin(\pi \hat{R} M \cos\theta)}{M \sin(\pi d \cos\theta)} \right|^2 \tag{7}$$

$$\text{where} \quad \hat{R} = \frac{d}{\lambda} = \frac{\text{microphone spacing}}{\text{wave length}} \tag{8}$$

Eq. (7) says that the ratio of microphone spacing and wave length is important in characterizing the looking direction gain $G(\theta)$. To see the details, several simulations are performed. Fig. 2, Fig. 3 and Fig. 4 show increasing effect of $\hat{R}$. But for the case of larger than 1/2, i.e., Fig. 4, it has two peak directions within π radian. Hence, to avoid this spatial aliasing, the ratio $\hat{R}$ should be less than 1/2. Also if the number of microphones are increased, then the resolution of bearing angle is increased (Figs. 3 and 5).

3. EIGENSTRUCTURE ANALYSIS

To obtain the high resolution in looking direction, eigenstructure analysis technique is introduced. To get the eigenstruture model from the typical array, a correlation matrix R is introduced as

$$x(t) = Au(t) + n(t) \tag{9}$$

$$\text{where} \quad A = \sqrt{M}[a(\theta_1), a(\theta_2), \cdots, a(\theta_K)]$$

$$u(t) = [u_1, u_2, \cdots, u_K]^T, \; n(t) = [n_1, n_2, \cdots, n_M]^T$$

$$a(\theta_k) = \left[1, \; e^{-j\omega d \cos\theta_k / c}, \cdots, e^{-j\omega d(M-1)\cos\theta_k / c} \right]^T$$

As mentioned in section 2, white noise in space domain, no correlation between true signal and noise, and independent noises between microphones are assumed. Then,

$$R = A\, R_u A^+ + \sigma^2 I \tag{10}$$

$$\text{where} \quad R_u = E[uu^+], \; \sigma^2 = E[nn^+]$$

The eigenvalues and eigenvectors of this matrix are arranged in accordance with decreasing order of eigenvalue. This is true when the signal to noise ratio is sufficiently high.

$$\lambda_k = \begin{cases} \mu_k + \sigma^2 & k=1,2,\cdots,K \\ \sigma^2 & k=K+1,\cdots,M \end{cases} \tag{11}$$

$$B = [\beta_1, \beta_2, \cdots \beta_K, \beta_{K+1}, \cdots \beta_M]^T$$

Then the eigenvalues and eigenvectors can be written as

$$\lambda_1 \geq \lambda_1 \geq \cdots \geq \lambda_K \geq \sigma^2 \tag{12}$$

$$\beta_1, \; \beta_2, \cdots, \beta_K, \; \beta_{K+1}, \cdots, \beta_M$$

$$u = [u_1, u_2, \cdots, u_K]^T$$

To apply eigenstructure analysis to Eq. 14, let's introduce the correlation matrix in terms of the grouped signal set.

$$R^f = \frac{1}{L}\sum_{l=1}^{L} R_l^f = AR_u^f A^* \tag{14}$$

where $$R_u^f = \frac{1}{L}\sum_{l=1}^{L} B^{l-1} R_u (B^{l-1})^*$$

This is called as forward spatial smoothing technique. If the direction of grouping is backward, then it is called as backward spatial smoothing technique and gives same result as forward spatial smoothing case. In both cases, the number of microphones to make the same dimension of correlationre matrix R is twice of those for the classical technique. To reduce the number of necessary microphones, forward and backward spatial smoothing techniques could be combined.
Table 1 is the relation between number of coherent sources and that of microphones.

Table 1 Relation between number of coherent sources and that of microphones

	Coherent source	Microphone	Grouping	smoothing
Forward(or backward) spatial smoothing	K	$M_0 (\geq 2K)$	N ($\geq$ K+1)	L ($\geq$ K)
Forward/backward spatial smoothing	K	$M_0 (\geq \frac{3}{2}K)$	N ($\geq$ K+1)	2L ($\geq$ K)

5. EXPERIMENTS AND RESULTS

To verify and understand the performance of the eigenstructure analysis, various simulations and experiments were performed. Anechoic room whose cutoff frequency is 80Hz is used for the experiments.

