Journal of Sound and Vibration (2001) **241**(1), 3–18
doi:10.1006/jsvi.2000.3272, available online at http://www.idealibrary.com on IDEAL®

A THEORY OF PRIMARY SENSATIONS AND SPATIAL SENSATIONS MEASURING ENVIRONMENTAL NOISE

Y. Ando

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

(*Accepted* 7 *August* 2000)

A theory of primary sensations and spatial sensations to environmental noise is proposed here. The model of the auditory–brain system includes the autocorrelation function (ACF), the interaural cross-correlation function (IACF) mechanisms and the specialization of the human cerebral hemispheres [1]. In addition to primitive sensations to the noise source-loudness, pitch and timbre-perception of duration is introduced here as a fourth. They are described by the temporal factors extracted from the ACF. It is worth noting that the information in the power density spectrum of a signal is identical to the ACF of the signal. In order to describe spatial sensations to the noise field, the spatial factors extracted from the IACF are taken into account. Timbre is defined as an overall sensation including the primitive sensations and spatial sensations.

1. INTRODUCTION

Fundamental subjective attributes for sound fields are well described by a model of the auditory–brain system. It includes autocorrelation function (ACF) and interaural cross-correlation function (IACF) mechanisms as shown in Figure 1 [1, 2]. Important mechanisms in this model were discovered in relation to the auditory–brain activity [1]. It is discussed that primitive and spatial sensations are described by temporal and spatial factors extracted from the ACF and the IACF respectively. In addition to primitive sensations to the noise source—loudness, pitch and timbre—perception of duration is introduced here as a fourth.

The speech intelligibility of spoken syllable and the delay time of a single reflection of sound fields can be expressed by temporal factors extracted from the ACF [3, 4]. Based on an auditory–brain system with two cerebral hemispheres, a theory of measuring environmental noise may be proposed.

2. PRIMARY SENSATIONS ARE MUTUALLY INDEPENDENT?

Let c_i ($i = 1, 2, \dots, I$) be physical factors representing cues influencing any primary sensations S_j ($j = 1, 2, \dots, J < I$), a sensation s_j may then be expressed by

$$s_j = f(c_1, c_2, \dots, c_i), \quad j = 1, 2, \dots, J. \tag{1}$$

If physical factors are orthogonal, then s_j may be expressed by a linear combination such that

$$s_j = f(c_1) + f(c_2) + \cdots + f(c_i), \quad j = 1, 2, \dots, J. \tag{2}$$

0022-460X/01/110003 + 16 $35.00/0

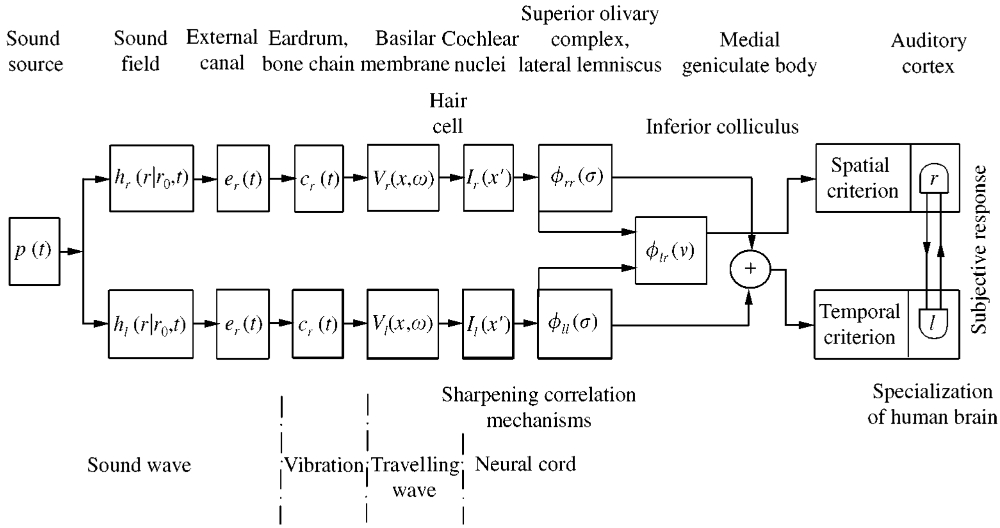

Figure 1. Model of the auditory–brain system with autocorrelation and interaural cross-correlation mechanisms and specialization of human cerebral hemispheres [1].

Here a question arises as to whether or not the sensations are independent of each other. For example, it is easily demonstrated by the simplest case, such as

$$s_j = f_j(c_1) + f_j(c_2), \quad s_k = f_k(c_1) + f_k(c_2), \tag{3}$$

that the correlation coefficient between s_j and s_k is given by

$$r_{jk} = \overline{s_j s_k} = f_j(c_1) f_k(c_1) + \overline{f_j(c_2) f_k(c_2)} + \overline{f_j(c_1) f_k(c_2)} + \overline{f_j(c_2) f_k(c_1)}, \tag{4}$$

and is not zero, because the first and second terms on the right-hand side are not always zero. Therefore, primitive sensations are subject to being not mutually independent. Particularly, timbre may not be independent from the loudness and the pitch as discussed below. Thus, timbre must be defined as an overall sensation, similar to the subjective preference of the sound field [1].

3. PRIMARY SENSATIONS IN RELATION TO FACTORS EXTRACTED FROM THE LONG-TERM ACF

3.1. THE LONG-TERM ACF AND FACTORS

Using the auditory–brain model (Figure 1), we now consider primary sensations of a given source signal $p(t)$ located in front of a listener in a free field. The long-term ACF is defined by

$$\Phi_p(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{+T} p'(t) p'(t + \tau)\, \mathrm{d}t, \tag{5}$$

where $p'(t) = p(t) * s(t)$, $s(t)$ being the ear sensitivity. For convenience, $s(t)$ may be chosen as the impulse response of an A-weighted network [1, 2]. The power density spectrum can also be obtained from the ACF and is defined by

$$P_d(\omega) = \int_{-\infty}^{+\infty} \Phi_p(\tau) \mathrm{e}^{-\mathrm{j}\omega\tau} \,\mathrm{d}\tau, \tag{6}$$

with

$$\Phi_p(\tau) = \int_{-\infty}^{+\infty} P_{\mathrm{d}}(\omega) \mathrm{e}^{\mathrm{j}\omega t} \,\mathrm{d}t. \tag{7}$$

Thus, the ACF and the power density spectrum mathematically contain the same information.

