inter-noise 83

SOUND SOURCE DETECTION AND LOCATION USING CROSS SPECTRA BETWEEN SIGNALS PICKED UP AT MANY POINTS

K. Kido, M. Abe, H. Noto (1) and Y. Ikegami (2)

(1) Research Center for Applied Information Sciences, Tohoku Univ. Katahira 2-1-1, Sendai 980, Japan
(2) Research Institute of technology, IHI

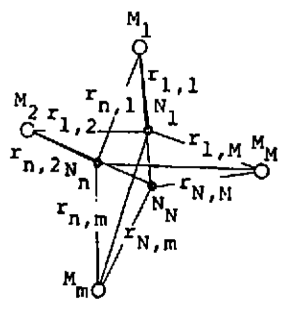
INTRODUCTION

We have developed an experimental system for simultaneous detection and location of sound sources in a noisy environment using cross spectra between signals picked up at many points. The performance of the system is first examined by computer simulations and an experimental system using 16 pick-ups is designed and developed. Preliminary experiments are carried out in air using 14 microphones distributed in a room.

PRINCIPLE

The signal radiated from N points is picked up by M pick-ups. The source signal has no correlation with one another. Figure 1 shows the relation between sources and pick-ups. No reflection from the boundary is first assumed. The output of the m-th microphone $x_m(t)$ is

Fig. 1 Relation between N sound sources and M pick-ups



expressed as follows:

$$x_m(t) = \sum_{n=1}^{N} \frac{1}{r_{n,m}} a_n(t - r_{n,m}/c), \qquad (1)$$

where $r_{n,m}$: distance between n-th source and m-th pick-up
c : sound velocity
$a_n(t)$: signal radiated from the n-th source.

$x_m(t)$ is windowed by an adequate function $w(t)$ and Fourier transformed as

$$X_m(\omega) = \frac{1}{T} \int_0^T w(t)\ x_m(t) \exp(-j\omega t)dt$$

$$= \sum_{n=1}^{N} \frac{1}{r_{n,m}}\ A_n(\omega) \exp(-j\omega r_{n,m}/c) \qquad (2)$$

where

$$A_n(\omega) = \frac{1}{T} \int_0^T w(t)\ a_n(t) \exp(-j\omega t)dt \qquad (3)$$

The sound source is first assumed to be at the p-th point. Then the modified spectrum $Y_{p,m}(\omega)$ is computed from $X_m(\omega)$ under the assumption that the transfer function $H_{p,m}(\omega)$ from the p-th source to the m-th pick-up is determined only by the difference in distance as follows:

$$Y_{p,m}(\omega) = X_m(\omega)/H_{p,m}(\omega)$$

$$= \sum_{n=1}^{N} \frac{r_{p,m}}{r_{n,m}}\ A_n(\omega) \exp\{-j\omega(r_{n,m} - r_{p,m})/c\} \qquad (4)$$

$Y_{p,m}(\omega)$ is computed for every pick-up. Then, the cross spectra between $Y_{p,k}(\omega)$ and $Y_{p,l}(\omega)$ are computed and averaged for all combinations of pick-ups as

$$W_p(\omega) = \frac{1}{M(M-1)} \sum_{k=1}^{N} \sum_{\substack{q=1 \\ k \neq q}}^{N} Y_{p,k}(\omega)\ Y^*_{p,q}(\omega)$$

$$= \sum_{n=1}^{N} \Big[\frac{|A_n(\omega)|^2}{M(M-1)} \sum_{k=1}^{M} \sum_{\substack{q=1 \\ k \neq q}}^{M} \frac{r_{p,k}\, r_{p,q}}{r_{n,k}\, r_{n,q}}$$

$$\exp[j\omega\{(r_{n,q} - r_{p,q}) - (r_{n,k} - r_{p,k})\}/c] \Big]$$

$$+ \sum_{n=1}^{N} \sum_{\substack{i=1 \\ n \neq i}}^{N} \Big[\frac{A_n(\omega) A_i^*(\omega)}{M(M-1)} \sum_{k=1}^{M} \sum_{\substack{q=1 \\ k \neq q}}^{M} \frac{r_{p,k}\, r_{p,q}}{r_{n,k}\, r_{i,q}}$$

$$\exp[j\omega\{(r_{i,q}-r_{p,q})-(r_{n,k}-r_{p,k})\}/c]] \qquad (5)$$

where A_n^* is the complex conjugate of $A_n(\omega)$.

