Journal of Sound and Vibration 330 (2011) 3526–3537

Contents lists available at ScienceDirect

Journal of Sound and Vibration

journal homepage: www.elsevier.com/locate/jsvi

Subjective diffuseness of music signals convolved with binaural impulse responses

Ryota Shimokura [a], Lamberto Tronchin [b], Alessandro Cocchi [b], Yoshiharu Soeta [a,*]

[a] Health Research Institute, National Institute of Advanced Industrial Science and Technology, 1-8-31 Midorigaoka, Ikeda, Osaka 563-8577, Japan
[b] Department of Energy, Nuclear and Environmental Control Engineering, Bologna University, Viale Risorgimento 2, Bologna 40136, Italy

ARTICLE INFO

Article history:
Received 28 May 2009
Received in revised form
2 February 2011
Accepted 9 February 2011
Handling Editor: R.E. Musafir
Available online 9 March 2011

ABSTRACT

The spatial impression of sound in a hall can be quantified using sound field factors such as the interaural cross-correlation coefficient (IACC) calculated from binaural impulse response (BIR), henceforth denoted by $IACC_{IR}$. The subjective diffuseness for the listener is a spatial attribute which depends on factors associated both with the source signal and with the actual sound field, and is quantified using the IACC of the signal received by the listener, henceforth denoted by $IACC_{SR}$. Therefore, the subjective diffuseness in a given hall may change with the music. The aims of this study are to estimate the $IACC_{SR}$ from the $IACC_{IR}$ and the factors, which is obtained from autocorrelation function (ACF) of music signal, and to evaluate the subjective diffuseness by these factors. First, the relationship between the $IACC_{IR}$ and $IACC_{SR}$ was investigated. Second, subjective diffuseness was measured by a psycho-acoustical experiment. As a result, the $IACC_{SR}$ could be estimated from the $IACC_{IR}$ of the BIR and the effective duration (τ_e) from the ACF of music signal. It was found that the effects of BIRs on subjective diffuseness could be evaluated by $IACC_{IR}$ for almost all subjects, while the effects of music signals could be evaluated by the τ_e and the width of the peak at τ=0 ($W_{\phi(0)}$) of the ACF.

1. Introduction

Spatial impression is an important subjective attribute of a sound field, which can be separated into two subjective dimensions: apparent source width (ASW) and listener envelopment (LEV), and has been evaluated from sound field factors calculated from binaural impulse responses (BIRs) measured in halls [1]. The ASW is defined as the width of a sound image fused from direct sound and early reflections. The LEV is defined as the sense of being surrounded by late arriving sound from a diffused array of distant sound images. The ASW has a high correlation with the maximum peak amplitude (IACC) and the width of the maximum peak (W_{IACC}) of the interaural cross-correlation function (IACF) [2–5]. The LEV has a high correlation with the late-arriving relative lateral sound level (LG or G_{LL}) and the late lateral energy fraction (LF_L) [8,9]. The common experimental method observed in these studies for the spatial impression is a fixed source signal with different sound fields. For example, for simulated sound fields, the source signal used in terms of ASW was Mozart's "Jupiter" [6,7], and the source signal used in terms of LEV was Handel's "Water Music" [8,9].

* Corresponding author. Tel.: +81 72 7518526; fax: +81 72 7518416.
E-mail address: y.soeta@aist.go.jp (Y. Soeta).

0022-460X/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsv.2011.02.014

On the other hand, the subjective diffuseness is a spatial impression considering also the effect of the source signals. Originally, subjective diffuseness was evaluated by the IACC of white noise which arrived from a different horizontal angle using multichannel loudspeaker reproduction [10,11]. A large IACC makes the listener perceive the well-defined direction of the incoming sound. A small IACC corresponds to subjectively diffused sound, and the listener has no impression of the clear direction of the sound. Researchers have investigated the effects of both source signal and source location on subjective diffuseness [12–14]. Octave bandpass noise with different center frequencies were used as source signal, while the sound field was altered by changing the locations of the loudspeakers [12,13]. In the results, the IACC of the bandpass noise arriving at left and right ears changed according to the center frequency even if the location of the loudspeakers was the same, and the subjective diffuseness showed a strong correlation with the IACC.

The IACF is defined by the correlation between the signals at the left, $p_l'(t)$, and right, $p_r'(t)$, ears as a function of delay time τ (Appendix A) [15,16]. When the binaural signals, $p_l'(t)$ and $p_r'(t)$, are a BIR, the calculated IACC is a sound field factor used to evaluate the spatial impression of a hall. Since the BIRs indicate the transfer function of sound fields, an anechoic source convolved with the BIRs represents the signal that listeners hear in that sound field. Therefore, when the binaural signals are music convolved with a BIR, the IACC is a combined factor of both sound field and source signal used to evaluate the subjective diffuseness. This study defines the former as IACC_{IR} (i.e. the IACC of impulse response) and the latter as IACC_{SR} (i.e. the IACC of other signal response through a hall). For example, the subjective rank ordering of concert halls has been determined by IACC_{IR} [17], while the subjective preference of concert halls, as an overall impression of sound fields, has been evaluated by IACC_{SR} [15].

During a musical performance, listeners listen to the music not to the impulse signal. Therefore, in the actual situation, sound arriving at listeners is not determined only by the BIR of a hall but also by the music signal. The IACC_{SR} changes according to the music even when the source signal is presented by the same loudspeaker location [18]. However, it has not been clarified yet which factors extracted from the music signal affect the IACC_{SR} and the subjective diffuseness. The aims of this study are to estimate the IACC_{SR} by the IACC_{IR} and factors obtained from music signal and to evaluate the subjective diffuseness by these factors. First, the relationship between IACC_{IR} and IACC_{SR} is investigated using eight BIRs and five music signals. Second, the subjective diffuseness was evaluated in a simulated sound field using four BIRs and three music signals. It is hypothesized that the subjective diffuseness has a correlation with the calculated IACC_{SR} by the IACC_{IR} and the factors from the music signal.