In the ACF analysis, there are three significant items: (1) energy represented at the origin of the delay, $\Phi_p(0)$; (2) effective duration of the envelope of the normalized ACF, τ_e, which is defined by the 10-percentile delay and which represents a repetitive feature or reverberation containing the noise source itself as shown in Figure 2; the normalized ACF is defined by $\phi_p(\tau) = \Phi_p(\tau)/\Phi_p(0)$; (3) fine structure, including peaks and dips and delays (see Figure 2(a)): for instance, τ_1 and ϕ_1 are the delay time and the amplitude of the first peak of ACF, τ_n and ϕ_n being the delay time and the amplitude of the nth peak. However, there are certain correlations between τ_n and τ_{n+1}, and between ϕ_n and ϕ_{n+1}. Thus, orthogonal and temporal factors that can be extracted from the ACF are $\Phi_p(0)$, τ_1, ϕ_1, and in addition, the effective duration of ACF, τ_e. In a manner as shown in Figure 2(b), this value is obtained by fitting a straight line for extrapolation of delay time at -10 dB, if the initial envelope of ACF decays exponentially.

3.2. LOUDNESS

Let us now consider primary sensations. Loudness s_L is given by [1, 5, 6]

$$s_L = f_L(\Phi(0), \tau_1, \phi_1, \tau_e), \tag{8}$$

where the value of τ_1 corresponds to pitch of the noise and/or the missing fundamental as discussed below.

When $p'(t)$ is measured with reference to the pressure 20 μPa leading to the level $L(t)$, the equivalent sound pressure level L_{eq}, defined by

$$L_{eq} = 10 \log \frac{1}{T} \int_0^T 10^{L(t)/10} \,\mathrm{d}t, \tag{9}$$

corresponds to $10 \log [\Phi_p(0)/\Phi_{ref}(0)]$, $\Phi_{ref}(0)$ being the reference energy. Since the sampling frequency of the sound wave is more than twice the maximum audio frequency, this value is much more accurate than the L_{eq} which is measured by the usual sound level meter.

Scale values of loudness within the critical band that were obtained in paired-comparison tests (with filters with the slope of 1080 dB/octave) under the condition of a constant $\Phi_p(0)$ are shown in Figure 3 [1, 5, 6]. Obviously, when a noise has a similar repetitive feature, τ_e becomes a great value, as for a pure tone, and then greater loudness results. Thus, a plot of loudness versus bandwidth is not flat in the critical band. This contradicts previous results of the frequency range centered on 1 kHz [7].

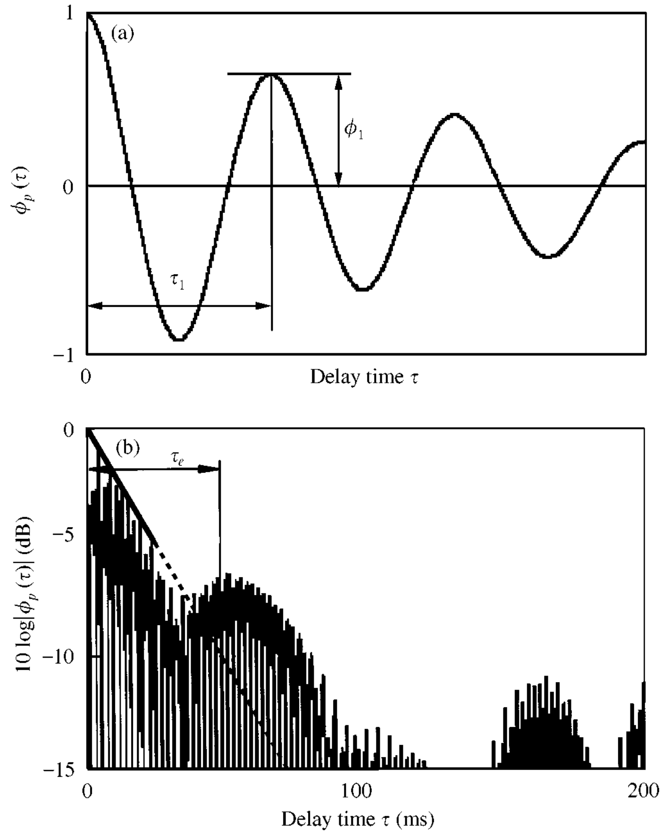

Figure 2. Definition of independent factors other than $\Phi(0)$ extracted from the normalized ACF. Values of τ_1 and ϕ_1 for the first peak (a); the effective duration of the ACFτ_e, which is defined by 10 percentile delay (at -10 dB) and which is obtained practically by the extrapolation of the envelope of the normalized ACF during the ecay, 5 dB initial (b).

3.3. PITCH

The second primary sensation relevant to the ACF is the pitch or the missing fundamental of the noise. It is given by [6, 8–10]

$$s_P = f_P(\tau_1, \phi_1). \tag{10}$$

When a sound signal contains only a number of harmonics without the fundamental frequency, one hears the fundamental as a pitch [3, 4, 9, 10]. This phenomenon is well explained by the delay time of the first peak in the ACF fine structure, τ_1. Also, according to experimental results on the pitch perceived when listening to bandpass noises without any fundamental frequency, the pitch s_p is also expressed by equation (10). The strength of the pitch sensation is described by the magnitude of the first peak of the ACF, ϕ_1, as mentioned below.

The bandwidths of each partial noise, which consist of the bandpass white noise with a cut-off slope of 1080 dB/Octave were changed. The center frequencies of the band noise components were 600, 800, 1000, 1200 and 1400 Hz. Here, the complex signals including bandpass noises with different center frequencies are called "complex noise". The bandwidths (Δf) of the four components were 40, 80, 120, and 160 Hz (see Figure 4(a)).

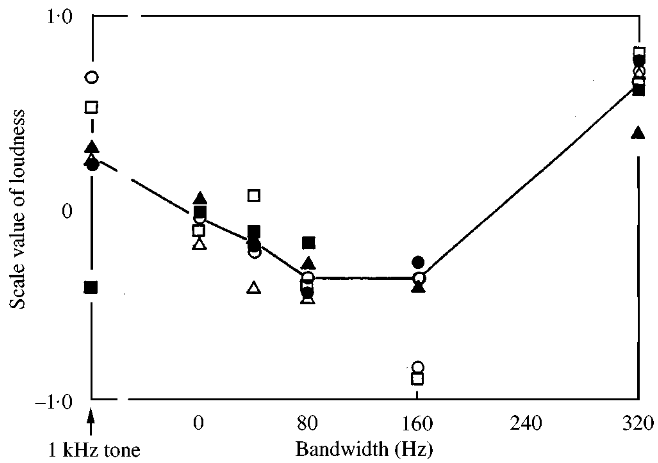

Figure 3. Scale values of the loudness of bandpass noises as a function of its bandwidth centered on 1 kHz. The cut-off slope of the filter used was 1080 dB/Oct. Different symbols indicate the scale values obtained with different subjects (six subjects).

Their waveforms without any specific periodical envelopes are shown in Figure 4(b)–4(e). The ACFs for four conditions are shown in Figure 4(b′)–4(e′). The amplitude of the peak (indicated by arrows in the figures) in the ACF increases with decreasing Δf. Five musicians (two male and three female, 20–25 years old) participated as subjects in this experiment.

