Journal of Sound and Vibration (2001) **241**(1), 53–56
doi:10.1006/jsvi.2000.3276, available online at http://www.idealibrary.com on IDEAL®

CUES FOR LOCALIZATION IN THE MEDIAN PLANE AS EXTRACTED FROM THE AUTO-CORRELATION FUNCTION

S. SATO AND Y. ANDO

Graduate School of Science and Technology, Kobe University, Rokkodai, Nada, Kobe 657-8501, *Japan. E-mail*: *satos@kobe-u.ac.jp*

AND

V. MELLERT

Department of Physics, University of Oldenburg, D-26111 *Oldenburg, Germany*

(*Accepted* 7 *August* 2000)

This study evaluates using factors extracted from the autocorrelation function (ACF) as cues for sound localization of the median plane. The amplitude of the transfer function for sound incident from the median plane to the ear entrances measured (see reference [6]) is translated into the ACF. It is shown that the sound localization in the median plane can be characterized by factors extracted from the ACF. These cues are the effective duration of the normalized ACF, τ_e, the delay time of the first peak τ_1, and its amplitude ϕ_1.

1. INTRODUCTION

A theory of primary sensations and spatial sensations to environmental noise has been proposed [1]. Primary sensations—loudness, pitch, and timbre—and spatial sensations can be described by temporal and spatial factors extracted from the autocorrelation function (ACF) and the interaural cross-correlation function (IACF) respectively. It has been shown that the characteristics of the environmental noises can be determined from these factors [2, 3]. Among the factors extracted from the IACF, the interaural delay time τ_{IACC} is a significant factor for determining the perceived horizontal sound localization of the source. However, the τ_{IACC} is theoretically zero when the sound source is incident from the median plane. Previously, sound localization in the median plane has been discussed in terms of the transfer function in the frequency domain with its peaks and dips [4–6]. However, the patterns of peaks and dips are insufficient as cues for directional information because they do not correspond directly to any spatial sensation. The ACF and the power density spectrum mathematically contain the same information. From the ACF analysis, (1) energy represented at the origin of delay, $\Phi(0)$, (2) effective duration of the envelope of the normalized ACF, τ_e, (3) the delay time of the first peak, τ_1, and (4) its amplitude, ϕ_1 were extracted. These factors may correspond to loudness, pitch, and timbre [1, 6–8]. The present study discusses whether or not the transfer function of the median plane can be characterized by these factors.

0022-460X/01/110053 + 04 $35.00/0

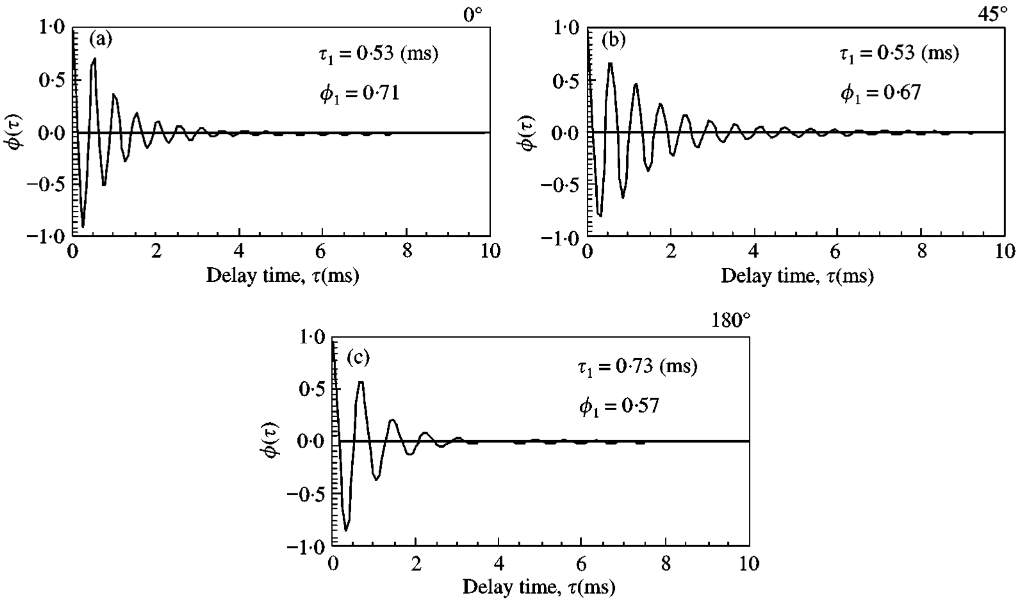

Figure 1. Normalized ACF for incident angles of 0, 45 and 180°.

