Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Sound and Vibration 277 (2004) 573–582

JOURNAL OF SOUND AND VIBRATION

www.elsevier.com/locate/yjsvi

Computer software for identification of noise source and automatic noise measurement

Kenji Fujii[a,*], Masatsugu Sakurai[a], Yoichi Ando[b]

[a] *Yoshimasa Electronic Inc., Daiichi-Nishiwaki Building, 1-58-10 Yoyogi, Shibuya, Tokyo 151-0053, Japan*
[b] *1-10-27 Yamanokami, Kumamoto 862-0915, Japan*

Accepted 25 March 2004
Available online 28 July 2004

Abstract

A new computational system for the environmental noise measurement and analysis has been developed. The system consists of binaural microphones, a laptop PC, and analysing software. A target noise is recorded automatically depending on the specified background noise level, and the acoustical parameters are calculated simultaneously. These functions allow for precise field measurements. The system is equipped with a template-matching algorithm for the identification of noise source. This function was implemented to avoid the effect of an interrupting sound such as voice and wind blowing during a measurement. Noise analyses in this system are based on the model of human auditory system. In addition to the time-series data of sound level, the important acoustical parameters of noise source are extracted from the running autocorrelation function (ACF) and the inter-aural cross-correlation function (IACF). It has been found that those parameters are strongly related to the auditory primary sensations and spatial sensations. Evaluation of the environmental noise based on these functions is another feature of this system. This paper describes the effectiveness of the ACF and the IACF analysis for analysing acoustical properties of noise and for evaluating the subjective response to noise.

1. Introduction

In this paper, we propose a noise measurement system for identifying a noise source and analysing a multidimensional sound quality of noise [1]. The first goal of the study is to create an

*Corresponding author. Tel.: +81-3-3370-5160; fax: +81-3-3370-6548.
E-mail address: fujii@ymec.com (K. Fujii).

0022-460X/$ - see front matter © 2004 Published by Elsevier Ltd.
doi:10.1016/j.jsv.2004.03.019

automatic noise measurement system for a long period of time. In practice, the automatic measurement is difficult, because there are many noise sources in the measurement field that may intervene. To perform an automatic noise measurement, identification of noise source is indispensable. The second goal of the proposed system is to overcome a limitation of the current procedure for evaluating the effect of noise. It is recognized that the sound level met can measure only one dimension of noise (i.e. loudness), but sometimes loudness may not evaluate subjective annoyance of noise. We need to analyse sound qualities other than loudness for evaluating the psychological effect of noise.

Noise analyses in the proposed system are based on a model of human auditory system [2]. It is recognized that the frequency analysis is performed in the periphery of the auditory system. This is a kind of bandpass filter bank in the basilar membrane. However, several psychophysical properties of sound cannot be fully explained by the frequency analysis. Experimental results on pitch and loudness suggest that some form of time domain analysis is performed in the more central area of the auditory system [3,4]. Recent neurophysiological studies on human hearing also suggest that there is a kind of temporal processor in the auditory pathway [5]. Time domain analysis of the firing rate from the auditory nerve of a cat revealed a pattern of the autocorrelation function rather than the frequency domain analysis [6], and the inter-spike interval distributions of the auditory nerve of cat resemble the short-time autocorrelation function (ACF) of the input sound [7]. A proposed model of the ACF processor is illustrated in Fig. 1. Also, the inter-aural cross-correlator is the most powerful model of the binaural hearing mechanism [2,8]. Neural activity regarding this mechanism has been found in the auditory brainstem response [9].

Considering the findings described above, a noise measurement system is proposed based on the temporal correlation mechanisms. The important acoustical parameters of noise source are extracted from the running ACF and the inter-aural cross-correlation function (IACF), in addition to the time-series data of the sound level. It has been found that those parameters are strongly related to the auditory primary sensations (loudness, pitch, and timbre) and spatial