The probabilities of the matching data counted for each 1/12-octave band are shown in Figure 5(a)–5(d). All histograms show that there is a strong tendency to perceive a pitch of 200 Hz for any stimulus. This agrees with the prediction based on the value of τ_1. The results in Figure 5(a)–5(d) indicate that a stimulus with a narrow bandwidth gives a stronger pitch corresponding to 200 Hz than does a stimulus with a wide bandwidth. The standard deviation (SD) for the perceived pitches increased because the value of ϕ_1 decreased as Δf increased. The probabilities matched to 400 Hz (one octave higher than 200 Hz) keep increasing as the bandwidth becomes narrow. This is caused by the similarity of the octave relation under the pitch perception which also appears in the previous experiment with the complex tone [3]. The probability of a pitch around 200 Hz being identified is plotted in Figure 6 as a function of ϕ_1. In this figure, the pitch-matching result from the previous experiment is also plotted at $\phi_1 = 1{\cdot}0$. For narrower-band noise, the probability of a pitch of the fundamental frequency increases as the magnitude of the 5-ms peaks in the ACF increases. Thus, as the ϕ_1 increases the pitch strength increases ($r = 0{\cdot}98$). This result also supports Wightman's theory [8].

Of particular interest is a discovery in the vision research that the pitch or subjective flicker rate with frequency components above 30 Hz can be described by the value of τ_1 also [11].

3.4. TIMBRE

The third primary sensation, timbre, is the most complicated, because it includes pitch and loudness:

$$s_T = f_T\,[\Phi(0),\ \tau_e,\ (\tau_1,\ \phi_1),\ \dots,\ (\tau_n,\ \phi_n)]. \tag{11}$$

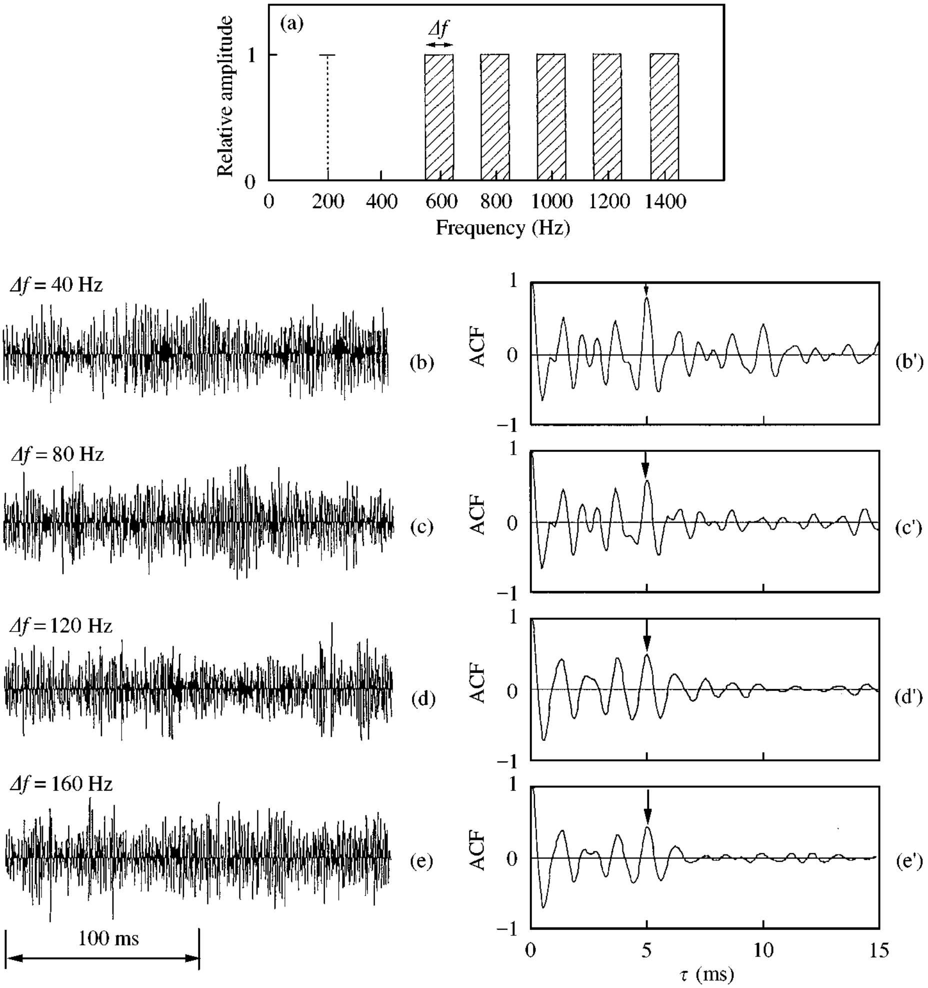

Figure 4. "Complex noise" containing the center frequencies: 600, 800, 1000, 1200 and 1400 Hz (a); its fundamental frequency is around 200 Hz. The Δf represents the bandwidth. Waveforms of the four complex noises applied: $\Delta f = 40$ Hz (b); $\Delta f = 80$ Hz (c); $\Delta f = 120$ Hz (d); and $\Delta f = 160$ Hz (e). The ACFs of the stimuli: $\Delta f = 40$ Hz (b′); $\Delta f = 80$ Hz (c′); $\Delta f = 120$ Hz (d′); and $\Delta f = 160$ Hz (e′).

since (τ_1, ϕ_1) are the most significant orthogonal factors in (τ_n, ϕ_n), $n = 1, 2, \ldots,$ equation (11) is rewritten as

$$s_T = f_T[\Phi(0), \tau_e, \tau_1, \phi_1]. \tag{12}$$

At present, there are no experimental results demonstrating timbre in relation to these temporal factors.

3.5. DURATION SENSATION

The fourth primitive sensation, which is introduced here, is the perception of signal duration, which is given by [12, 13]