2. Relationship between IACC_{IR} and IACC_{SR}

2.1. Concept

The cross-correlation of two real functions $f(t)$ and $g(t)$ of a real variable t, denoted Φ_{fg}, is defined by

$$\Phi_{fg} = f(-t)^* g(t), \tag{1}$$

where * denotes convolution. Correlation and convolution are the same except for the flip of t in one of the two functions. When a music signal, $m(t)$, is convolved with a BIR at the left, $h_l(t)$, and right, $h_r(t)$, ears, the echoic music arriving at left and right ears can be expressed as

$$mh_l(t) = m(t)^* h_l(t), \tag{2}$$

$$mh_r(t) = m(t)^* h_r(t). \tag{3}$$

The IACF of the music signal response through a hall (IACF_{SR}) is calculated by Eqs. (1)–(3) as follows:

$$\text{IACF}_{\text{SR}} = \Phi_{mh_l mh_r} = mh_l(-t)^* mh_r(t) = m(-t)^* m(t)^* h_l(-t)^* h_r(t) = \Phi_{mm}{}^* \Phi_{lr}, \tag{4}$$

where Φ_{mm} is the autocorrelation function (ACF) of the music signal, and Φ_{lr} is the IACF of the BIR (IACF_{IR}). The IACC_{IR} and IACC_{SR} is calculated by

$$\text{IACC}_{\text{IR}} = \frac{\Phi_{lr}(\tau_{\text{IACC}})}{\sqrt{\Phi_{ll}(0)\Phi_{rr}(0)}}, \tag{5}$$

$$\text{IACC}_{\text{SR}} = \frac{\int_{-T}^{T} \Phi_{mm}(t)\Phi_{lr}(t+\tau_{\text{IACC}})\,\mathrm{d}t}{\sqrt{\int_{-T}^{T} \Phi_{mm}(t)\Phi_{ll}(t)\,\mathrm{d}t \int_{-T}^{T} \Phi_{mm}(t)\Phi_{rr}(t)\,\mathrm{d}t}}, \tag{6}$$

where τ_{IACC} is the interaural delay time of the maximum peak of the IACF_{IR}. Φ_{ll} (Φ_{rr}) is the ACF of the BIR at the left (right), and $2T$ is the integral time of the ACF and IACF. When $m(t)$ is a white noise, the ACF becomes a Dirac delta function; thus, the IACF_{SR} (or IACC_{SR}) is equal to the IACF_{IR} (or IACC_{IR}). When $m(t)$ is a pure tone, the ACF becomes a cosine function, so the IACF_{SR} is also a cosine function, and IACC_{SR} is always 1 regardless of the IACC_{IR}. Therefore, the periodicity included in the signal affects the IACC_{SR}.

The periodicity of music lies between those of white noise (no periodicity) and pure tone (perfect periodicity). The degree of periodicity can be expressed by the effective duration (τ_e) of the ACF (Appendix B) [15]. The τ_e is defined by the delay time at which the envelope along the early decay of the normalized ACF falls by 10 dB. The τ_e of white noise and of a pure tone is 0 and ∞, respectively, and the τ_e of music lies between 0 and ∞. The τ_e can be controlled by using noise with different bandwidths [19]. When the bandwidth of the noise becomes narrower, the τ_e becomes longer. For example, when the bandwidths were changed to 640, 320, 160, 80, 40, 20, and 10 Hz, the τ_e of the bandpass noises changed by 5, 10, 20, 40, 97, 159 and 318 ms, respectively [19]. Fig. 1 shows the relationship between $IACC_{SR}$ and τ_e of bandpass noise. The used BIRs are explained in Section 2.3. The $IACC_{SR}$ of bandpass noise becomes higher as the τ_e becomes longer. Thus, the $IACC_{SR}$ of music may also become higher as the τ_e calculated from the music become longer.

2.2. Music signal

Eight music signals generated using MIDI were used as source signals. Two melodies (A and B) were played by four musical instruments (Table 1). The scores of Melodies A and B are shown in Fig. 2. The duration of music was 10 s. The sampling rate and size were 44.1 kHz and 16 bit, respectively.

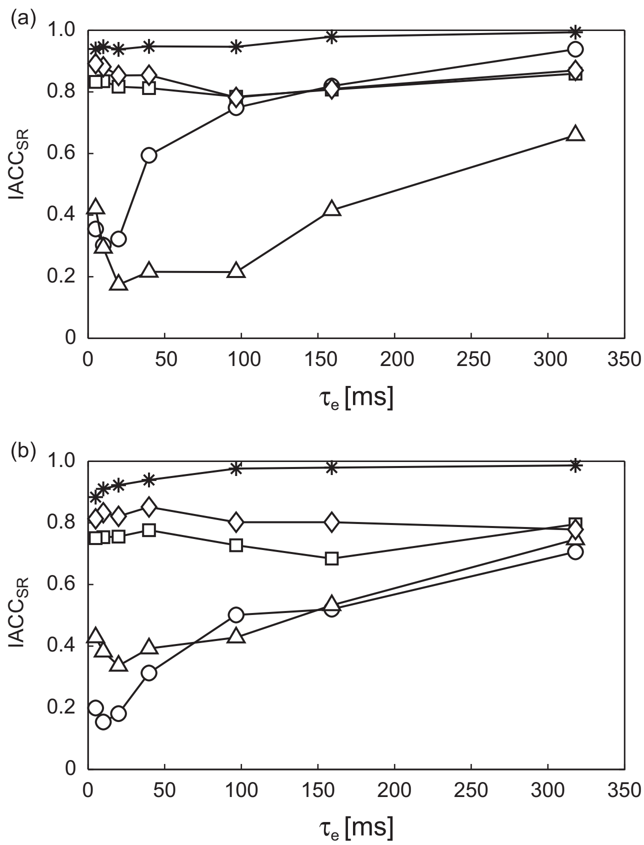

Fig. 1. $IACC_{SR}$ from bandpass noise convoluved with BIRs as a function of τ_e from the bandpass noise. The center frequencies of the bandpass filters were (a) 500 Hz and (b) 1 kHz, and the bandwidths were 640, 320, 160, 80, 40, 20, and 10 Hz. The different symbols indicate BIR 1 (○), BIR 2 (△), BIR 3 (□), BIR 4 (◇) and BIR 5 (*).

Table 1
ACF factors of the music signals.

Music signal	$\Phi(0)$ (dB)	τ_1 (ms)	ϕ_1	τ_e (ms)	$W_{\phi(0)}$ (ms)
A by glockenspiel	43.60	9.05	0.68	404.49	0.20
A by harpsichord	41.73	8.28	0.37	165.56	0.22
A by piano	36.33	5.49	0.35	236.31	0.38
A by trumpet	42.73	6.19	0.34	66.28	0.19
B by harpsichord	36.05	7.60	0.71	223.37	0.16
B by organ	38.79	5.51	0.82	627.03	0.18
B by piano	31.41	5.51	0.52	402.60	0.39
B by strings ensemble	41.22	7.60	0.61	85.91	0.18

Fig. 2. Scores of (a) Melody A and (b) Melody B.

Music signals are characterized by ACF factors, which are (1) the energy represented at the origin of the delay [$\Phi(0)$], (2) the delay time of the maximum peak (τ_1), (3) the amplitude of the first maximum peak (ϕ_1), (4) the width of the peak at τ=0 ($W_{\phi(0)}$), and (5) τ_e (Appendix B) [15,20]. The $\Phi(0)$, τ_1 and ϕ_1 are related to perceived loudness, pitch and pitch strength of a music signal, respectively [15]. The $W_{\phi(0)}$ indicates the spectral centroid of a signal [20]. The broad and narrow $W_{\phi(0)}$ mean that the low and high spectral components are strong, respectively. The τ_e indicates the degree of tonal components and reverberation included in the music signal, and is dependent on the type of musical instrument, the tempo of the melody, and the pattern of playing [15]. Because ACF factors of music vary as a function of time, the ACF of the music was calculated in the integral interval (2T) that slides along the duration of music. In this study, we calculated the running ACF using a 2T of 0.5 s with 0.1 s sliding steps (Fig. 3). The represented ACF factors were determined from the median values (Table 1).