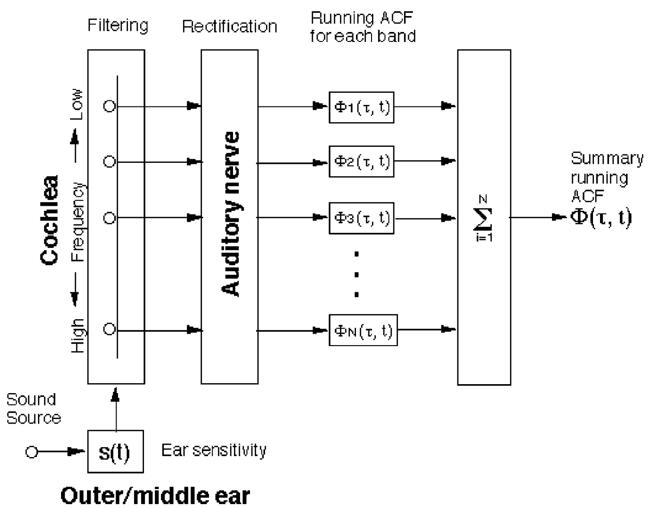

Fig. 1. Neural processing model of the running ACF.

sensations (localization, apparent source width, and subjective diffuseness) [2]. Evaluation of noise based on these hearing sensations is the goal of our study.

In the following, it is described how the ACF and the IACF analysis are effective for analysing acoustical properties of noise and for evaluating the subjective response to noise. Then the outline of the proposed measurement system is described.

2. Noise analysis by ACF and IACF

Previously, we have measured many kinds of noise and analysed them by the ACF [10–14]. These studies have shown that the acoustical properties of noise are well represented by the ACF analysis. The long-time ACF is defined by

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

where $p'(t) = p(t)*s(t)$, $s(t)$ is the ear sensitivity practically represented by the impulse response of the A-weighting filter. The value of τ represents the time delay and the value of $2T$ is the integration time. The ACF is also calculated as the Fourier transform of the power spectrum. In the ACF analysis, we can extract four significant parameters. First, is the sound energy, $\Phi(0)$, calculated as the ACF value at $\tau = 0$. Second, is the effective duration of the ACF, τ_e, which is defined as the 10-percentile time delay of the normalized ACF. It is the decay rate of the ACF, representing the repetitive nature and reverberation component contained in the signal. The third and fourth parameters are the time delay and the amplitude of the maximum peak in the normalized ACF, τ_1 and ϕ_1. These two parameters represent perceived pitch and its strength as described below.

For analysing the spatial properties of sound, three parameters are extracted from the IACF. The IACF is given by

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

where p'_1 and p'_r are A-weighted sound received by left and right ear. To measure the IACF, binaural recording is performed by using dummy head or human head microphones. The IACF is usually normalized by $\Phi(0)$. The magnitude of the IACF is defined by

$$\mathrm{IACC} = \left|\phi_{lr}(\tau)\right|_{\max}, \quad |\tau| \le 1\ \mathrm{ms}. \tag{3}$$

IACC represents the degree of similarity of sound arriving at each ear that is a significant factor to determine the degree of subjective diffuseness in a sound field [2]. When IACC decreases, the subjective diffuseness increases. The inter-aural time delay (ITD) between -1 and $+1$ ms is defined as τ_{IACC}. The value of τ_{IACC} represents the horizontal sound direction.

To illustrate how the ACF analysis is effective for analysing the acoustical properties of noise, the power spectra and the ACFs calculated for the sound from the motorcycle and the passenger car are shown in Fig. 2. The former gives a clear pitch sensation and the latter has a weak pitch. These sounds are referred to as "tonal noise" and "atonal noise". Generally, the spectrum of complex sounds consists of the tonal component (discrete peak) and the noise component

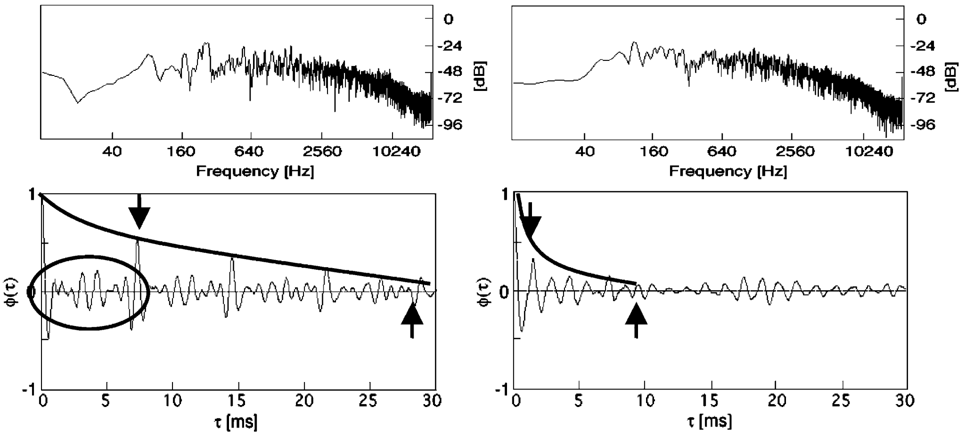

Fig. 2. Power spectra and the ACF of motorcycle (left) and passenger car (right) sound.