$$s_D = f_D[\Phi(0), \tau_e, \tau_1, \phi_1]. \tag{13}$$

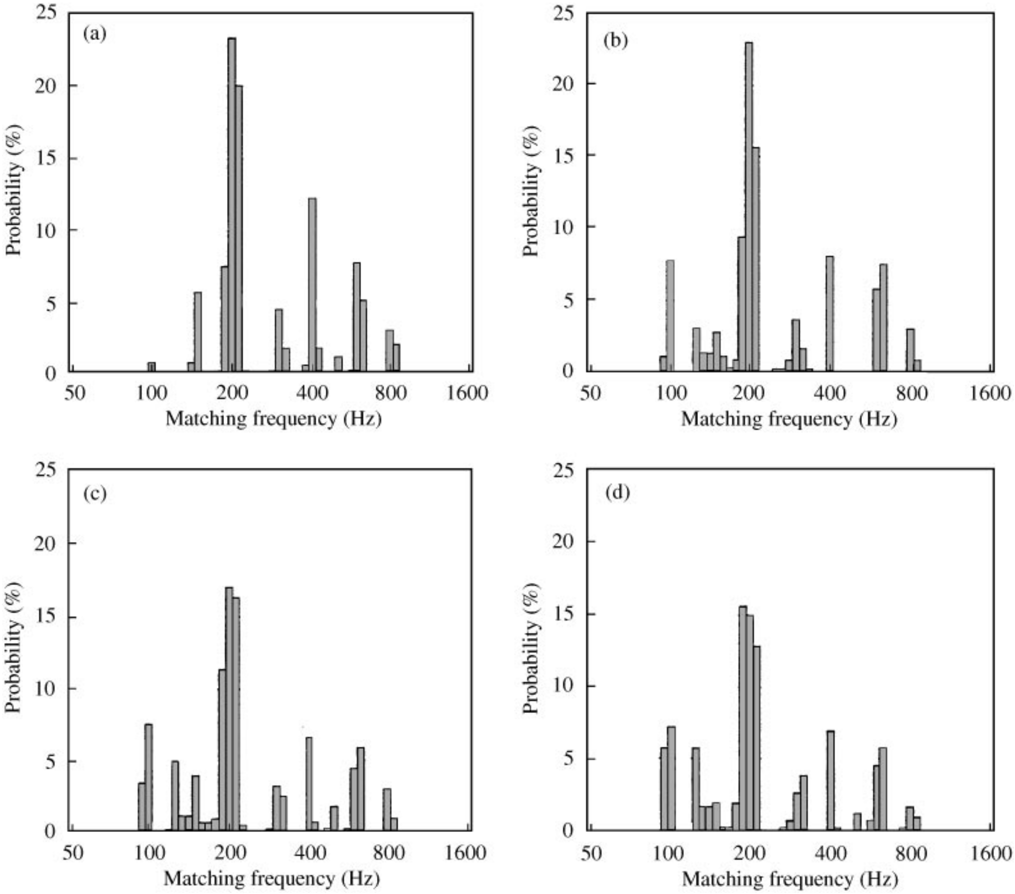

Figure 5. Results of the pitch-matching tests for the global data of five subjects: $\Delta f = 40$ Hz (a); $\Delta f = 80$ Hz (b); $\Delta f = 120$ Hz (c); and $\Delta f = 160$ Hz (d).

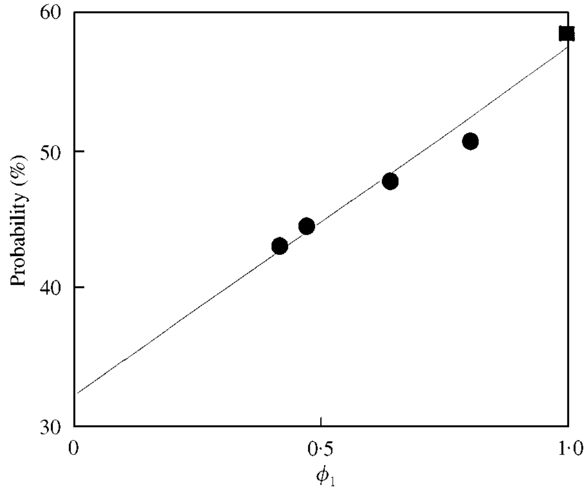

Figure 6. Relationship between ϕ_1 and probability of the pitch being within 200 ± 16 Hz ($r = 0{\cdot}98$, $p < 0{\cdot}01$). Just for reference, the plot (■) at $\phi_1 = 1{\cdot}0$ is the result of the first tests.

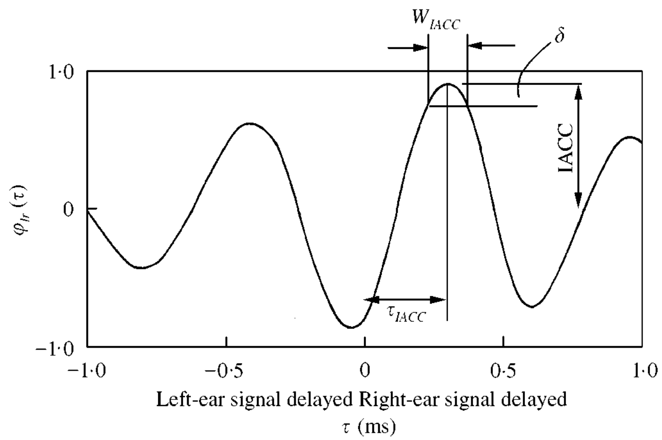

Figure 7. Definition of independent factors IACC, τ_{IACC} and W_{IACC} extracted from the normalized IACF.

4. SPATIAL SENSATIONS IN RELATION TO FACTORS EXTRACTED FROM THE LONG-TERM IACF

4.1. LOCALIZATION OF NOISE SOURCE

The long-term IACF is given by

$$\Phi_{lr}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{+T} p'_l(t)\, p'_r(t+\tau)\,\mathrm{d}t, \tag{14}$$

where $p'_{l,r}(t) = p(t)_{l,r} * s(t)$, $p(t)_{l,r}$ being the sound pressure at the left- and right-ear entrances. Spatial factors extracted from the IACF are defined in Figure 7 [1].

The perceived direction of a noise source in the horizontal plane is assumed to be described as

$$s = f(LL, \text{IACC}, \tau_{IACC}, W_{IACC}), \tag{15}$$

where

$$LL = 10\log_{10}\{\Phi_{ll}(0), \Phi_{rr}(0)\}. \tag{16}$$

The symbol { } signifies a set, $\Phi_{ll}(0)$ and $\Phi_{rr}(0)$ being ACFs at $\tau = 0$ (sound energies), of the signals arriving at the left- and right-ear entrances. Mathematically, LL is expressed by the geometrical mean of the energies of sound signals arriving at both the ear entrances. The listening level is given by

$$LL = 10\log_{10}\frac{\sqrt{\Phi_{ll}(0)\,\Phi_{rr}(0)}}{\Phi_{ref}(0)} \text{ (dB)}. \tag{17}$$

In these four orthogonal factors in equation (15), the τ_{IACC} defined within the possible interaural delay time is a significant factor in determining the perceived horizontal direction of the source. A well-defined direction is perceived when the normalized interaural cross-correlation function has one sharp maximum, a high value of the IACC and a narrow value of the W_{IACC}, due to high-frequency components. On the other hand, subjective diffuseness or no spatial directional impression corresponds to a low value of IACC ($< 0{\cdot}15$) [14].