A distance of sound sources and microphones were enough to confirm the far field condition (ka >>1). A sound sources with narrow band width were used.

To get the signal concerning the time delay, cross-spectrum to the reference microphone is used. The other method to get this information is the transfer function method to the reference microphone. In this study, both techniques were tested.

Figure 7 shows the experimental result by MUSIC algorithm and shows the good agreement with the true value which was measured by rulers and calculated by simple trigonometric relation.

In case of coherent sources, we derived two speakers with same signal generator and measured the sound pressure by using aforementioned array microphones. Figure 8 shows the result by forward/backward spatial smoothing technique and the good agreement with true bearing angles.

6. CONCLUSION AND RECOMMENDATION

In this study, the signal processing method for array microphones is analyzed and the basic techniques to the detection of coherent sound sources are derived based on the doppler shift effect in space domain. The performances are tested to the simple two coherent sources case in anechoic room.

The eigenstructure analysis with spatial smoothing technique is proved to be very useful technique to detect the coherent sources.

REFERENCES

1. S. L. Marple Jr., *Digital Spectral Analysis*, PRENTICE-HALL. Inc., 1987, Chap.6-10

2. M. S Bartlett, 'Periodogram analysis and continous spectra' , Biometrica, Vol. 37, pp1-6, 1950

3. J. Capon, 'High Resolution Frequency-wave number Spectrum Analysis', Proc. IEEE, Vol. 57, pp1408-1418, Aug. 1969

4. S.W. Lang and J.H. McClellan, 'Frequency Estimation with Maximum Entropy Spectral Estimators', IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-28, pp716-724, Dec. 1980

5. R.O. Schmidt, 'MUltiple Emitter Location and Signal Parameter Estimation', IEEE Trans. Antenna and Propagation, Vol. AP-34. No. 3, March 1986, pp276-280

6. S.U. Pillai, *Array Signal Processing*, Springer-verlag, 1988, Chap. 2

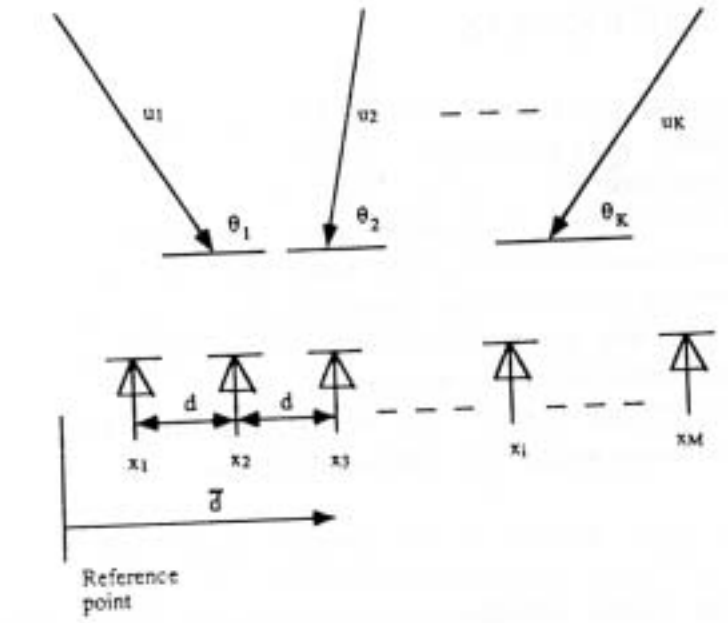

Fig. 1 A typical microphone array scene

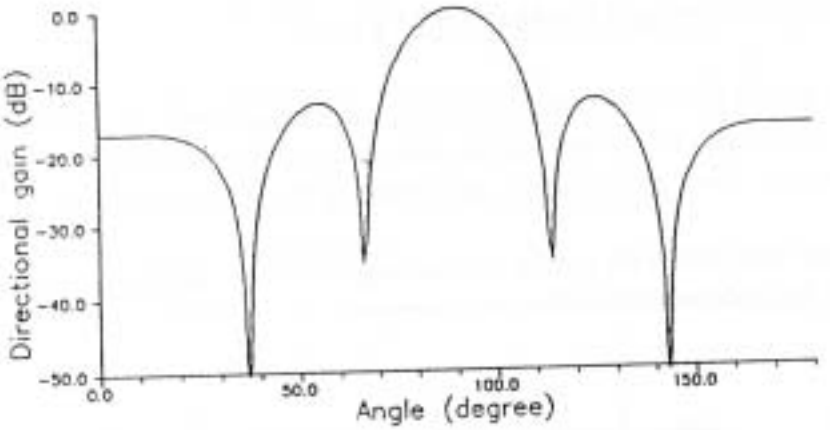

Fig. 2 Directional gain pattern
(10 microphones, $\frac{1}{4}\lambda$ spacing)