(continuous part). In the tone correction for the aircraft noise, tone/noise ratio or difference is calculated. However, it is sometimes difficult to identify which peak in the spectrum is a fundamental frequency that is perceived as a pitch. When the same sound is analysed by the ACF, its harmonic structure is clearly extracted. Strong periodical peaks in the ACF show that a periodicity corresponding to the pitch is present in the sound. Reciprocal of the time delay of the peak corresponds to the pitch frequency. The amplitude of the peak corresponds to the perceived pitch strength. Minor peaks within a period in the ACF give useful information about the higher frequency components, or timbre of sound [7,15]. For the atonal noise, there is no particular peak in the spectrum. This means that the sound has no particular periodicity perceived as a pitch. In this case, the ACF decreases to zero without exhibiting strong periodical peaks. Thus, the decay rate of the ACF, τ_e, is a good measure of the periodical structure of sound.

For analysing the temporally fluctuating noise, the acoustical parameters should be calculated for short-time intervals. For this purpose, the short-time running ACF analysis is useful. To capture the fluctuation of sound quality, selection of the integration time (temporal window, $2T$ in Eq. (1)) and calculation interval becomes important. In the case of music, $2T$ was set between 2 and 5 s [16]. This is based on the study of the ‘‘psychological present’’, which is the time length in which human perceives successive events as one thing [17]. But in calculating the ACF for a single syllable of Japanese speech, a much shorter $2T$ (about 30 ms) was needed to capture the very fast change in the speech signal [18].

To capture the fluctuation of aircraft noise, $2T$ for calculating the ACF has been examined. Fig. 3 shows the time course of SPL for two aircraft noises with different τ_e values (Fig. 3, top), integrated for three different $2T$s. The calculation interval was set to the half of $2T$ (for example, when $2T$ was 1.0 s, calculation interval was 0.5 s). It is clear that the $2T$ of 1.0 s is too long to capture the fluctuation of SPL. It seems that a fluctuation could be caught by $2T$ of 0.25 or 0.5 s, but a finer variation could be measured only in the case of 0.25 s. Informal listening tests showed that such a finer variation could not be heard for a sound with long τ_e value. For the sound with short τ_e, on the other hand, 0.25 s matched the perceived sound fluctuation. This implies that the proper length of temporal window should be chosen depending on the periodical structure of the

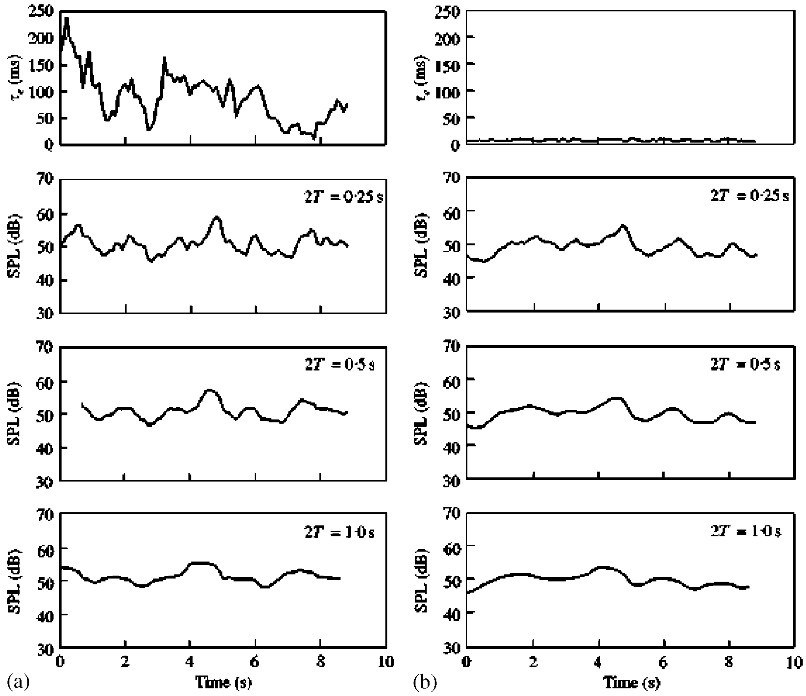

Fig. 3. Examples of measured SPL for two types of noise with (a) $(\tau_e)_{\min} = 20$ ms and (b) $(\tau_e)_{\min} = 10$ ms with three integration times ($2T$), from the second row, 0.25, 0.5, and 1.0 s.