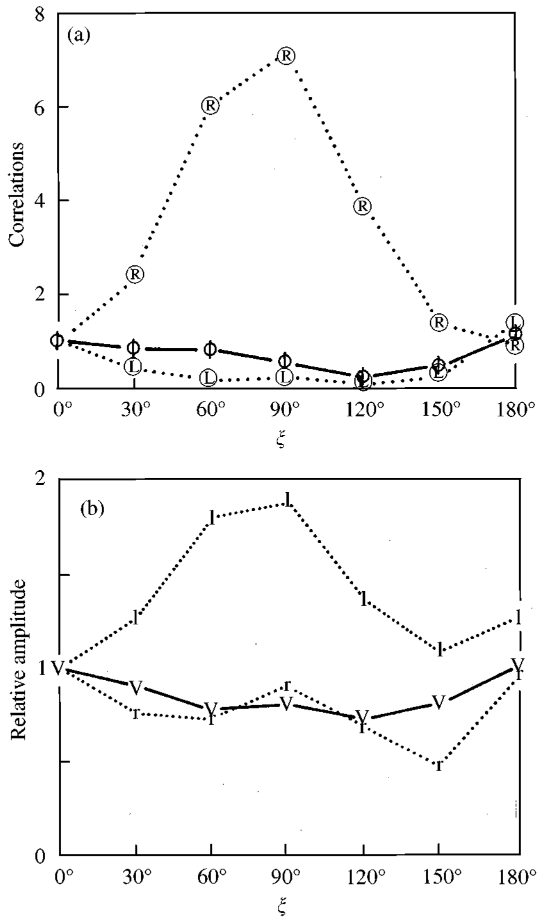

Figure 8. Evidence for auditory brainstem responses related to the IACF. L: $\Phi_{ll}(0)$ measured at the left-ear; R: $\Phi_{rr}(0)$ measured at the right-ear (a); Φ: Maximum interaural cross-correlation, $|\Phi_{lr}(\tau)|_{max}$, $|\tau| < 1$ ms. Averaged amplitudes of waves IV_l (l) and $IV_r(r)$, and averaged amplitudes of waves V_l and $V_r(V)$ normalized to the amplitudes at the frontal incidence, four subjects (b).

For the perception of a noise source located in the median plane, the temporal factors extracted from the long-term ACF of sound signal arriving at the ear entrances should be added into equation (15) [15].

As shown in Figure 8, a remarkable finding is that there are neural activities at the inferior colliculus corresponding to the IACC and sound energies for sound signals that arrive at the two ear entrances [16]. Also, it is discovered that the *LL* and the IACC are dominantly associated with the right cerebral hemisphere, and the temporal factors, Δt_1 and T_{sub}, the sound field in a room are associated with the left [2]. These support the model shown in Figure 1.

4.2. SUBJECTIVE DIFFUSENESS OF NOISE FIELD

The scale value of subjective diffuseness is assumed to be given by equation (15) also. In order to obtain the scale value of subjective diffuseness, paired-comparison tests with

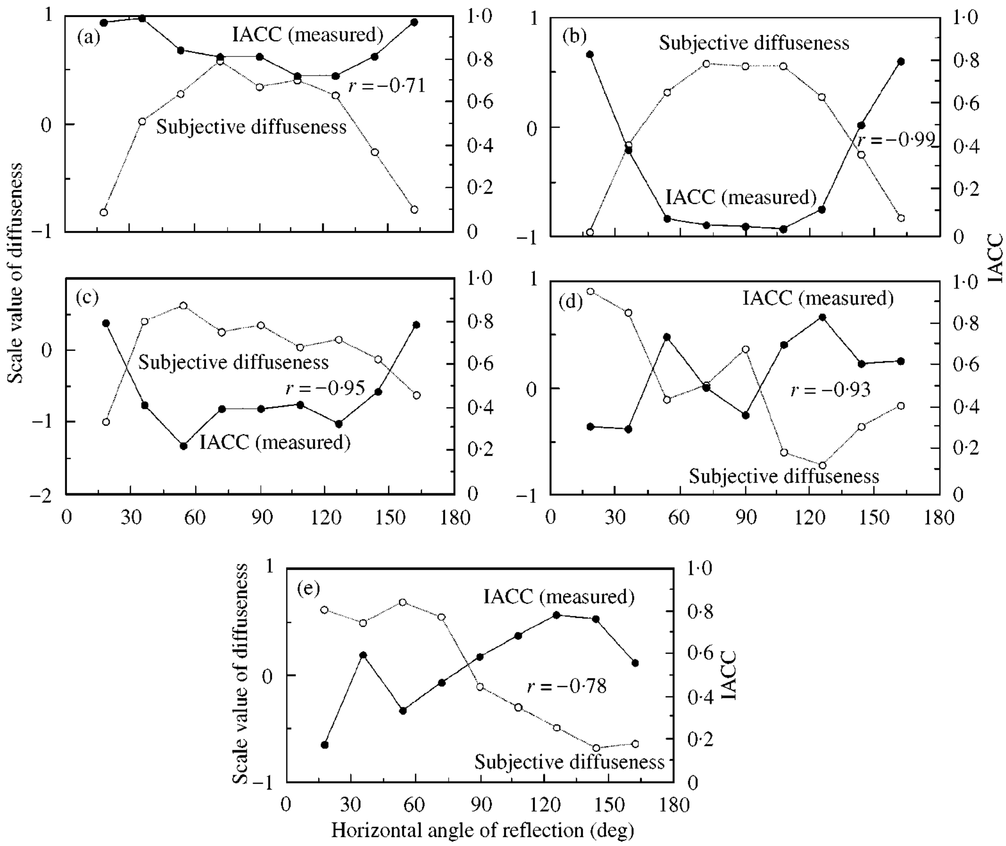

Figure 9. Scale values of subjective diffuseness and the IACC as a function of the horizontal angle of incidence to a listener, with $\frac{1}{3}$ octave band noise of center frequencies. 250 Hz (a); 500 Hz (b); 1 kHz (c); 2 kHz (d); and 4 kHz (e).

bandpass Gaussian noise, varying the horizontal angle of two symmetric reflections have been conducted. Listeners judged which of two sound fields were perceived as more diffuse, under the constant conditions of LL, τ_{IACC}, and W_{IACC} [17]. The strong negative correlation between the scale value and the IACC can be found in the results with frequency bands between 250 Hz and 4 kHz, as shown in Figure 9. Scale values of subjective diffuseness are inversely proportional to the IACC as shown in Figure 10. The scale value of subjective diffuseness may be well formulated in terms of the $\frac{3}{2}$ power of the IACC in a manner similar to the subjective preference for the sound field, i.e.,

$$S_{\text{diffuseness}} \approx -\alpha(\text{IACC})^{\beta}, \tag{18}$$

where coefficients $\alpha = 2{\cdot}9$ and $\beta = \frac{3}{2}$.

