Journal of Sound and Vibration (2002) **258**(3), 443–450
doi:10.1006/jsvi.5267, available online at http://www.idealibrary.com on IDEAL®

EFFECT OF MODULATED DELAY TIME OF REFLECTION ON AUTOCORRELATION FUNCTION AND PERCEPTION OF ECHO

J. Atagi

Graduate School of Science and Technology, Kobe University, Rokkodai, Nada, Kobe, 657-8501, *Japan. E-mail: atagi@gamma.ocn.ne.jp*

AND

R. Weber and V. Mellert

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

(*Accepted* 30 *May* 2002)

An indoor sound field can be regarded more or less as a time-variant system. In this paper, the time-variant model of the impulse response, which consists of a direct sound and a reflection with a modulated delay time, is applied to simulate such a time-variant sound field. The effects of the modulated delay time of reflection are investigated in terms of physical properties and subjective evaluation of the temporal fluctuation of a sound field. The physical properties of the temporal fluctuation are analyzed by using the autocorrelation function (ACF) of the sound field and identified by the factors and behavior of the ACF. To examine subjective evaluation of the temporal fluctuation of a sound field, listening tests on the perception of echo were conducted. Such subjective evaluation may be associated with the ACF of a sound field. The relationship between the physical properties and the subjective evaluation of the temporal fluctuation of a sound field is also discussed.

1. INTRODUCTION

An indoor sound field can be regarded more or less as a time-variant system because of air currents and other movements. Though there have been many studies on the temporal fluctuation of outdoor sound propagation, it has rarely been investigated in the field of room acoustics. In fact, temporal fluctuation of sound propagation does exist in an indoor sound field. Ueda and Ando [1] investigated the fact that the air currents produced by air conditioners influenced sound propagation in a gymnasium. The fluctuation of the sound pressure level, especially for high frequencies, was confirmed in large indoor space. They proposed a time-variant model of the impulse response, based on statistical analysis of the sound-pressure-level fluctuation.

In this paper, a time-variant sound field is simulated by applying the time-variant model of the impulse response in an anechoic chamber. The model consists of a direct sound and a single reflection with a modulated delay time. There are difficulties in the investigations of subjective attributes for a sound field in a room because the sound field consists of a great number of reflections. However, the fundamental attributes are contained in the

0022-460X/02/$35.00

simplest sound field, which consists of a direct sound and a single reflection as representative of a set of reflections [2].

The physical properties of a time-variant sound field with a modulated delay time of reflection are analyzed by using the autocorrelation function (ACF) of the sound field for comparison to a time-invariant sound field with a fixed delay time of reflection. To examine the psychological effects of the modulated delay time of reflection, listening tests on the perception of echo were also conducted.

2. PHYSICAL PROPERTIES OF TIME-VARIANT SOUND FIELD

2.1. PROCEDURE

A musical motif was used as a sound source (Royal Pavane by Gibbons, 10 s). A sound field consists of a direct sound and a single reflection representing a set of reflections. The time-variant model of the impulse response was applied to simulate a time-variant sound field (Figure 1). It includes definitions of the delay time of reflection (Δt_1) and the modulation interval of the delay time (Δ). The initial value of Δt_1 was set to 240 ms, so that the echo could be perceived [2]. The Δt_1 was modulated by a sinusoidal wave for the time-variant sound field. The frequency of the sinusoidal wave was set to 0·2 Hz, because it was found to be 0·1–0·2 Hz in a statistical analysis of the SPL fluctuation in a real room [1].

Five different values of Δ (Table 1) were determined by a preliminary experiment on the just noticeable difference (*jnd*) of the modulation of the delay time of reflection. The preliminary experiment was conducted using the same sound source, measurement conditions and subjects as the following ACF analysis and listening tests. The subjects were instructed to increase or decrease the value of Δ in each of ten trials and judge the *jnd* of the modulation of the delay time of reflection. The Δ value for the *jnd* (Δ_{jnd}) was obtained for each trial for each subject. As the result of the preliminary experiment, the value of Δ_{jnd} was determined from the mean value for all the subjects, which was 10·6 ms.