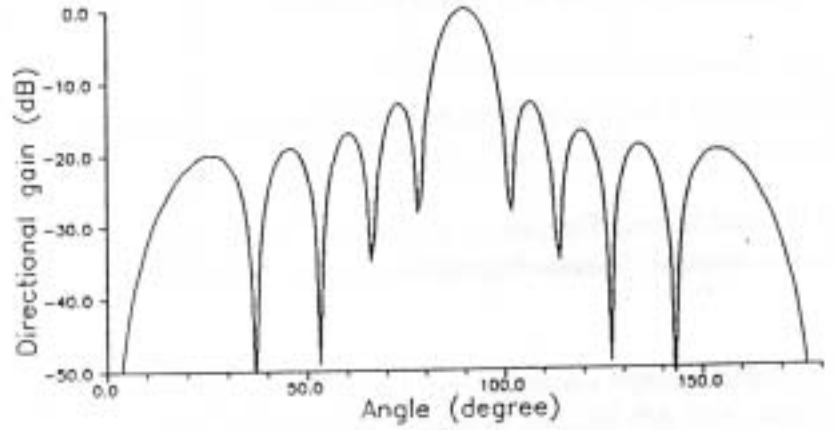

Fig. 3 Directional gain pattern
(10 microphones, $\frac{1}{2}\lambda$ spacing)

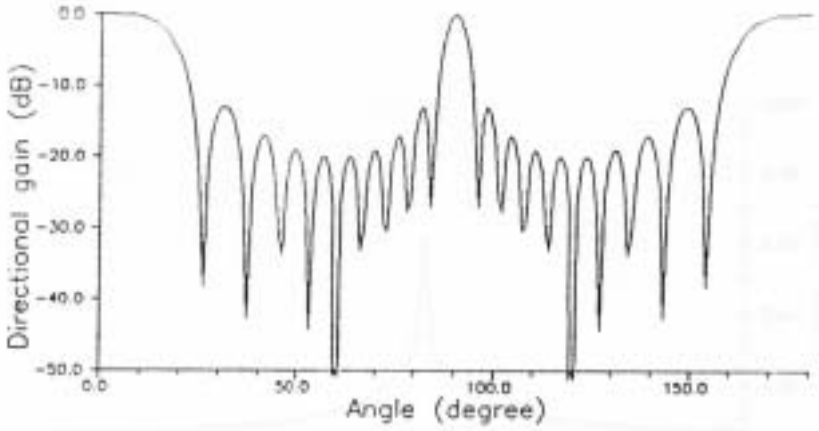

Fig. 4 Directional gain pattern
(10 microphones, λ spacing)

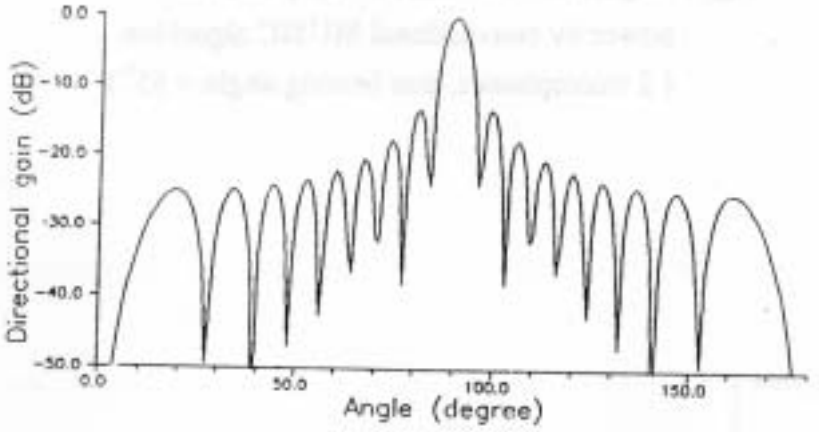

Fig. 5 Directional gain pattern
(18 microphones, $\frac{1}{2}\lambda$ spacing)