2.3. BIR

Five BIRs measured in Kirishima International Concert Hall (BIR 1 and 2), Tsuyama Concert Hall (BIR 3 and 4) in Japan, and Delphi Ancient Theatre (BIR 5) in Greece were used as sound fields. In the measurements, an omni-directional loudspeaker was located on the center of the stage, and all BIRs were measured with a dummy head in a central seat of the hall. The signal was a time-stretched pulse with exponentially varied frequency from 40 Hz to 20 kHz over 10 s (a sine sweep), and the BIRs were obtained by deconvolution of the recorded signal. The sampling rate and size were 44.1 kHz and 16 bit, respectively.

Table 2 shows the IACF factors for each sound field. To examine only the effect of the IACC, we selected BIRs that have different $IACC_{IR}$ but similar IACF factors, $(W_{IACC})_{IR}$ and $(\tau_{IACC})_{IR}$. Since the $IACC_{IR}$ of BIR 1 was quite low, $(\tau_{IACC})_{IR}$ was not close to 0 in spite of putting the source and receiver in the center of the hall. The difference of the reverberation times of the BIRs in the halls (BIR 1 to 4) was small (1.93 $\pm$ 0.03 s). The reverberation time of BIR 5 was 0.53 because the theater is outdoors.

2.4. Estimation of $IACC_{SR}$

The 40 echoic music signals were obtained by the convolutions between eight anechoic music signals and five BIRs. Since the echoic music signals lasted for 10 s, the running IACF was calculated in the same manner as the running ACF. Table 3 shows the IACF factors calculated from the examples of the echoic music signals. The represented values were determined from the median values. Fig. 4 shows the relationship between $IACC_{IR}$ and $IACC_{SR}$. The $IACC_{SR}$ increased with increasing $IACC_{IR}$. The relationship can be expressed by a power function as

$$IACF_{SR} = IACC_{IR}^{\alpha}. \quad (7)$$

Table 4 shows the exponent, α, and the correlation coefficient, r, for the fitting for each music signal. Almost all music signals fitted Eq. (7) with high correlations and the α values were different according to the music signals. Fig. 5 shows α as a function of τ_e of the anechoic music signals. The relationship between α and τ_e is determined by an asymptotic curve regression and expressed by

$$\alpha = \frac{286}{\tau_e + 286} \quad (8)$$

with high correlation (r=0.95), and the probability of significance (p) for the correlation coefficient was lower than 0.01. For white noise, τ_e is theoretically 0, and $IACC_{SR}$=$IACC_{IR}$ (α=1). For pure tones, τ_e is ∞, and $IACC_{SR}$=1 (α=0). When the τ_e of a musical motif is short, the value of α approaches 1. On the other hand, when the τ_e of a musical motif is long, the value of α approaches 0. The correlation between the $IACC_{SR}$ obtained from the convolved music with the BIR and estimated $IACC_{SR}$ by Eqs. (7) and (8) was high and significant (r=0.95, $p < 0.01$).

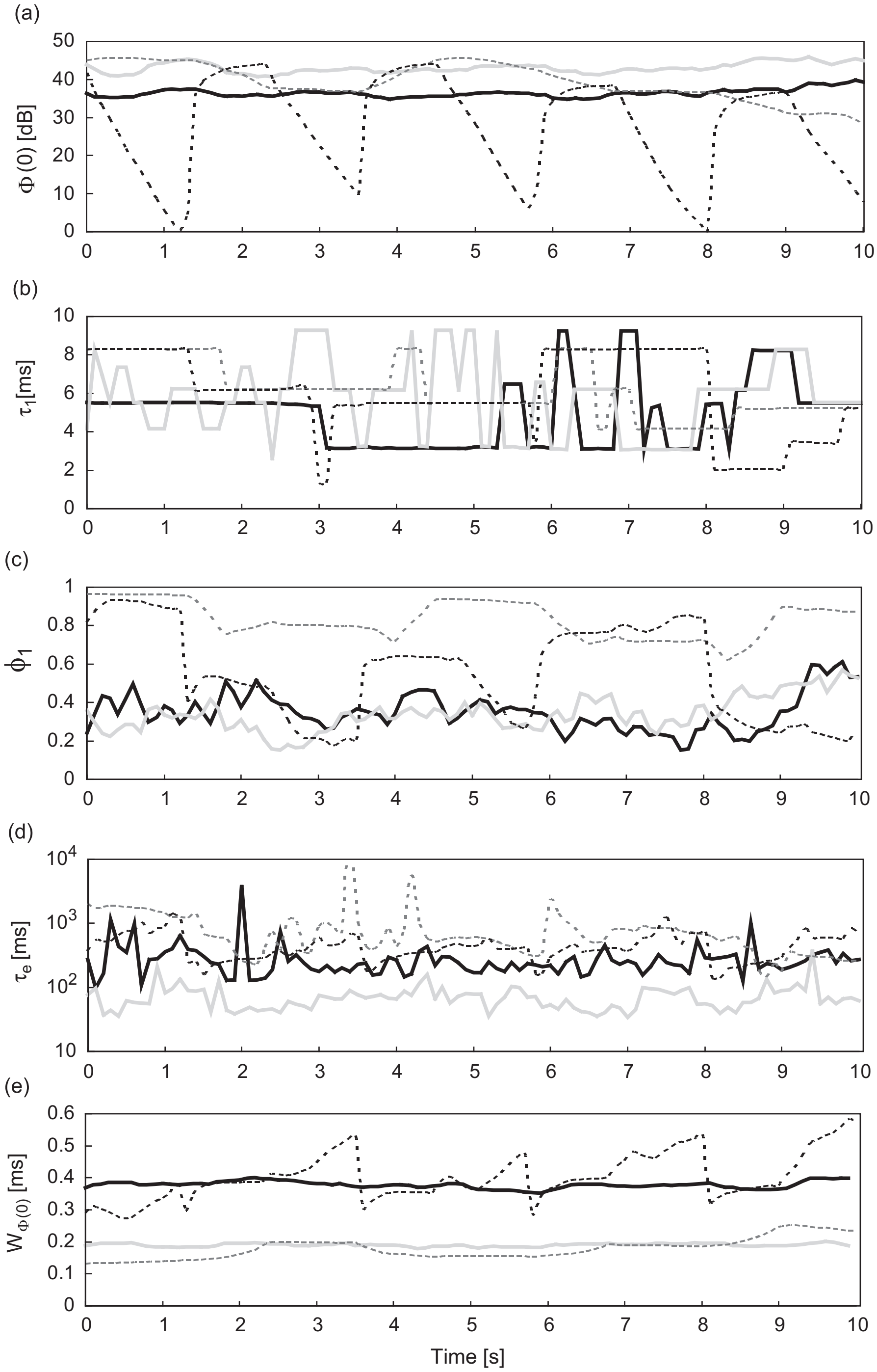

Fig. 3. ACF factors as a function of time for Melody A by piano (—), Melody A by trumpet (—), Melody B by piano (- - -), and Melody B by organ (- - -).