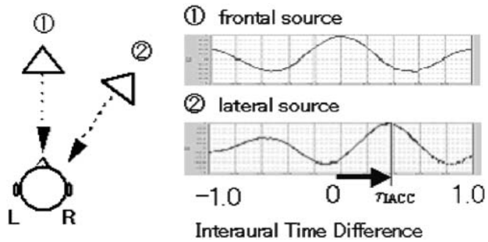

Fig. 4. Relationship between the sound direction and IACF parameter, τ_{IACC}.

sound. Mouri et al. [19] have reported that the integration time should be set as $2T \approx (\tau_e)_{\min}$, but this is a tentative value. More experimental data is needed to find the optimal temporal window for analysing time-varying noise. As reported by Fujii et al. [12], the ACF parameters can be possible measures to evaluate subjective annoyance of noise.

In the field measurement, information about the direction of noise is very useful. The maximum peak of the IACF, IACC, represents direction of a sound source because it corresponds to the inter-aural time difference. As a sound source moves from left to right, the value of τ_{IACC} varies from minus to plus value, as illustrated in Fig. 4. By the τ_{IACC} measurement, moving sound sources such as a car or an airplane can be tracked and the hidden sound source in a factory can be detected. Also, it has been found recently that the temporal fluctuation in spatial sensations such as direction and diffuseness strongly affects subjective annoyance [13]. This implies that binaural measurement is necessary to evaluate the effect of such spatial sensations of subjective annoyance.

3. Measurement system

Based on the basic theory of Ando [2], the prototype of a noise measurement system has been developed during 1999 and 2000 as a first step of the project [1]. Here, our outline of the system is described.

3.1. Automatic measurement

Fig. 5 is a block diagram of the measurement system. The system consists of binaural microphones, a laptop computer, and software calculating the ACF and the IACF parameters from a real time noise data. The system can measure noise event automatically by using the peak detection algorithm. As shown in Fig. 6, the noise level is continuously monitored, and a target noise event is extracted when the noise level exceeds the trigger level, L_{trig}. The appropriate L_{trig} value varies according to the kind of target noise and the distance between the noise source and the receiver. It must therefore be determined by a preliminary measurement. The noise data with duration of t_{s} centred on its peak level is recorded on the hard disk as a single session. The duration t_{s} should be set so as to include the peak level after exceeding L_{trig}. For aircraft noise, it is set to about 10 s. This value is different between a steady-state noise with longer duration and an intermittent noise. Note that the current system does not work well when there are interfering noises.

3.2. Running ACF and IACF calculation

For each session of noise, the running ACF and IACF are calculated with the integration time ($2T$) and the calculation interval (running step: t_{step}). Appropriate values of $2T$ and t_{step} are determined before the measurement. As you can see in Fig. 6, the ACF and IACF are calculated in every step ($n = 1, 2, \ldots, M$) for the data length of $2T$ which shifts in every t_{step}, as $\{(0,\ 2T),\ (t_{\mathrm{step}},\ t_{\mathrm{step}} + 2T), (2t_{\mathrm{step}},\ 2t_{\mathrm{step}} + 2T), \ldots, ((M-1)\,t_{\mathrm{step}},\ (M-1)\,t_{\mathrm{step}} + 2T)\}$. Acoustical parameters described before are extracted from each step of the ACF and IACF and used for the identification process.