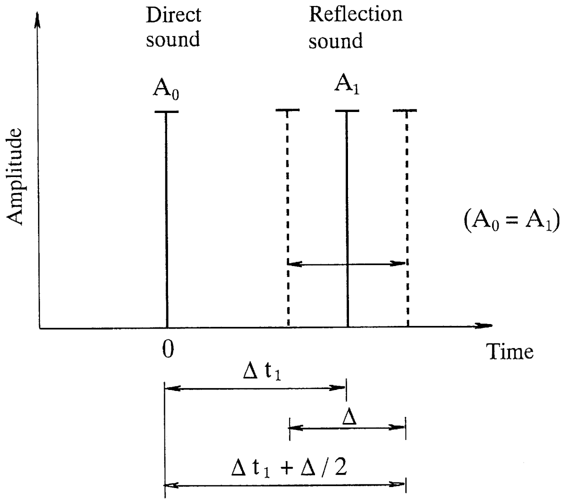

Figure 1. Time-variant model of the impulse response, which consists of a direct sound and a single reflection with a modulated delay time. The model includes definitions of the delay time of reflection (Δt_1) and the modulation interval of the delay time (Δ).

TABLE 1

The modulation intervals (Δ) used for the analyzed sound fields. Δ_{jnd} is the Δ value for the jnd of the modulation of the delay time of reflection

Sound field	Δ (ms)	
1	0	—
2	5·3	$\frac{1}{2}\,\Delta jnd$
3	10·6	Δjnd
4	21·2	$2\,\Delta jnd$
5	42·4	$4\,\Delta jnd$

2.2. AUTOCORRELATION FUNCTION ANALYSIS

To examine the physical properties of a time-variant sound field, the ACF was analyzed. Since time-variant sound signals were studied and a certain degree of coherence exists in the time sequence of these signals, which may greatly influence the subjective evaluation of a sound field, the short-time moving ACF was analyzed here [2]. The normalized ACF is calculated as a function of time t as follows:

$$\begin{aligned}\phi_p(\tau) &= \phi_p(\tau; t, T)\\ &= \frac{\Phi_p(\tau; t, T)}{[\Phi_p(0; t, T)\Phi_p(0; \tau + t, T)]^{1/2}},\end{aligned} \tag{1}$$

where

$$\Phi_p(\tau; t, T) = \frac{1}{2T}\int_{t-T}^{t+T} p'(s)p'(s+\tau)\,\mathrm{d}s, \tag{2}$$

and $p'(t) = p(t) * s(t)$, where $p(t)$ is a signal and $s(t)$ is the ear sensitivity. The short-time moving ACF was obtained for a signal duration $2T = 2{\cdot}0$ s with a moving interval. The signal duration was examined by changing between 1·0 and 4·0 s, based on a "psychological present," and determined to correspond to the hearing impression of the sound source [2, 3]. The running step for the moving interval was 0·1 s.

Four factors were extracted from the ACF: (a) the energy at the origin of the delay, $\Phi(0)$; (b) the effective duration of the envelope of the normalized ACF, τ_e; (c) the delay time of the first peak, τ_1; and (d) its amplitude, ϕ_1 [2].

2.3. EXTRACTED FACTORS OF THE ACF

Figure 2 shows the four factors extracted from the ACF of the sound source, a time-invariant sound field with a fixed delay time of reflection (sound field 1), and time-variant sound fields with modulated delay times of reflection (sound fields 2–5). It was found that the τ_e values of the time-variant sound fields were very different from that of the time-invariant sound field. On the other hand, there were no significant differences for the other three factors: $\Phi(0)$, τ_1 and ϕ_1. Therefore, τ_e should be a significant factor in representing the physical properties of a time-variant sound field.

2.4. RESULTS

To understand the ACF behavior for a time-variant sound field, it was compared with that for a time-invariant sound field. Figure 3 shows the log magnitude of the normalized ACF $\phi(\tau)$ for the sound source, the time-invariant sound field with a fixed delay time of reflection and a modulation interval Δ of zero (sound field 1), and the time-variant sound field with a modulated delay time of reflection and a Δ value of 10·6 ms (sound field 3). It was found that the log magnitude of the ACF of the sound field with the fixed delay time had a peak at 240 ms, which corresponds to Δt_1 and is much suppressed by adding the modulation.