3. Psycho-acoustical experiment

3.1. Stimuli and presentation method

The stimuli were four music signals (Melody A played by piano or trumpet, and Melody B played by piano or organ) convolved with three BIRs (BIRs 1 to 3). A total of 12 stimuli were obtained.

Two loudspeakers were located in the anechoic chamber (3.2 m in width, 3.2 m in length, and 2.6 m in height), and the BIRs measured in the different halls were presented using the stereo dipole method (Fig. 6). To create the parallel

Table 2
IACF factors of the BIRs.

BIR	$IACC_{IR}$	$(W_{IACC})_{IR}$ (ms)	$(\tau_{IACC})_{IR}$ (ms)
BIR 1	0.05	0.05	0.88
BIR 2	0.29	0.04	0.00
BIR 3	0.57	0.04	0.04
BIR 4	0.68	0.05	0.02
BIR 5	0.86	0.04	0.02

Table 3
IACF factors of the music signals convolved with BIRs.

Music signal	BIR	LL_{SR} (dB)	$IACC_{SR}$	$(w_{IACC})_{SR}$ (ms)	$(\tau_{IACC})_{SR}$ (ms)
A by piano	BIR 1	51.06	0.19	0.13	0.20
A by trumpet		51.68	0.13	0.09	0.50
B by organ		53.17	0.46	0.09	−0.71
B by piano		41.19	0.35	0.14	0.01
A by piano	BIR 2	44.67	0.42	0.14	0.00
A by trumpet		46.02	0.37	0.09	0.02
B by organ		48.49	0.63	0.08	0.07
B by piano		34.76	0.47	0.13	−0.05
A by piano	BIR 3	45.69	0.85	0.14	0.02
A by trumpet		46.53	0.74	0.10	0.02
B by organ		49.64	0.82	0.12	0.00
B by piano		37.28	0.86	0.13	0.02

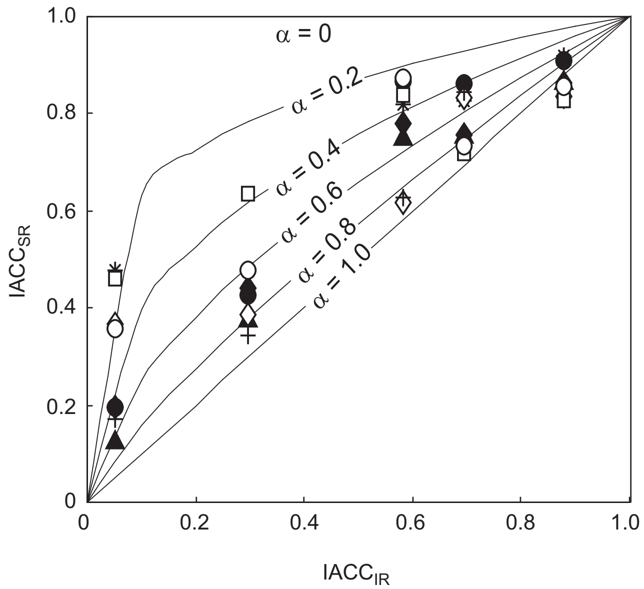

Fig. 4. Relationship between $IACC_{IR}$ and $IACC_{SR}$ for Melody A by glockenspiel (*), harpsichord (◆), piano (●), trumpet (▲), and Melody B by harpsichord (◇), organ (□), piano (○), strings ensemble (+).

presentation from the left (right) loudspeaker to the left (right) ear, crosstalk canceling (CTC) filters were convolved before the output [21–23]. The CTC filter is generated by converting a BIR measured in the anechoic chamber [23] and cancels only the crosstalk pathway sounds using the sound interference effect. The LL_{SR} at the listener's position, which corresponds to the ACF factor $\Phi(0)$, was 70 dBA.

Before the psycho-acoustical experiment, the 12 stimuli were presented in the chamber, and their $IACC_{SR}$, $(W_{IACC})_{SR}$, and $(\tau_{IACC})_{SR}$ were calculated using binaural data obtained by the dummy head (Neumann KU-100). The differences between the IACF factors obtained from the convolved music and the presented music by the stereo dipole procedure, $\Delta IACC_{SR}$, $\Delta(W_{IACC})_{SR}$, and $\Delta(\tau_{IACC})_{SR}$, were 0.025, 0.024, and 0.068, respectively. The $\Delta IACC_{SR}$ was much lower than the just

Table 4
Exponential coefficient α and correlation coefficient r.

Music signal	α	r
A by glockenspiel	0.38	0.83
A by harpsichord	0.62	0.98**
A by piano	0.54	0.97**
A by trumpet	0.72	0.99**
B by harpsichord	0.61	0.82
B by organ	0.34	0.61
B by piano	0.46	0.87*
B by strings ensemble	0.75	0.98**

* Indicates 5% significant levels.
** Indicates 1% significant levels.

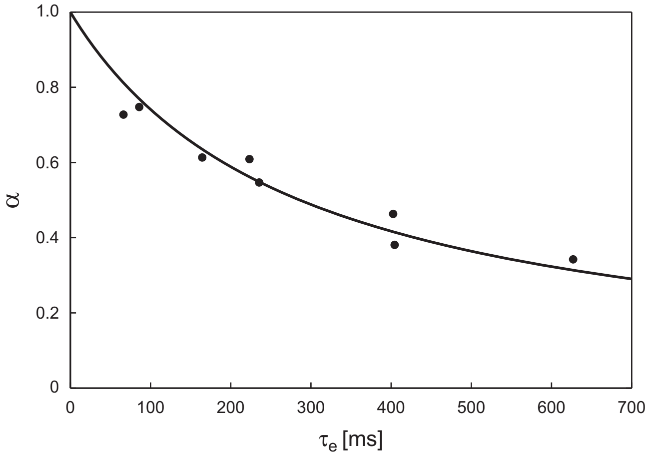

Fig. 5. Relationship between the exponential coefficient α and τ_e for the music (●).