3.3. Identification of noise source

In the proposed system, noise source is identified based on the similarity of the acoustical parameters by using the template-matching algorithm. The basic concept of the identification algorithm in this system is illustrated in Fig. 7. Three dimensions in the figure represent the acoustical parameters considered. A fundamental assumption here is that the noises from the

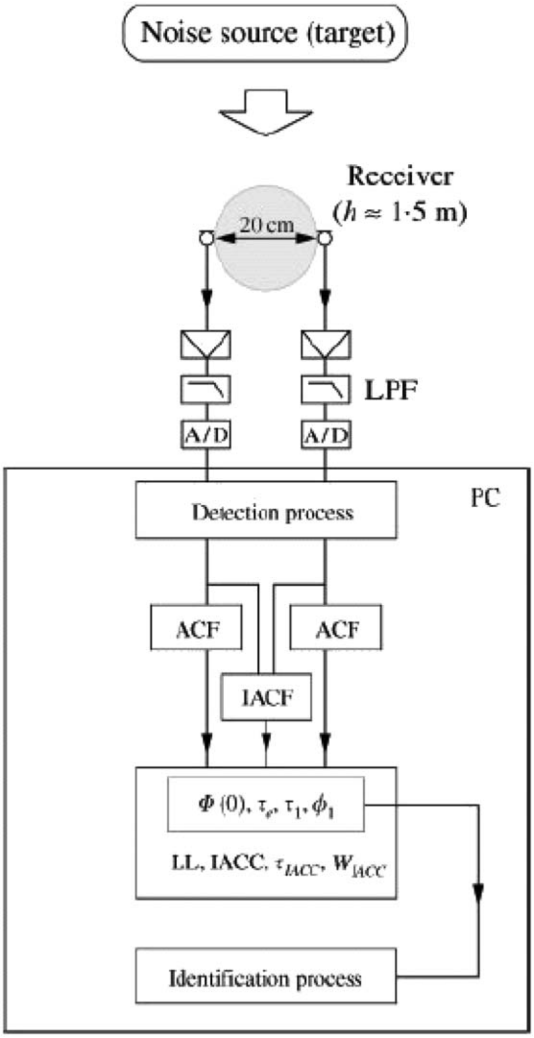

Fig. 5. A flow chart of the measurement system. The ACF and the IACF parameters are extracted after the process of automatic detection of noise event. Noise source is identified by using the ACF parameters.

same source have similar acoustical qualities and therefore they are mapped into the same cluster in the feature space. Now, suppose that there are two noise sources, and they are represented as the parameter set A and B. When new data X is measured, the distance of parameters $|A - X|$ and $|B - X|$ are calculated. Distance represents the similarity of parameters. Thus, if the distance $|A - X|$ is smaller than $|B - X|$, the input data X is categorized in A.

The acoustical parameters used in the identification algorithm are derived from the running ACF. As described above, timbre information of noise is extracted from the ACF. Distance $D(x)$ ($x : \Phi(0)$, τ_e, τ_1, and φ_1) between the unknown target data (indicated by symbol a in Eqs. (4–7) and the template (already calculated data of noise sources, indicated by symbol b) is calculated at