2. METHOD

The amplitudes of the transfer functions for sound incident from the median plane to the ear entrances, as measured by Mehrgardt and Mellert [6], were rearranged. The ACFs of the transfer functions were obtained using the following steps.

1. Data sets were obtained from the figures from Mehrgardt and Mellert's paper using an optical image reader (scanner), with 300 data points each.
2. Amplitudes as a function of the frequency of the logarithmic scale were replotted as a function of the frequency in the real number.
3. Amplitudes in decibel scale were converted to real numbers, and the ACFs were calculated by an inverse Fourier transform after passing through an A-weighting filter.

3. RESULTS AND DISCUSSION

Examples of the normalized ACF are shown in Figure 1. There is a certain degree of correlation between both τ_n and τ_{n+i}, and ϕ_n and ϕ_{n+i}, as shown in Table 1, where τ_n and ϕ_n are the delay time and amplitude, respectively, of the nth peak of the ACF. Thus, τ_1, and ϕ_1 can be regarded as representing sets of τ_n and ϕ_n. Examples of plotting the absolute value of the ACF on a logarithmic scale are shown in Figure 2. The envelope of the decay of the ACF can be fitted by a straight line, and τ_e is easily obtained from the delay at which the envelope drops below -10 dB. In the figures from Mehrgardt and Mellert's paper, the amplitude values were plotted on a relative scale, but $\Phi(0)$ is not considered important as a cue for sound localization. As shown in Figures 1 and 2, the τ_1, which corresponds to pitch, for incident angles of 0 and 45° was equal, but τ_e for an incident angle of 45° ($\tau_e = 3{\cdot}14$ ms) was much larger than that of 0° ($\tau_e = 2{\cdot}16$ ms). The τ_1 for an incident angle of 180° was different from those for the above two angles; however, the ϕ_1, which corresponds to pitch strength, was relatively small.

Table 1

Correlation coefficients between τ_n and τ_{n+i} (a) *and ϕ_n and ϕ_{n+i}* (b)

(a)

	τ_1	τ_2	τ_3
τ_1	1·000	0·963	0·981
τ_2	—	1·000	0·992
τ_3	—	—	1·000

(b)

	ϕ_1	ϕ_2	ϕ_3
ϕ_1	1·000	0·798	0·806
ϕ_2	—	1·000	0·987
ϕ_3	—	—	1·000

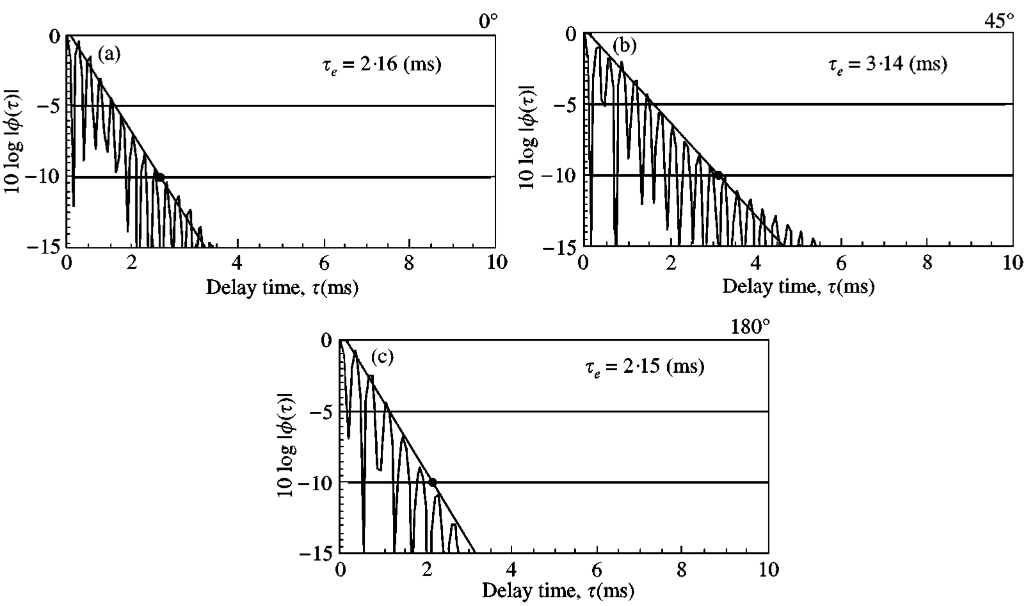

Figure 2. Effective duration of the ACF τ_e for incident angles of 0, 45 and 180°.