The second term of Eq. (5) approaches zero when the number of combination M(M-1) is large because the distance from the assumed source to each microphone is different. The first term of Eq. (5) approaches a definite value if the source is located at the assumed point, but, otherwise it approaches zero. The following equation is obtained when the source is at the assumed point.

$$W_p(\omega) = |A_p(\omega)|^2 \qquad (6)$$

The assumed point is scanned over the space where the inspection should be carried out. The average cross spectrum takes a considerable value if the assumed point coincides with that of the real source.

Some image sources can be generated if the estimation is carried out using only one frequency component. However, the image source moves by the change in frequency. Therefore, the real source can be discriminated from the image sources observing the average cross spectrum over a range of frequency.

COMPUTER SIMULATION

A few computer simulations were carried out using the model shown in Fig. 2, where 32 pick-ups are used. In the computer simulation, the sound velocity assumed is 333 m/s, the sampling frequency is 30 kHz and the interval between successive assumed points is 25 cm. The source is assumed to be a point radiating white noise. When only a single frequency component is used (750 Hz), some image sources are

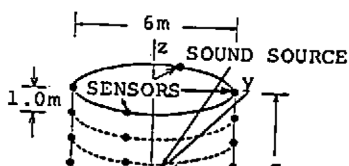


Fig. 2 Arrangement of 32 pick-ups and a source

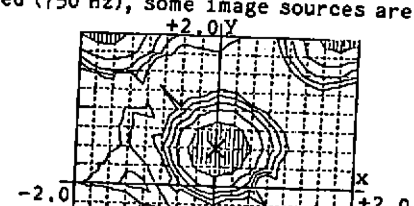


Fig. 3 Distribution of Average cross spectrum displayed by contour lines

generated as shown in Fig. 3, where x shows the point of real source. The hatched area shows the possibility of source point and there appear four image sources.

The next simulation is carried out for the investigation on the effect of two sources and of the external noise. The cross spectra are averaged both for every combination of pick-ups and for 16 frequency components between 937.5 Hz and 15 kHz to diminish image sources, and also averaged 10 times in the time domain to decrease external noise given to every pick-ups. Figure 4 shows the distribution of computed magnitude on the plane of sources. Two peaks are observed at the source points by the close scann (10 cm).

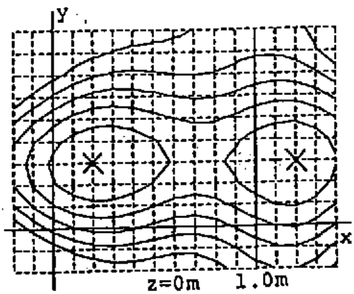


Fig. 4 Calculated distribution of cross spectrum

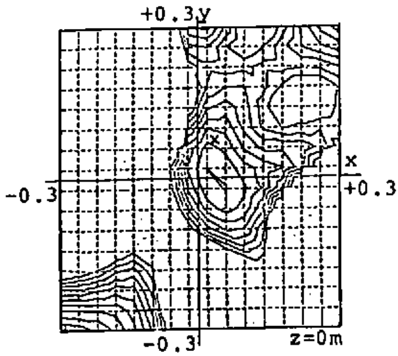


Fig. 5 Measured distribution of cross spectrum

FIELD EXPERIMENT

An experiment is carried out in a conference room using 14 pick-ups. The sampling frequency is 3.3 kHz and the cross spectra are averaged every 13 Hz in the frequency domain between 417 Hz and 1.25 kHz. Figure 5 shows the result computed for the plane where there is a real source.

CONCLUSION

A method for the sound detection and location is proposed with the description of principle. The usefulness of the method is shown both by computer simulations and by an experiment.

According to the investigation, not only the averaging of pick-ups but also the averaging in the frequency domain is necessary for diminishing image sources.

The effects of diffraction in the space, the boundary wall, the flow of the medium etc. should be investigated for practical use.