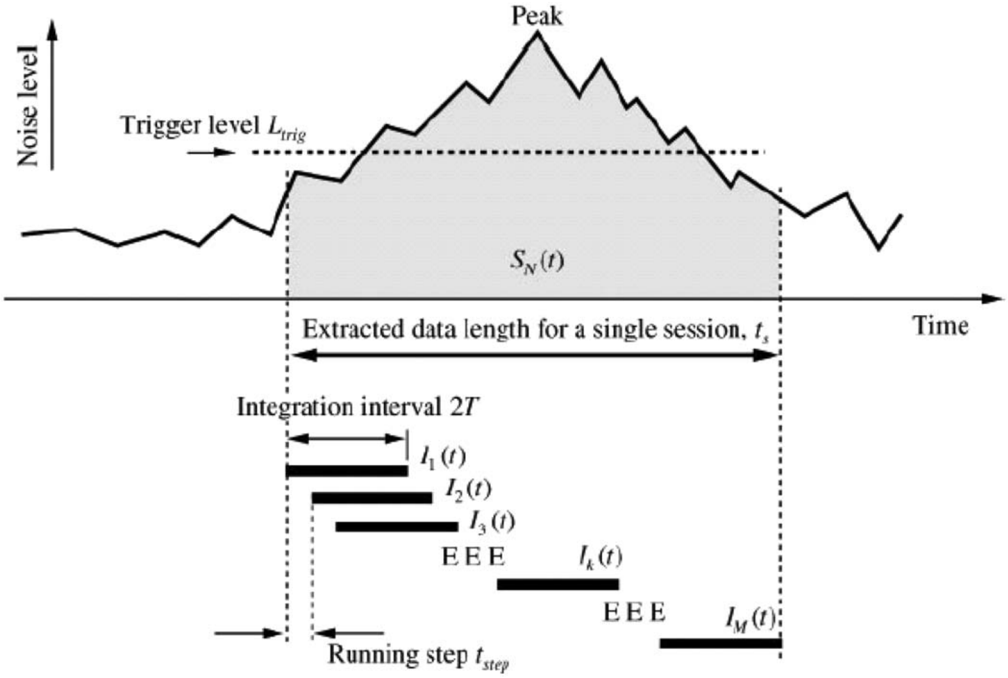

Fig. 6. Process of the extraction of noise event and the running ACF analysis.

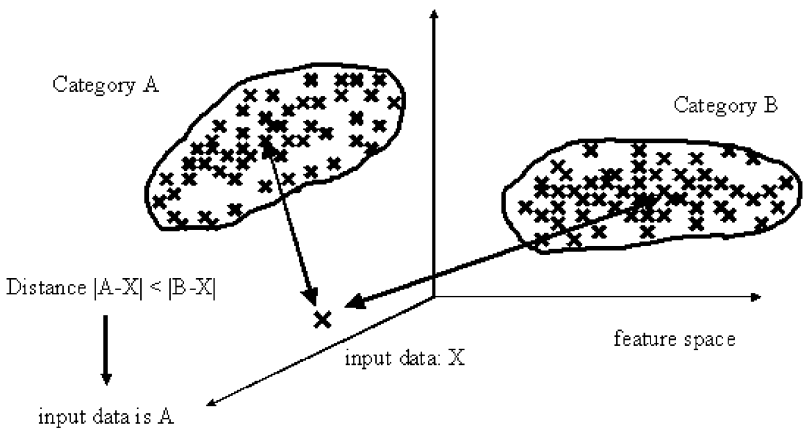

Fig. 7. Basic concept of the identification algorithm.

SPL by the following equations:

$$D(\Phi(0)) = \left|\log(\Phi(0)^a - \log\left(\Phi(0)^b\right)\right|, \tag{4}$$

$$D(\tau_e) = \left|\log(\tau_e)^a_{\min} - \log(\tau_e)^b_{\min}\right|, \tag{5}$$

$$D(\tau_1) = \left|\log(\tau_1)^a - \log(\tau_1)^b\right|, \tag{6}$$

$$D(\phi_1) = \left|\log(\phi)^a - \log(\phi_a)^b\right|. \tag{7}$$

$D(x)$ is calculated at the maximum SPL of measured noise, because the noise is considered as steady state at this point in many cases. The total distance D is calculated as the sum of the right-hand terms of Eqs. 4–7, so

$$D = W^{\Phi(0)} D\Phi(0)) + W^{\tau_e} D(\tau_e) + W^{\tau_1} D(\tau_1) + W^{\phi_1} D(\phi_1), \tag{8}$$

where $W^{(x)}$ shows the weighting coefficient of each parameter. The template with the smallest D is defined as the noise source of the target data. Weighting coefficients $W^{(x)}$ in Eq. (8) are obtained by using statistical values $S_1^{(x)}$ and $S_2^{(x)}$, as $(S_2/S_1)^{1/2}$. Here, $S_1^{(x)}$ is the arithmetic mean of the standard deviations (SD) for all categories (i.e. SD within a certain category), and $S_2^{(x)}$ is the SD of the arithmetic means in each category (i.e. SD between categories), and category means a set of data for the same noise source. It means that a parameter (x) becomes important when the SD between categories becomes large and the SD among a category becomes small.

The performance of the identification algorithm was tested for several kinds of sounds, such as car engine sound, aircraft noise, motorcycle noise, and human voice. Good identification performance was obtained in the measurement of the idling car engine sound. It is considered that the identification worked well because the acoustical quality of idling sound is almost constant during measurement and the ACF parameters differed well for different sounds. The proposed algorithm is thus applicable for the steady-state noise. Noise with non-stationary sound qualities was sometimes misidentified. It is because the algorithm calculates the distance by using sound parameters only at the maximum SPL point. This problem would be solved by using the whole time pattern of the ACF parameters for identification. A more important problem in the present algorithm is that it cannot deal with the mixed noise. Separation of sound source is indispensable in order to apply the measurement system in the real environment. It has been found that the binaural localization cues are very effective in separating a target sound from an interfering sound [20]. We plan to implement the algorithm for sound separation in future research.