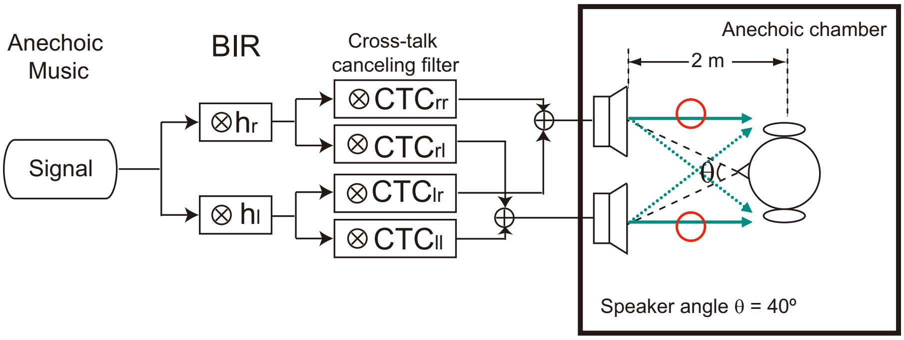

Fig. 6. Block diagram of presentation with stereo dipole.

noticeable difference (JND) of the IACC [24]. These confirmed that the stereo dipole method reproduced the desired sound field. To investigate the individual differences of the reproduction accuracy, the stimuli were presented for three subjects (A, B and F), and the IACF factors were calculated using a binaural microphone (Type 4101; B&K). The subjects put the small microphones at their ear canals, and sat at the same location of the dummy head. As results of analysis of variance (ANOVA), the difference of IACF factors for the dummy head and the subjects was not significant ($F_{3,\,44}$=0.38 for $\Delta IACC_{SR}$, $F_{3,\,44}$=0.77 for $\Delta(W_{IACC})_{SR}$, and $F_{3,\,44}$=0.38 for $\Delta(\tau_{IACC})_{SR}$). These results mean that these subjects were essentially in the same sound field. The subjects were alerted to avoid head movements during the psycho-acoustical experiments, although the stereo dipole method has been shown to be particularly robust to head movement [25–29]. Even though the stereo dipole method may introduce misalignment causing localization of a frontal source in the back [27], all subjects reported that the music came from the front.

3.2. Subjects and procedure

Nine subjects (Subjects A–I), 21–38 years old with normal hearing, participated in the experiment. Subjects B and G play the saxophone and piano as a hobby, respectively. All subjects were naive regarding the hypothesis of the experiment.

The subjects were seated in the anechoic chamber and were presented the sound stimuli. Two alternative forced choice (2AFC) tasks were performed for all combinations of the 12 stimuli, (i.e. 66 pairs [$N(N-1)/2$, $N=12$]) in random order in a session. Five sessions were conducted for each subject. The duration of the stimuli was 10 s including the linear rise and

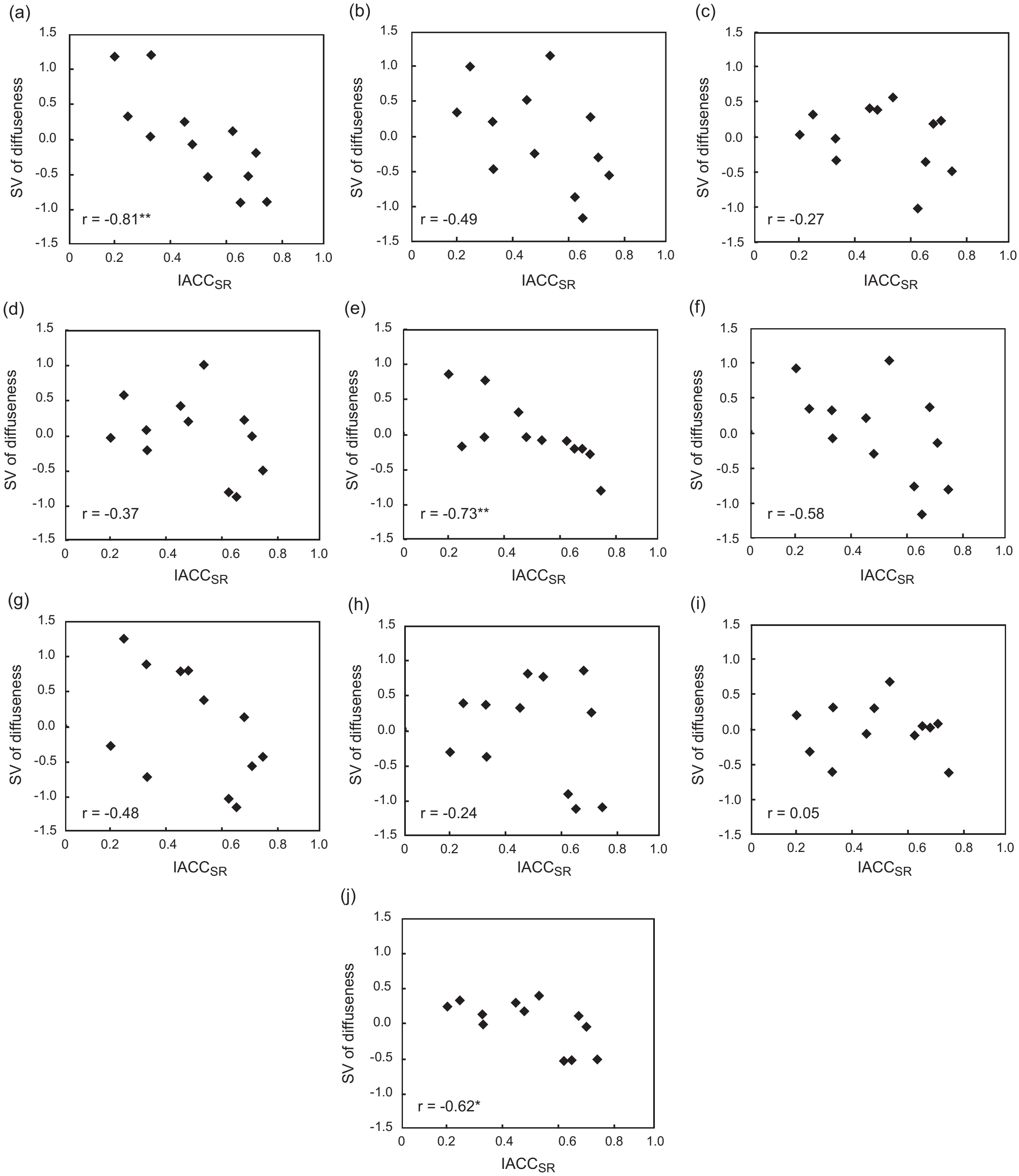

Fig. 7. Individual scale values of subjective diffuseness as a function of $IACC_{SR}$ of the stimuli for subjects: (a) A, (b) B, (c) C, (d) D, (e) E, (f) F, (g) G, (h) H, (i) I and (j) Total. *5%, **1% significant levels.

Table 5
Individual and total results of standardized regression coefficients for each IACF or ACF factors.