4.3. APPARENT SOURCE WIDTH OF NOISE FIELD

It is considered that the scale value of apparent source width (ASW) is given by equation (15) as well. For a noise field with a predominate low-frequency range, the long-term IACF has no sharp peaks for the delay range of $|\tau| < 1$ ms, and W_{IACC} becomes wider. The values

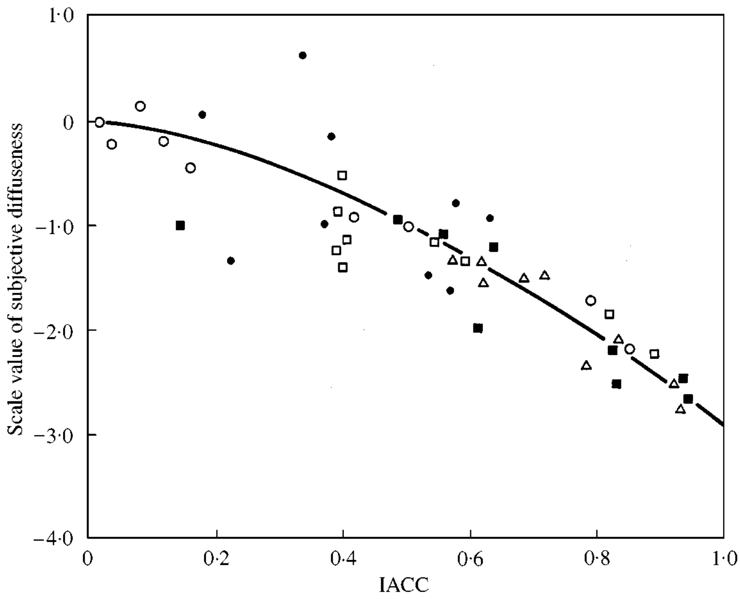

Figure 10. Scale values of subjective diffuseness as a function of the IACC calculated. Different symbols different frequencies of the $\frac{1}{3}$ octave-bandpass noise: indicate (△) 250 Hz; (○) 500 Hz; (□) 1 kHz; (●) 2 kHz; and (■) 4 kHz: (—) Regression line by equation (18).

are theoretically calculated by the equation [1, 3, 4]

$$W_{IACC} \approx \frac{4}{\Delta\omega_c} \cos^{-1}\left(1 - \frac{\delta}{\mathrm{IACC}}\right)(\mathrm{s}), \tag{19}$$

where $\Delta\omega_c = 2\pi(f_1 + f_2)$, and f_1 and f_2 are the lower and upper frequencies of an ideal filter. For simplicity, δ is defined by 0·1(IACC).

Of particular interest is that a wider ASW may be perceived with low-frequency bands and by decreasing the IACC. More clearly, the ASW may be well described by both factors, IACC and W_{IACC} [3, 4], under the conditions of a constant LL and $\tau_{IACC} = 0$. The scale values of ASW were obtained by paired-comparison tests with ten subjects. In order to control the values of W_{IACC}, the center frequencies of $\frac{1}{3}$ octave bandpass noises were changed. Its center frequencies were 250, 500 Hz, 1, and 2 kHz. The values of IACC were adjusted by controlling the sound pressure ratio of the reflections to the level of the direct sound. The listening level affects ASW [18]: therefore, the total sound pressure levels at the ear canal entrances of all noise fields were kept constant at a peak of 75 dB(A). Listeners judged which of two noise sources they perceived to be wider. The results of the analysis of variance for the scale values s_{ASW} indicate that both of the factors IACC and W_{IACC} are significant ($p < 0{\cdot}01$), and contribute to the s_{ASW} independently; thus,

$$s_{ASW} \approx a(\mathrm{IACC})^{3/2} + b(W_{IACC})^{1/2}, \tag{20}$$

where coefficients $a \approx -1{\cdot}64$ and $b \approx 2{\cdot}44$ are obtained by regressions of the scale values with ten subjects as shown in Figure 11(a) and 11(b) respectively. Calculated scale values s_{ASW} by equation (20) and measured scale values are in good agreement (see Figure 12; $r = 0{\cdot}97$, $p < 0{\cdot}01$).

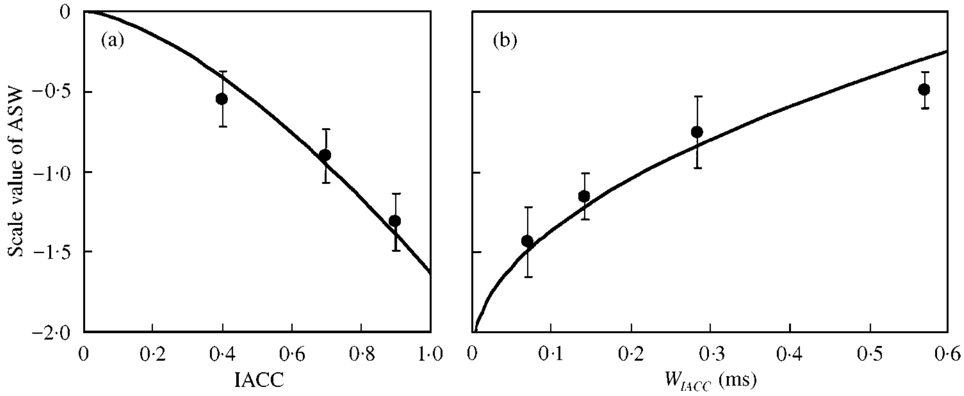

Figure 11. Average scale values of ASW for $\frac{1}{3}$ octave-bandpass noises with 95% reliability, as a function of the IACC (a). The regression curve is expressed by the first term of equation (20) with $a \approx -1{\cdot}64$. Average scale values of ASW for $\frac{1}{3}$ octave-bandpass noises with 95% reliability, as a function of W_{IACC} (b). The regression curve is expressed by the second term of equation (20) with $b \approx 2{\cdot}44$.

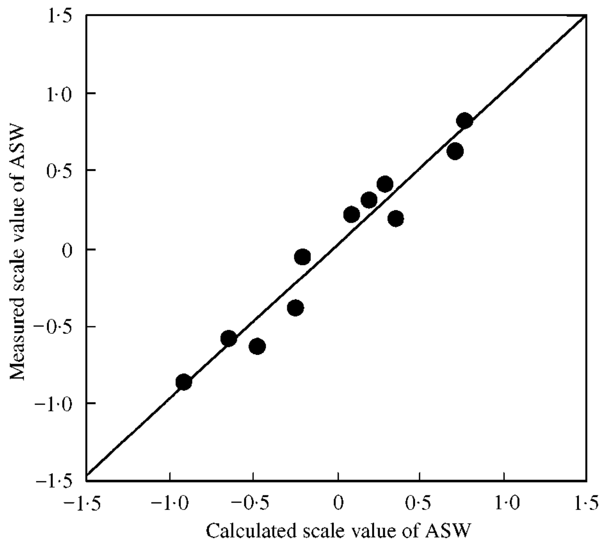

Figure 12. Relation between the measured scale values of ASW and the scale values calculated by equation (20) for the noise ($r = 0{\cdot}97$, $p < 0{\cdot}01$).

5. PRIMARY SENSATIONS OF THE FLUCTUATING-ENVIRONMENTAL NOISE

5.1. SHORT-TERM ACF

To assess fluctuating-environmental noise, we use the running short-time ACF as well as the short-time IACF. Running short-time spatial and temporal factors extracted in a similar manner to the above are used to describe the primitive sensations of a fluctuating-noise field.