4. Summary

This paper described a noise measurement system for identifying a noise source and analysing a multidimensional sound quality of noise. It was shown how the ACF and IACF analysis are effective for analysing acoustical properties of noise. The identification process was found to be effective for the steady-state noise. Further research is planned to deal with the fluctuating noise and mixed noise for applying the system in the real environment.

Reference

[1] M. Sakurai, H. Sakai, Y. Ando, Computational software for noise measurement and toward its identification, *Journal of Sound and Vibration* 241 (2001) 19–27.

[2] Y. Ando, A theory of primary sensations and spatial sensations measuring environmental noise, *Journal of Sound and Vibration* 258 (2001) 405–417.
[3] M. Inoue, Y. Ando, T. Taguti, The frequency range applicable to pitch identification based upon the auto-correlation function model, *Journal of Sound and Vibration* 241 (2001) 105–116.
[4] Y. Ando, S. Sato, H. Sakai, Fundamental subjective attributes of sound fields based on the model of auditory–brain system, in: J.J. Sandra (Ed.), *Computational Acoustics in Architecture*, WIT Press, Southampton, 1999, pp. 63–99 (Chapter 4).
[5] G. Langner, Periodicity coding in the auditory system, *Hearing Research* 60 (1992) 115–142.
[6] H.E. Secker-Walker, C.L. Searle, Time domain analysis of auditory-nerve-fiber firing rates, *Journal of the Acoustical Society of America* 88 (1990) 1427–1436.
[7] P.A. Cariani, B. Delgutte, Neural correlates of the pitch of complex tones. I. Pitch and pitch salience, *Journal of Neurophysiology* 76 (1996) 1698–1716.
[8] J. Blauert, *Spatial hearing*, MIT Press, Cambridge, MA, 1996.
[9] Y. Ando, K. Yamamoto, H. Nagamatsu, S.H. Kang, Auditory brainstem response (ABR) in relation to the horizontal angle of sound incidence, *Acoustic Letters* 15 (1991) 57–64.
[10] K. Fujii, Y. Soeta, Y. Ando, Acoustical properties of aircraft noise measured by temporal and spatial factors, *Journal of Sound and Vibration* 241 (2001) 69–78.
[11] H. Sakai, S. Sato, N. Prodi, R. Pompoli, Measurement of regional environmental noise by use of a PC-based system. An application to the noise near airport "G. Marconi" in Bologna, *Journal of Sound and Vibration* 241 (2001) 57–68.
[12] K. Fujii, J. Atagi, Y. Ando, Temporal and spatial factors of traffic noise and its annoyance, *Journal of Temporal Design* 2 (2002) 33–41.
[13] T. Kitamura, R. Shimokura, S. Sato, Y. Ando, Measurement of temporal and spatial factors of a flushing toilet noise in a downstairs bedroom, *Journal of Temporal Design* 2 (2002) 13–19.
[14] H. Sakai, T. Hotehama, Y. Ando, Diagnostic system based on the human auditory–brain model for measuring environmental noise–an application to railway noise, *Journal of Sound and Vibration* 250 (2002) 9–21.
[15] R. Meddis, M.J. Hewitt, Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: pitch identification, *Journal of the Acoustical Society of America* 89 (1991) 2866–2882.
[16] Y. Ando, *Architectural Acoustics, Blending Sound Sources, Sound Fields, and Listeners*, AIP Press/Springer, New York, 1998.
[17] P. Fraisse, Rhythm and tempo, in: D. Deutsch (Ed.), *Psychology of Music*, Academic Press, Orlando, FL, 1982.
[18] T. Shoda, Y. Ando, Calculation of speech intelligibility using four orthogonal factors extracted from the autocorrelation function of sound source and sound field signals, *Journal of the Acoustical Society of America* 103 (1998) 2999.
[19] K. Mouri, K. Akiyama, Y. Ando, Preliminary study on recommended time duration of source signals to be analyzed, in relation to its effective duration of autocorrelation function, *Journal of Sound and Vibration* 241 (2001) 87–95.
[20] N. Roman, D.L. Wang, G.J. Brown, Localization-based sound segregation, in: *Proceedings of IJCNN-02*, 1998, pp. 2299–2303.