Subject	$a(\mathrm{IACC_{IR}})$	$b(\tau_e)$	$c(W_{\phi(0)})$	R
A	−0.49*	−0.66**	−0.35	0.87*
B	−0.88**	0.27	−0.05	0.92**
C	−0.69**	0.64**	0.15	0.94**
D	−0.81**	0.46*	−0.02	0.93**
E	−0.55*	−0.46	−0.38	0.79
F	−0.85**	0.14	−0.35*	0.94**
G	−0.71**	0.24	0.45*	0.87*
H	−0.60*	0.58*	−0.04	0.84*
I	−0.30	0.28	−0.53	0.69
Total	−0.90**	0.27	−0.05	0.94**

* Indicates 5% significant levels.
** Indicates 1% significant levels.

fall times of 50 ms. The silent interval between the stimuli was 2.0 s, and the interval between the pairs was 4.0 s, which was the time allowed for the subject to respond by marking an answer sheet. Subjects were asked to judge which of the two stimuli were more diffused [12]. The subjects were trained giving some examples before the experiment. On this occasion, they were instructed to image the position of the performer. If the subjects could not image the position clearly and perceived to be embedded in the sound, the experimenter explained that the stimulus was diffused. The experimenter encouraged them to evaluate the clarity of the sound image.

The scale values of subjective diffuseness were calculated according to Case V of Thurstone's theory [30]. The model of Case V for all data was confirmed by the goodness of fit test [31]. The higher scale value indicates that stimulus is perceived to be more diffused.

3.3. Results

Fig. 7 shows the scale value (SV) of subjective diffuseness as a function of the $\mathrm{IACC_{SR}}$ for the 12 stimuli. The SV had a negative correlation with the $\mathrm{IACC_{SR}}$. The correlation coefficients were significantly high for subjects A and E, while they were low for the other subjects. The $(W_{\mathrm{IACC}})_{\mathrm{SR}}$ and $(\tau_{\mathrm{IACC}})_{\mathrm{SR}}$ were not correlated with the subjective diffuseness.

To examine which factor of music signal affects the subjective diffuseness, the contributions of the ACF factors calculated from the music signals were investigated using multiple regression analysis. The explanatory variables were $\mathrm{IACC_{IR}}$, τ_e, and $W_{\phi(0)}$ as follows:

$$\mathrm{SV} = a\mathrm{IACC_{IR}} + b\tau_e + cW_{\varphi(0)} + d \tag{9}$$

Because the $(W_{\mathrm{IACC}})_{\mathrm{IR}}$, $(\tau_{\mathrm{IACC}})_{\mathrm{IR}}$, and $\Phi(0)$ were fixed as the experimental condition, they were excluded from the explanatory variables. Because τ_1 ($r = -0.71$, $p < 0.05$) and ϕ_1 ($r = 0.93$, $p < 0.01$) were highly correlated with τ_e, they were excluded. The standardized partial regression coefficients are listed in Table 5. The contribution of $\mathrm{IACC_{IR}}$ was significant for the eight subjects, and the regression coefficient a was negative for all subjects. In the factors of music signal, the contributions of τ_e and $W_{\phi(0)}$ were significant for four and two subjects, respectively. However, the regression coefficients b and c were different individually. For example, the subjects A and C perceived the music signal with shorter and longer τ_e more diffused, respectively. And the subjects F and G perceived the music signal with narrower and broader $W_{\phi(0)}$ more diffused, respectively. For the subjects B, E, and I, the subjective diffuseness could not be explained by the ACF factors.

4. Discussion

The $\mathrm{IACC_{SR}}$ can be estimated from the $\mathrm{IACC_{IR}}$ and τ_e of the source signal with high accuracy. The $\mathrm{IACF_{SR}}$ can be also estimated from the sum of IACF_{SR} of bandpass noises with the frequency weighting of the source signal [32]. Putting a loudspeaker in any direction, the $\mathrm{IACF_{SR}}$ of bandpass noises is previously measured by a dummy head. However, this method needs to measure the horizontal angle of each reflection, so it is difficult to apply the method for the measured BIR of a hall. From the measured BIR, several kinds of IACC were introduced to evaluate the spatial impression of a hall [33,34]. For example, the IACC_E is calculated from only the early part (within 80 ms) of the BIR, and the IACC_{E3} is calculated by averaging the values of the IACC_E in the three octave bands (0.5, 1 and 2 kHz). The IACC_{E3} is particularly correlated with the subjective rank-orderings of halls [33,34]. Although the IACC_{E3} is useful to compare the acoustical qualities of concert halls in which symphonic music is performed mainly, the $\mathrm{IACC_{SR}}$ estimated from the $\mathrm{IACC_{IR}}$ and τ_e of the sound source can compare the sound images of different performances in a same hall (e.g. singing and symphonic music in an opera house, or speech and organ music in a church).

To investigate the relationship between $\mathrm{IACC_{SR}}$ and subjective diffuseness, psycho-acoustical experiments were conducted using four music sources convolved with three BIRs. The results indicated that listeners perceive the sound image of music in a hall to be more diffused as the $IACC_{SR}$ decreases. This result is consistent with the results of previous experiments measuring the subjective diffuseness for bandpass noise sources [12,13]. Unlike the bandpass noise sources, the correlation between $IACC_{SR}$ and subjective diffuseness is not high for the music source as shown in Fig. 7 suggesting that the subjective diffuseness perceived by seven subjects is not explained by the $IACC_{SR}$.

The next point of interest is which factors affect subjective diffuseness. To investigate the contributions of IACF and ACF factors, the multiple regression analysis were carried out. As results, the standardized regression coefficient for $IACC_{IR}$ was negative and the largest observed ($a = -0.9$, $p < 0.01$), while the coefficients for the ACF factors (i.e. b and c) were different among the subjects. The standardized regression coefficients for the ACF factor suggest that the subjective diffuseness perceived by subjects A, C, D, and H was influenced by τ_e (Table 5). However, the contributions were not uniform among the subjects. Subject A perceived the music with shorter τ_e to be more diffused, while subjects C, D, and H perceived the music with longer τ_e to be more diffused. The judgments of subject A can be explained by the $IACC_{SR}$, as shown in Fig. 7(a). The music with shorter τ_e makes the $IACC_{SR}$ smaller, and the subject A perceives it to be more diffused. On the other hand, the judgments of subjects C, D, and H can be explained by the τ_e of sound source, which indicates its degree of reverberation included in the source. For example, when legato chordal music such as Melody B is played by an organ, the sound reverberates well, and it seems that the subjects C, D, and H perceive it to be more diffused. Since the stereo dipole presentation cannot control the sound level of lateral reflections, the LG and G_{LL}, which are related to the judgment of LEV [8,9], were the same in all the simulated sound fields. However, subjects C, D, and H may have perceived to be surrounded due to the reverberation in the source signal.

The results for multiple regression analysis suggest an effect of $W_{\phi(0)}$ on the subjective diffuseness perceived by subjects F and G (Table 5). $W_{\phi(0)}$ is related to the spectral centroid of a signal. The broad and narrow $W_{\phi(0)}$ mean that the low and high spectral components are strong, respectively. The low spectral components of the source signal are considered to be important in evaluating the ASW [14]. The W_{IACC} of a bandpass noise in the sound field becomes broader as the low spectral components of the noise increase, and the ASW can be explained by the W_{IACC} [4]. For the music, the $W_{\phi(0)}$ was highly correlated with the $(W_{IACC})_{SR}$ ($r=0.85$, $p < 0.01$). Subject G perceived the stimulus with the broader $W_{\phi(0)}$ and $(W_{IACC})_{SR}$ to be more diffused, so his perceived subjective diffuseness might be based on the ASW. However, subject F's judgment, that the music with narrow $W_{\phi(0)}$ was diffused, does not correspond to this explanation. Regarding the judgments of the other three subjects (B, E and I), the factors of ACF (i.e., τ_e and $W_{\phi(0)}$) are insufficient to explain the effect of the sound source on subjective diffuseness.