The short-time ACF is defined by

$$\Phi_p(\tau) = \frac{1}{2T}\int_{-T}^{+T} p'(t)p'(t+\tau)\,\mathrm{d}t, \tag{21}$$

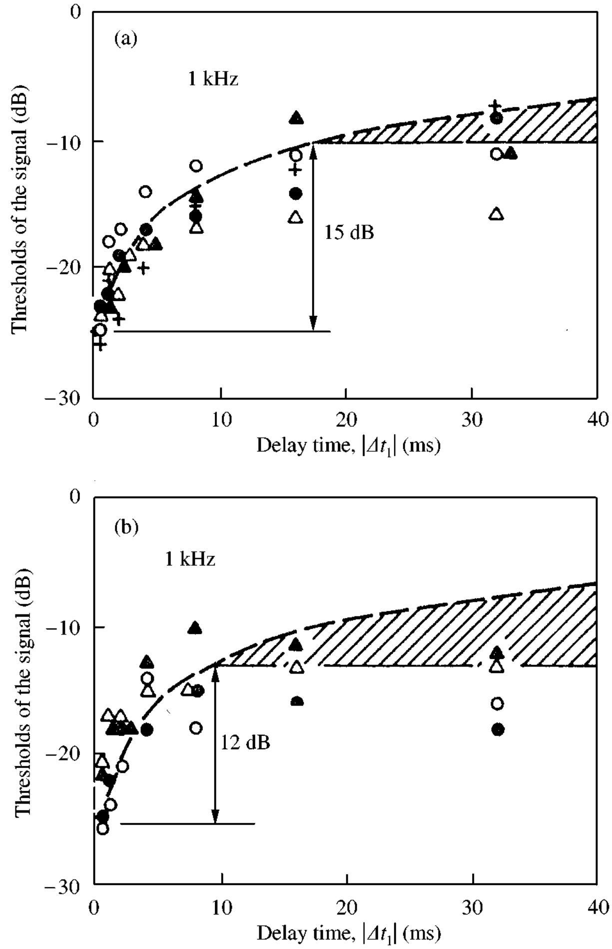

Figure 13. Thresholds of the single weak reflection as a function of its delay time Δt_1 [20]. For the reflection from the front (a); for the reflection from the side (b). The dashed curve shows the threshold obtained from the ACF envelope of the source signal of the band-limited noise centered at 1 kHz. The open and filled symbols show the threshold values judged by two subjects.

where $2T$ is determined by the signal duration to be analyzed. This duration $2T$ should be selected covering at least the minimum value of effective duration, $(\tau_e)_{min}$, of the running ACF [19]. The noise piece indicating $(\tau_e)_{min}$ contains the most rapid movement in the signal; thus, this particular piece influences most greatly the subjective responses.

5.2. REMARKS OF PRIMARY SENSATIONS

(i) For the loudness s_L of each noise piece, equation (8) can be replaced by

$$s_L = f_L(LL, \tau_1, \phi_1, \tau_e). \tag{22}$$

Here each factor is obtained for each noise piece, and the $\Phi(0)$ in equation (8) has been replaced by LL.

It is worth noticing that the temporal factors extracted from the ACF must be influenced by repetitive reflections (Δt_1, Δt_2, …) and the subsequent reverberation time (T_{sub}) in a room.

(ii) In describing the pitch of the environmental noise field, the significant temporal factors of the noise field are τ_1 and ϕ_1; thus, equation (10) holds.

(iii) The timbre of the environmental noise field may be expressed by all of the temporal and spatial factors, so that

$$s_T = f_T(\tau_e, \tau_1, \phi_1; LL, \mathrm{IACC}, \tau_{IACC}, W_{IACC}). \tag{23}$$

Considering the fact that the human cerebral hemispheres are specialized in such a way that temporal factors are associated with the left hemisphere and spatial factors are associated with the right hemisphere [1], one can rewrite equation (23) as

$$s_T = f_T(\tau_e, \tau_1, \phi_1)_{left} + f_T(LL, \mathrm{IACC}, \tau_{IACC}, W_{IACC})_{right}. \tag{24}$$

(iv) Since the perceived duration of the noise field is the temporal sensation, equation (13) may hold. To describe these primary sensations effectively for a certain duration of fluctuating noise, it is assumed that τ_e in equations (13) and (22–24) can be replaced by $(\tau_e)_{min}$ and its corresponding factors.

6. REMARKS

The ear sensitivity may be characterized by the physical system including the external ear and the middle ear [1, 2]. Before analyzing the sound signal, we can use an A-weighting network for the sake of convenience.

The threshold of the weak reflection is shown in Figure 13 as a function of Δt_1. The spatial direction of reflection to the listener (IACC and τ_{IACC}) and the delay time of reflection Δt_1, as included in equation (24), express this threshold [20].

It is worth noticing that the intelligibility of single syllables as a function of the delay time of a single reflection is well calculated by the four orthogonal factors extracted from the short-term ACF analyzed for the piece between consonant and vowel sounds [3, 4]. A recent investigation [21] clearly shows that timbre or dissimilarity judgment is an overall subjective response similar for the subjective preference of sound fields in concert hall. The subjective preference as well as timbre can be described by the use of the minimum value of τ_e [22, 23; see also Appendix A]. Concerning a recommended short-term-integration time, it is given in this special issue [19], such that

$$(2T)_r \approx 30\ (\tau_e)_{min}. \tag{25}$$

Effects of noise on mental tasks can be interpreted as an interference phenomenon between the task performance and the hemispheric dominance [24, 25]. Temporal factors extracted from the ACF are associated with the left cerebral hemisphere and spatial factors extracted from IACF are much concerned with the right [1]. The method of measuring environmental noise based on this theory is proposed [26]. It is supplemented [27, 28], and some of the measured examples are demonstrated in this special issue [29–32].

ACKNOWLEDGMENTS

The author wishes to thank Prof. Roberto Pompoli for his useful suggestions in improving the manuscript of this paper. This work was supported by the JST, 1999.