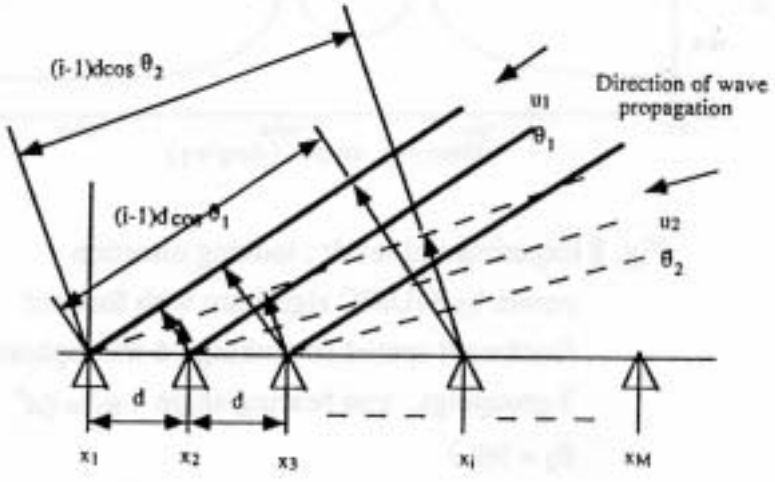

Fig. 6 Time delay between microphones

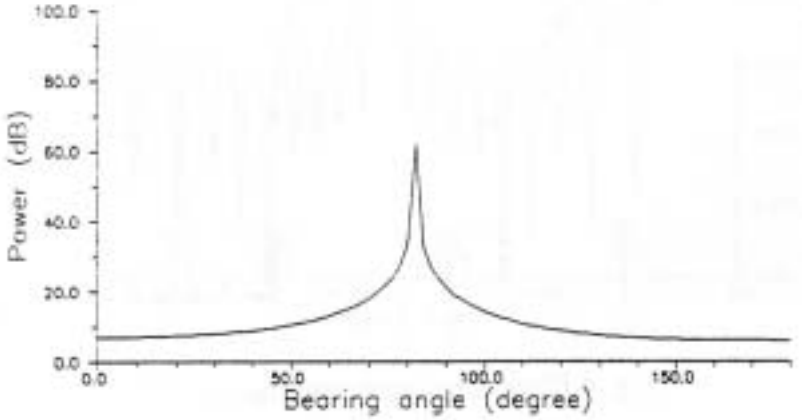

Fig. 7 Experimental result ; looking direction power by conventional MUSIC algorithm (2 microphones, true bearing angle = 85°)

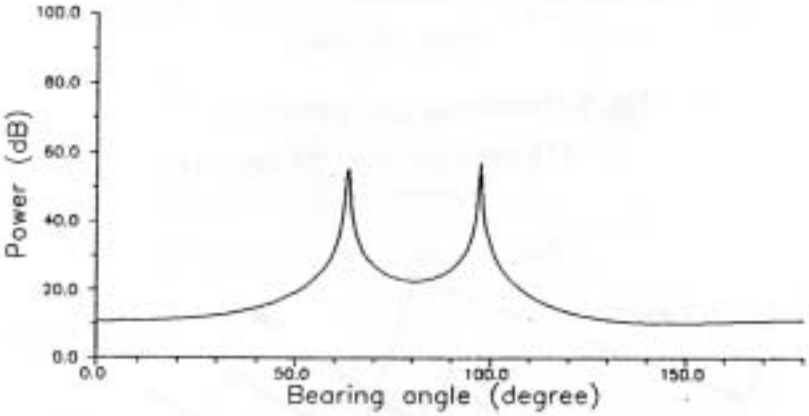

Fig. 8 Experimental result ; looking direction power by MUSIC algorithm with forward /backward spatial smoothing (5 microphones, 3 groupings, true bearing angle ; $\theta_1 = 64°$, $\theta_2 = 96°$)