5. Conclusions

Taken together, our results indicate the following:

- The $IACC_{SR}$ can be estimated from the $IACC_{IR}$ of BIR and the τ_e of music signal with high accuracy.
- When the subjective diffuseness is evaluated based on BIR and music signal separately, the contributions of $IACC_{IR}$ from the BIR and τ_e or $W_{\phi(0)}$ from the music signal were significant for eight and for six in nine subjects, respectively.
- The contributions of τ_e or $W_{\phi(0)}$ show the individual differences.

The aim of this study is to clarify whether the subjective diffuseness in a hall depends on the music signal. In this study, the $(W_{IACC})_{IR}$, $(\tau_{IACC})_{IR}$, and reverberation time were constant in the all simulated sound fields, while the music signals with different ACF factors were used to find which ACF factors influence the subjective diffuseness. Although the experimental result indicated that the ACF factors, τ_e and $W_{\phi(0)}$, were influential on the subjective diffuseness, the contributions were individual among the subjects. This might be due to the fact that the $IACC_{SR}$ and $(W_{IACC})_{SR}$ change at the same time according to the τ_e and $W_{\phi(0)}$, respectively. In subsequent studies, music stimuli with similar $W_{\phi(0)}$ should be used to observe the unique contributions of $IACC_{SR}$ on the subjective diffuseness.

Acknowledgments

The authors thank the subjects who participated in the experimental sessions. The first author thanks Dr. Norio Emura who instructed him on the writing style of music scores. This work was supported by an academic grant from Bologna University in Italy and a Gant-in-Aid for Young Scientists (A) from the Japan Society for the Promotion of Science (18680025).

Appendix A. Definitions of IACF factors

The normalized IACF of a binaural signal is

$$\varphi_{lr}(\tau) = \frac{\Phi_{lr}(\tau)}{\sqrt{\Phi_{ll}(0)\Phi_{rr}(0)}} \tag{A1}$$

where

$$\Phi_{lr}(\tau) = \frac{1}{2T}\int_{-T}^{T} p'_l(t)p'_r(t+\tau)\,\mathrm{d}t \tag{A2}$$

in which, $p_l'(t)$ and $p_r'(t)$ are the left and right signals after passing through the A-weighting filter, τ is the time delay, $\Phi_{ll}(0)$ and $\Phi_{rr}(0)$ are the ACFs of the signals at τ=0, and 2T is the integral interval. The IACF factors are: (1) the amplitude of the maximum peak within the delay time 1 ms (IACC), (2) the interaural delay time of the maximum peak ($\tau_{\rm IACC}$), and (3) the width of the maximum peak ($W_{\rm IACC}$), defined by the interval of interaural delay time at a value 10% below the IACC. Definitions of the IACF factors are illustrated in Fig. A1.

Appendix B. Definitions of ACF factors

The normalized ACF of a signal is

$$\varphi(\tau) = \frac{\Phi(\tau)}{\Phi(0)}, \tag{B1}$$

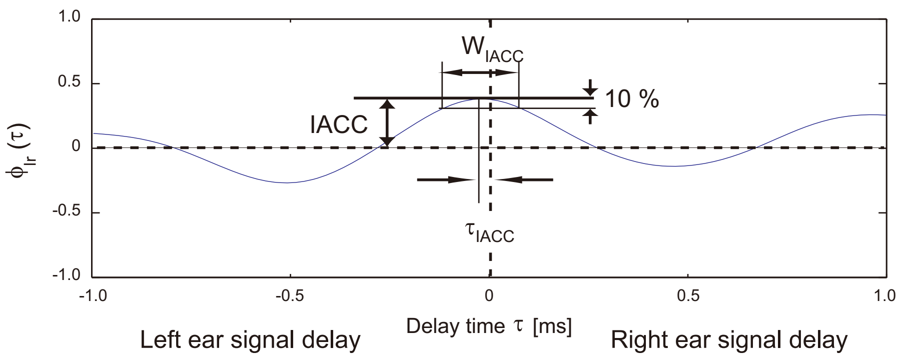

Fig. A1. Definition of IACF factors.

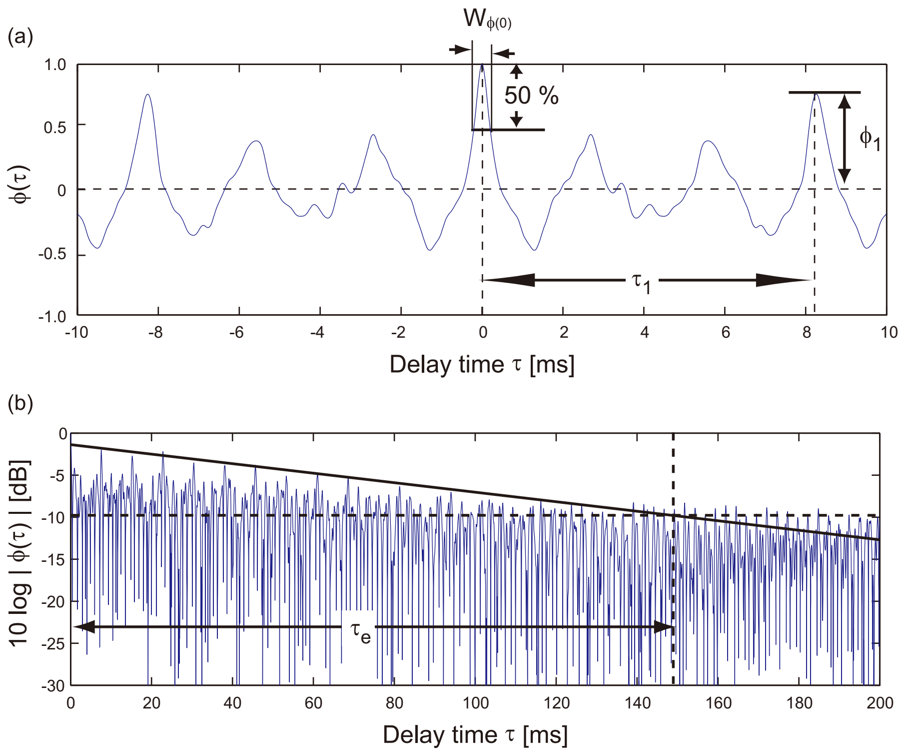

Fig. B1. Definition of ACF factors. (a) τ_1, ϕ_1, and $W_{\phi(0)}$; and (b) τ_e.

where

$$\Phi(\tau)=\frac{1}{2T}\int_{-T}^{T}p'(t)p'(t+\tau)\,\mathrm{d}t, \tag{B2}$$

where $2T$ is the integral interval, τ is the time delay, and $p'(t)$ is the signal after passing through the A-weighting filter. The ACF factors are (1) the energy represented at the origin of the delay [$\Phi(0)$], (2) the delay time of the maximum peak (τ_1), (3) the amplitude of the first maximum peak (ϕ_1), (4) the effective duration (τ_e), defined by the delay time at which the envelope along the early decay of the normalized ACF becomes -10 dB, and (5) the width of the peak at τ=0 ($W_{\phi(0)}$), defined by the double of delay time at which the normalized ACF becomes 0.5 at the beginning. Definitions of ACF factors are illustrated in Fig. B1.