The remarkable finding here is that each angle from the median plane can be distinguished by the three τ_1, ϕ_1, and τ_e, as shown in Figure 3. The difference between incident angles in the median plane roughly corresponds to the distance between points in the three-dimensional space. The τ_1, ϕ_1, and τ_e-space can be interpreted as a "phase-state representation" of the special acoustic environment. The median plane incidence seems to be continuously mapped on a rather simple trajectory.

4. CONCLUSIONS

After the amplitudes of the transfer functions for the sound incident from the median plane to the ear entrances were translated into the ACF, the factors extracted from the ACF

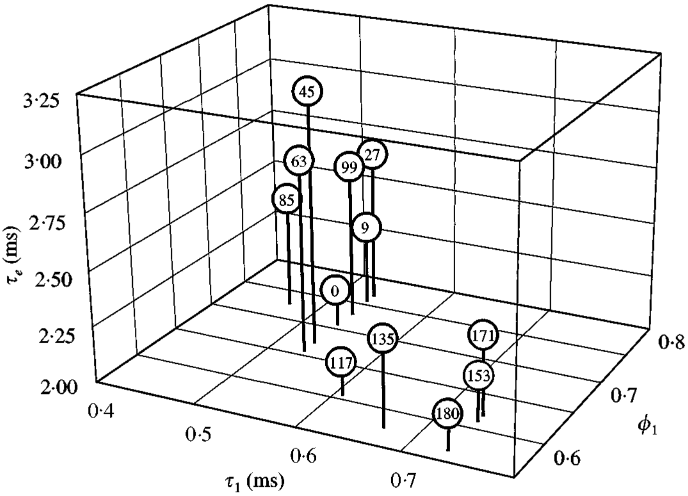

Figure 3. Three-dimensional illustration of τ_1, ϕ_1, and τ_e for each incident angle in the median plane. The number indicates the incident angle in degree.

were analyzed to clarify cues for sound localization in the median plane. The sound localization can be characterized by three factors: the effective duration of the normalized ACF and the delay time and amplitude of the first peak of the normalized ACF. These factors may correspond to timbre, pitch, and strength of the pitch.

ACKNOWLEDGMENTS

This work was supported by a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science.

REFERENCES

1. Y. Ando 2000 *Journal of Sound and Vibration* **241**, 3–18. A theory of primary sensations and spatial sensations measuring environmental noise, this is in this special issue.
2. H. Sakai, S. Sato, N. Prodi and R. Pompoli 2000 *Journal of Sound and Vibration* **241**, 57–68. Measurement of regional environmental noise by use of a PC-based system, this is in this special issue.
3. K. Fujii, Y. Soeta and Y. ando 2000 *Journal of Sound and Vibration* **241**, 69–78. Acoustical properties of aircraft noise measured by temporal and spatial factors, this is in this special issue.
4. J. Hebrank and D. Wright 1974 *Journal of the Acoustical Society of America* **56**, 1829–1834. Spectral cues used in the localization of sound sources on the median plane.
5. R. A. Butler and K. Belendiuk 1977 *Journal of the Acoustical Society of America* **61**, 1264–1269. Spectral cues utilized in the localization of sound in the median sagittal plane.
6. S. Mehrgardt and V. Mellert 1977 *Journal of the Acoustical Society of America* **61**, 1567–1576. Transformation characteristics of the external human ear.
7. Yoichi Ando 1998 *Architectural Acoustics, Blending Sound Sources, Sound Fields, and Listeners*, New York: AIP Press/Springer-Verlag; chapter 5.
8. Y. Ando, S. Sato and H. Sakai 1999 In *Computational Architectural Acoustics in Architecture* (J. J. Sendra, editor), Southampton: WIT Press; chapter 4. Fundamental subjective attributes of sound fields based on the model of auditory-brain system.