To represent such differences in ACF behavior among sound fields with different values of Δ, the ratio of the magnitude of the ACF of a sound field to that of a sound source is defined as follows:

$$E_{p+1}(\tau) = \frac{|\phi_{p+1}(\tau)|}{|\phi_p(\tau)|}, \tag{3}$$

where $\phi_p(\tau)$ and $\phi_{p+1}(\tau)$ are the normalized ACF of a sound source and a sound field with a direct sound and a single reflection, respectively.

In Figure 4, the ratio $E_{p+1}(\tau)$ for the sound fields with fixed (sound field 1) and modulated delay times of reflection (sound fields 2–5) are compared for a delay τ of

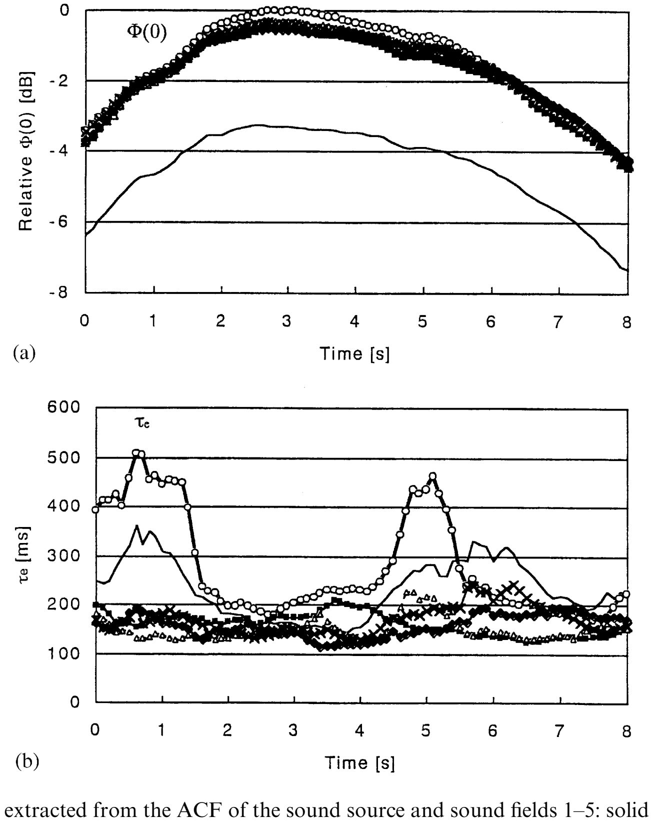

Figure 2. Factors extracted from the ACF of the sound source and sound fields 1–5: solid line (sound source), ○, 1; ■, 2; △, 3; ◆, 4; ×, 5. The τ_1 of the ACF is related to the pitch sensation [Y. Ando 2000 *Journal of Sound and Vibration* **241**, 3–18. A theory of primary sensations and spatial sensations measuring environmental noise]. The jumps in the curves below 2 s are due to the change of the pitch of the sound source.

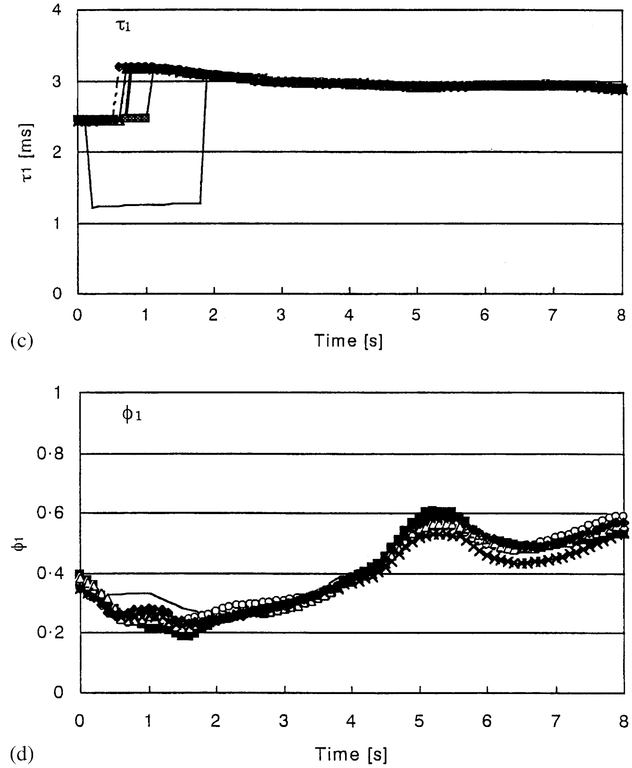

Figure 2. Continued.