References

[1] J.S. Bradley, G.A. Soulodre, The influence of late arriving energy on spatial impression, *Journal of the Acoustical Society of America* 97 (1995) 2263–2271.
[2] W.V. Keet, The influence of early lateral reflections on the spatial impression, *Proceedings of the Sixth International Congress on Acoustics,* Tokyo, E-2–4, 1968.
[3] Y. Ando, H. Sakai, S. Sato, Formulae describing subjective attributes for sound fields based on a model of the auditory-brain system, *Journal of Sound and Vibration* 232 (2000) 101–127.
[4] S. Sato, Y. Ando, Apparent source width (ASW) of complex noises in relation to the interaural cross-correlation function, *Journal of Temporal Design in Architecture and the Environment* 2 (2002) 29–32.
[5] Y. Ando, *Section 7.2.2. Apparent width of multiband noise*, *Auditory and Visual Sensations*, Springer-Verlag, New York, 2009, pp. 131–136pp. 131–136.
[6] M. Barron, The subjective effects of first reflections in concert halls: the need for lateral reflections, *Journal of Sound Vibration* 15 (1971) 475–494.
[7] M. Barron, A.H. Marshall, Spatial impression due to early lateral reflections in concert halls: the derivation of a physical measure, *Journal of Sound Vibration* 77 (1981) 211–232.
[8] J.S. Bradley, Comparison of concert hall measurements of spatial impression, *Journal of the Acoustical Society of America* 96 (1994) 3525–3535.
[9] J.S. Bradley, R.D. Reich, S.G. Norcross, On the combined effects of early- and late-arriving sound on spatial impression in concert halls, *Journal of the Acoustical Society of America* 108 (2000) 651–661.
[10] P. Damaske, Subjektive Untersuchungen von Schallfeldern, *Acustica* 19 (68) (1967) 199–213.
[11] P. Damaske, Y. Ando, Interaural cross-correlation for multichannel loudspeaker reproduction, *Acustica* 27 (1972) 232–238.
[12] Y. Ando, Y. Kurihara, Nonlinear response in evaluating the subjective diffuseness of sound fields, *Journal of the Acoustical Society of America* 80 (1986) 833–836.
[13] P.K. Singh, Y. Ando, Individual subjective diffuseness responses of filtered noise sound fields, *Acustica* 80 (1994) 471–477.
[14] T. Okano, L.L. Beranek, T. Hidaka, Relations among interaural cross-correlation coefficient ($IACC_E$), lateral fraction (LF_E), and apparent source width (ASW) in concert halls, *Journal of the Acoustical Society of America* 104 (1998) 255–265.
[15] Y. Ando, *Architectural Acoustics—Blending Sound Sources, Sound Fields, and Listeners*, AIP Press/Springer-Verlag, New York, 1998.
[16] Y. Ando, Correlation factors describing primary and spatial sensations of sound fields, *Journal of Sound Vibration* 258 (2002) 405–417.
[17] L.L. Beranek, *Concert and Opera Halls: How They Sound*, Acoustical Society of America, New York, 1996.
[18] Y. Ando, Subjective preference in relation to objective parameters of music sound fields with a single echo, *Journal of the Acoustical Society of America* 62 (1977) 1439–1441.
[19] Y. Ando, H. Alrutz, Perception of coloration in sound fields in relation to the autocorrelation function, *Journal of the Acoustical Society of America* 71 (1982) 616–618.
[20] Y. Soeta, R. Shimokura, Comparison of noise characteristics in airplanes and high-speed trains, *Journal of Temporal Design in Architecture and the Environment* 9 (2009) 22–25.
[21] M.R. Schroeder, Digital simulation of sound transmission in reverberant space, *Journal of the Acoustical Society of America* 47 (1970) 424–431.
[22] B.B. Bauer, Stereophonic earphones and binaural loudspeakers, *Journal of the Audio Engineering Society* 9 (1961) 148–151.
[23] O. Kirkeby, P.A. Nelson, H. Hamada, The "Stereo Dipole"—a virtual source imaging system using two closely spaced loudspeakers, *Journal of the Audio Engineering Society* 46 (1998) 387–395.
[24] M. Morimoto, K. Iida, A practical evaluation method of auditory source width in concert halls, *Journal of the Acoustical Society of Japan* 16 (1995) 59–69.
[25] P.A. Nelson, O. Kirkeby, T. Takeuchi, H. Hamada, Sound fields for the production of virtual acoustic images, *Journal of Sound Vibration* 204 (1997) 386–396.
[26] O. Kirkeby, P.A. Nelson, H. Hamada, Local sound field reproduction using two closely spaced loudspeakers, *Journal of the Acoustical Society of America* 104 (1998) 1973–1981.
[27] T. Takeuchi, P.A. Nelson, Robustness to head misalignment of virtual sound imaging systems, *Journal of the Acoustical Society of America* 109 (2001) 958–971.
[28] J. Rose, P. Nelson, B. Rafaely, T. Takeuchi, Sweet spot size of virtual acoustic imaging systems at asymmetric listener locations, *Journal of the Acoustical Society of America* 112 (2002) 1992–2002.
[29] N. Prodi, S. Velecka, The evaluation of binaural playback systems for virtual sound fields, *Applied Acoustics* 64 (2003) 147–161.
[30] L.L. Thurstone, A low of comparative judgment, *Psychological Review* 31 (1927) 273–289.
[31] F. Mosteller, Remarks on the method of paired comparisons. III, *Psychometrika* 16 (1951) 207–218.
[32] T. Nakajima, J. Yoshida, Y. Ando, A simple method of calculating the interaural cross-correlation function for a sound field, *Journal of the Acoustical Society of America* 93 (1993) 885–891.
[33] T. Hidaka, L.L. Beranek, T. Okano, Interaural cross-correlation, lateral fraction, and low- and high-frequency sound levels as measures of acoustical quality in concert halls, *Journal of the Acoustical Society of America* 98 (1995) 988–1007.
[34] T. Hidaka, L.L. Beranek, Objective and subjective evaluations of twenty-three opera houses in Europe, Japan, and the Americans, *Journal of the Acoustical Society of America* 107 (2000) 368–383.