REFERENCES

1. Y. ANDO 1998 Architectural Acoustics, Blending Sound Sources, Sound Fields, and Listeners. New York: AIP Press/Springer.
2. Y. ANDO 1985 *Concert Hall Acoustics*. Heidelberg: Springer.
3. Y. ANDO, S. SATO and H. SAKAI 1999 in *Fundamental subjective attributes of sound fields based on the model of auditory–brain system*, J. J. Sendra editor, *Computational acoustics in architecture*. Southampton: WIT, chapter 4.
4. Y. ANDO, H. SAKAI and S. SATO 2000 *Journal of Sound and Vibration* **232**, 101–127. Formulae describing subjective attributes for sound fields based on a model of the auditory–brain system.
5. I. GDE, N. MERTHAYASA, H. HEMMI and Y. ANDO 1994 A12, 147–156. *Memoirs of Graduate School Science and Technology*, *Kobe University*. Loudness of a 1 kHz pure tone and sharply (1080 dB/Oct.) filtered noises centered on its frequency.
6. S. SATO, T. KITAMURA, H. SAKAI and Y. ANDO 2001 *Journal of Sound and Vibration* **241**, 97–103. The loudness and pitch of "complex noise" in relation to the factors extracted from the auto-correlation function.
7. E. ZWICKER, G. FLOTTORP and S. S. STEVENS 1957 *Journal of Acoustics Society of America* **29**, 548–557. Critical band width in loudness summation.
8. F. L. WIGHTMAN 1973 *Journal of Acoustics Society of America* **54**, 397–406. Pitch and stimulus fine structure.
9. T. SUMIOKA and Y. ANDO 1996 *Journal of Acoustics Society of America* A**100**, 2720. On the pitch identification of the complex tone by the autocorrelation function (ACF) model.
10. M. INOUE, Y. ANDO and T. TAGUTI 2001 *Journal of Sound and Vibration* **241**, 105–116. The frequency range applicable to pitch identification based upon the auto-correlation model, .
11. K. FUJII, S. KITA, T. MATSUSHIMA and Y. ANDO 2001 *Psychological Research*. The missing fundamental phenomenon in temporal vision, in press.
12. K. SAIFUDDIN, H. SAKAI and Y. ANDO 2001 *Journal of Sound and Vibration* **241**, 117–127. Duration sensation of the simulated environmental noise in relation to its auto-correlation function.
13. Y. ANDO and K. SAIFUDDIN 2000 *Japanese and Swedish Symposium on Medical Effects of Noise*, *Stockholm*. Duration experience during listening to bandpass noise.
14. P. DAMASKE and Y. ANDO 1972 *Acustica* **27**, 232–238. Interaural crosscorrelation for multichannel loudspeaker reproduction.
15. S. SATO, V. MELLERT and Y. ANDO 2001 *Journal of Sound and Vibration* **241**, 53–56. Cues for localization in the median plane as extracted from the auto-correlation function.
16. Y. ANDO, K. YAMAMOTO, H. NAGAMASTU and S. H. KANG 1991 *Acoustic Letters* **15**, 57–64. Auditory brainstem response (ABR) in relation to the horizontal angle of sound incidence.
17. Y. ANDO and Y. KURIHARA 1986 *Journal of Acoustics Society of America* **80**, 833–836. Nonlinear response in evaluating the subjective diffuseness of sound field.
18. M. V. KEET 1968 *Proceedings of the* 6*th International Congress on Acoustics*, *Tokyo*, *Paper* E-2-4. The influence of early lateral reflections on the spatial impression.
19. K. MOURI, K. AKIYAMA and Y. ANDO 2001 *Journal of Sound and Vibration* **241**, 87–95. Preliminary study on recommended time duration of source signals to be analyzed, in relation to its effective duration of auto-correlation function, to be published.
20. Y. ANDO and H. ALRUTZ 1982 *Journal of Acoustics Society of America* **71**, 616–618. Perception of coloration in sound fields in relation to the autocorrelation function.
21. T. HOTEHAMA, S. SATO and Y. ANDO 2000 Dissimilarity judgements switching loudspeakers located on the stage in a concert hall (unpublished).
22. Y. ANDO, T. OKANO and Y. TAKEZOE 1989 *Journal of Acoustics Society of America* **86**, 644–649. The running autocorrelation function of different music signals relating to preferred temporal parameters of sound fields.
23. K. MOURI, K. AKIYAMA and Y. ANDO *Journal of Sound and Vibration* **232**, 139–147. Relationship between subjective preference and the alpha-brain wave in relation to the initial time delay gap with vocal music.
24. Y. ANDO 2001 *Journal of Sound and Vibration* **241**, 129–140. Differential effects of noise and music signals on the behavior of children.
25. K. MOURI, K. AKIYAMA and Y. ANDO 2001 *Journal of Sound and Vibration* **241**, 141–145. Effects of telephone ring on two mental tasks relative to an office, to be published.
26. M. SAKURAI, H. SAKAI and Y. ANDO 2001 *Journal of Sound and Vibration* **241**, 19–27. A computational software for noise measurement and towards its identification.

27. M. KAZAMA and M. TOHYAMA 2001 *Journal of Sound and Vibration* **241**, 41–52. Estimation of speech components by ACF analysis in noisy environment.
28. A. HEISS 2001 *Journal of Sound and Vibration* **241**, 29–39. A supplement of sound measurement technique by an online quality test procedure for the indices Leq and Lx%, to be published.
29. H. SAKAI, S. SATO, N. PRODI and R. POMPOLI 2001 *Journal of Sound and Vibration* **241**, 57–68. Measurement of regional environmental noise by use of a PC-based system—An application to the noise near airport 'G. Marconi' in Bologna.
30. K. FUJII, Y. SOETA and Y. ANDO 2001 *Journal of Sound and Vibration* **241**, 69–78. Acoustical properties of aircraft noise measured by temporal and spatial factors.
31. Y. SOETA, K. OHTORI, K. FUJII and Y. ANDO 2001 *Journal of Sound and Vibration* **241**, 79–86. Measurement of sound transmission by hall doors.
32. J. Y. JEON 2001 *Journal of Sound and Vibration* **241**, 147–155. Subjective evaluation of floor impact noise based on the model of ACF/IACF.
33. M. R. SCHROEDER, D. GOTTLOB and K. F. SIEBRASSE 1974 *Journal of Acoustics Society of America* **56**, 1195–1201. Comparative study of European concert halls: correlation of subjective preference with geometric and acoustic parameters.

APPENDIX A. THE MOST PREFERRED CONDITIONS FOR THE SOUND FIELD IN A CONCERT HALL

For the reader's information, the most preferred conditions for the sound field in a concert hall are briefly described here by both temporal and spatial factors [2].

(1) The most preferred initial time delay gap between the direct sound and the first reflection is expressed by

$$[\Delta t_1]_p \approx [1 - \log_{10} A](\tau_e)_{min}, \tag{A.1}$$

where $(\tau_e)_{min}$ is the minimum value of the effective duration of the running ACF of the source signal, and A is the total amplitude of reflections given by

$$A = \left\{ \sum_{n=1}^{\infty} A_n^2 \right\}^{1/2}, \tag{A.2}$$

A_n being the pressure amplitude of the nth reflection, $n = 1, 2, \ldots$.

(2) The most preferred subsequent reverberation time is expressed by

$$[T_{sub}]_p \approx 23\ (\tau_e)_{min}. \tag{A.3}$$

(3) One of the typical spatial factors is the IACC. The consensus preference is obtained at a small value of the IACC, so that signals arriving at both ears should be dissimilar. But, the peak value of the IACF must be maintained at the origin of the delay time, i.e.,

$$\tau_{IACC} = 0, \tag{A.4}$$

so that the sound field is well balanced.

(4) Concerning the listening level, LL, the sound power level is dependent upon the musical notes and performance. It is recommended that each individual listener may easily listen to music by use of the seat-selection system [1].

It is worth noticing that large individual differences in the subjective preferences were observed for the temporal factors expressed by equations (A.1) and (A.2), and LL [1], but not for the IACC. The IACC is called, therefore, the consensus factor [2, 33].