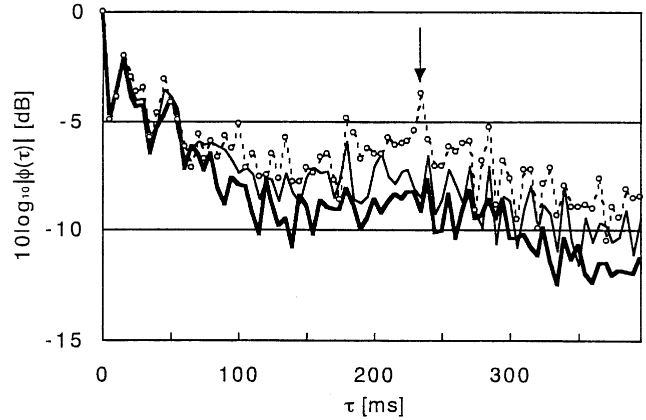

Figure 3. Log magnitude of the normalized ACF $\phi(\tau)$ for the sound source (solid line), the time-invariant sound field with a fixed delay time of reflection and a modulation interval Δ of zero (sound field #1, open circles), and the time-variant sound field with a modulated delay time of reflection and a Δ value of 10·6 ms (sound field #3, bold line). The arrow shows the delay τ corresponding to Δt_1.

240 ms, which corresponds to Δt_1. It was found that the $E_{p+1}(240\,\text{ms})$ values for the sound field with the fixed delay time of reflection (1) were larger than 1·0, while those for the sound fields with the modulated delay times of reflection (2–5) were smaller than 1·0.

By further examining the ACF behavior, it was found that the magnitude of the ACF was much suppressed by adding the modulation to the delay time of reflection. As the Δ

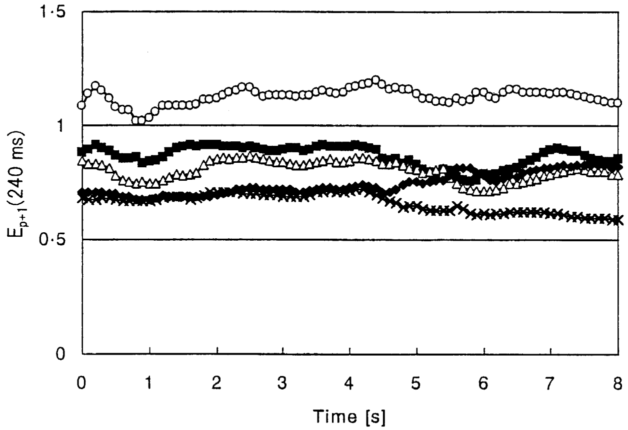

Figure 4. Ratio $E_{p+1}(\tau)$ for the sound fields with fixed (#1) and modulated delay times of reflection (2–5) for a delay τ of 240 ms: ○, 1; ■, 2; △, 3; ◆, 4; ×, 5.

value increased, the ratio E_{p+1}(240 ms) decreased. A small value of $E_{p+1}(\tau)$ means that the total coherence between the direct sound and the reflection is low. Adding the modulation to the delay time of the reflection in a time-variant sound field is considered to make the coherence low.

3. PERCEPTION OF ECHO IN TIME-VARIANT SOUND FIELD

3.1. PROCEDURE

Paired-comparison tests on perception of the echo were conducted. The sound stimuli were the same as those used in the ACF analysis. The stimuli were paired and presented to subjects in an anechoic room by two loudspeakers for a direct sound and a reflection, respectively. The subjects judged which of two stimuli caused them to feel a stronger echo. This experiment was carried out ten times repetitively for each of ten subjects, aged 22–27.

3.2. RESULTS

Figure 5 shows the mean scale values of echo for all the subjects as a function of Δ. The scale values were obtained by the law of comparative judgment (Case V; Thurston) [4]. It was found that the scale value of echo significantly increased with larger Δ ($p < 0.01$). The mean values of E_{p+1}(240 ms) for each sound field with different values of Δ are also shown in Figure 5. The values of E_{p+1}(240 ms) decreased with larger Δ. The correlation coefficient between the scale values and E_{p+1}(240 ms) is -0.78.

4. DISCUSSION

A model of the auditory–brain system for subjective responses was proposed for the major independent acoustic factors, classified by temporal and spatial factors [2]. The model consists of the autocorrelation mechanism, the interaural cross-correlation mechanism between the two auditory pathways, and the specialization of human cerebral hemispheres for temporal and spatial factors of the sound field. In this model, subjective

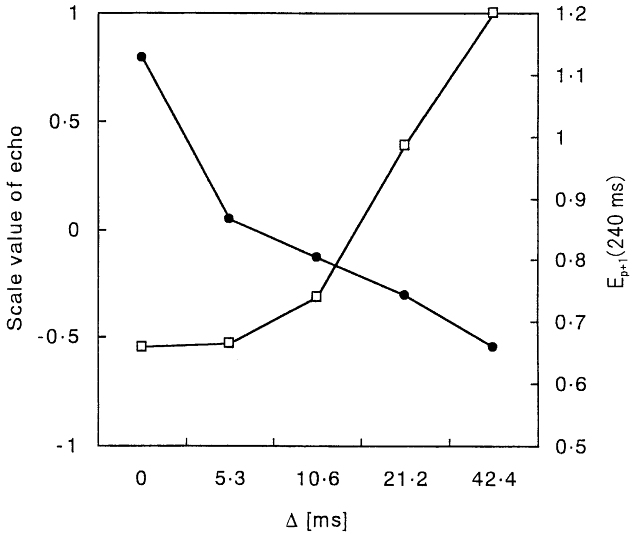

Figure 5. Mean scale values of echo for all the subjects (□) and mean values of $E_{p+1}(240\,\mathrm{ms})$ (●) as a function of Δ.

responses to temporal factors are considered to be dominantly processed in the left hemisphere and related to the ACF of the sound signals.

In this paper, the perception of echo may be associated with the ACF behavior of the sound field. From the listening tests conducted here, it was found that the perception of echo increased with longer modulation interval, while the magnitude of the ACF at the delay time corresponding to Δt_1 decreased. We think that the echo can be easily perceived because adding the modulation to the delay time of reflection causes the total coherence between the direct sound and the reflection to become lower than that of the direct sound alone around the delay time corresponding to Δt_1. The ACF behavior is thus significant in examining the relationship between the physical properties and the perception of echo in a time-variant sound field.

5. CONCLUSIONS

The ACF of a sound field was analyzed to examine the physical properties of its temporal fluctuation. It was found that the magnitude of the ACF is much suppressed by adding a modulation to the delay time of reflection. The effect of the modulated delay time of reflection can be identified by the effective duration τ_e and the ACF behavior of the sound field.

The perception of echo in a time-variant sound field was also investigated. The perception of echo was found to increase with longer modulation interval, while the magnitude of the ACF at the delay time corresponding to Δt_1 decreases. The effect of the modulated delay time of reflection on subjective evaluation can be associated with the ACF behavior of the sound field.

ACKNOWLEDGMENT

The authors wish to thank Professor Yoichi Ando for his valuable suggestions and advice.

REFERENCES

1. Y. Ueda and Y. Ando 1997 *Journal of the Acoustical Society of America* **102**, 2771–2775. Effects of air conditioning on sound propagation in a large space.
2. Y. Ando 1998 *Architectural Acoustics—Blending Sound Sources, Sound Fields, and Listeners*. New York: AIP Press/Springer-Verlag.
3. P. Fraisse 1982 in *Psychology of Music* (D. Deutsch, editor), Orlando, Fl: Academic Press; chapter 6: Rhythm and Tempo.
4. L. L. Thurston 1927 *Psychological Review* **34**, 273–289. A law of comparative judgment.