ICIC Express Letters
Part B: Applications
Volume **10**, Number **4**, April **2019**

ICIC International ©2019 ISSN 2185-2766
pp. 335–342

OBJECTIVE EVALUATION OF SOUND QUALITY FOR AUDIO SYSTEM IN CAR

Arimasa Kanda[1], Shunsuke Ishimitsu[1], Koji Wakamatsu[2]
Masanori Nakashima[2] and Hiroshi Yamanaka[2]

[1]Graduate School of Information Science
Hiroshima City University
3-4-1, Ozuka-Higashi, Asaminami-ku, Hiroshima 731-3194, Japan
a-kanda@hfce.info.hiroshima-cu.ac.jp; ishimitu@hiroshima-cu.ac.jp

[2]Mazda Motor Corporation
3-1, Shinchi, Fuchu-cho, Aki-gun, Hiroshima 730-8670, Japan

Received September 2018; accepted December 2018

Abstract. *The perception of sound is very multidimensional in nature. Among these dimensions, we focus on sound clarity whose objective parameter has not yet been specified for spatial impression. To evaluate sound clarity, an analysis of the physical properties is needed. In this study, the sound clarity of a layout of loudspeakers in a car was investigated and evaluated using transfer functions and time-frequency analyses. Four pairs of loudspeakers were mounted around the driver's seat. Furthermore, an auditory experiment was conducted to examine the relationship between physical factors and auditory impression. We believe that sound clarity correlates with the magnitude property of the transfer function, attack property, and sound separation. Additionally, prediction models were calculated using multiple regression analysis. It is believed that the proposed method enables an objective evaluation of sound clarity in car audio systems.*
Keywords: Clarity, Transfer function, Wavelet analysis, Multiple regression analysis

1. **Introduction.** Sound quality has a very multidimensional nature, especially for sound perception. The multidimensional nature of auditory impression consists of clarity, localization, and presence. In this study, we focused on sound clarity. Sound clarity is the precision with which the details of sound can be heard [1]. Sound clarity can be analyzed through subjective and objective evaluations. Subjective evaluation has a demerit because it depends on age and on individual differences in audio perception. Moreover, a lot of participants and time are required for a subjective evaluation experiment. Conversely, in objective evaluation, various parameters have been proposed for use in evaluating room acoustics [2,3]. However, an objective parameter for spatial impression has not been fully studied yet. The objective parameter of sound clarity makes it possible to evaluate sound quality objectively irrespective of individual differences. To evaluate sound clarity, an analysis of the physical properties of sound is needed.

In this study, we investigated loudspeaker layout in a car and evaluated its sound clarity using transfer function and time-frequency analysis. 1) The transfer function was further converted to obtain the maximum value based on the listening resolution of each critical bandwidth. In addition, 2) envelope analysis was performed on the time waveforms using results from the time-frequency analysis. Furthermore, 3) an auditory experiment was conducted to investigate the relationship between physical factors and auditory impression. Finally, 4) a prediction model for sound clarity using multiple regression analysis was calculated, and we discussed the possibility of quantification of sound clarity.

DOI: 10.24507/icicelb.10.04.335

2. Evaluation Using Transfer Function in Car and Anechoic Room.

2.1. Experiment methods. First, a binaural recording was performed in a car. Four pairs of identical loudspeakers (Genuine product) were mounted around the driver's seat. Figure 1 shows the layout of the loudspeakers in the car. The loudspeakers were numbered from Nos. 1 to 4, where No. 1 was the loudspeaker on the dashboard, No. 2 was fixed to the upper part of the front door, No. 3 was placed below No. 2, and No. 4 represented the original layout position for the car (below No. 3). In addition, the binaural recording was conducted in an anechoic room. The loudspeaker was positioned exactly in front of a dummy torso. White noise was fed to a microphone (Brüel & Kjær/TYPE 4101-A) to calculate the transfer function, and it was recorded at a sampling frequency of 65,536 Hz with 24-bit quantization. The equivalent sound level for the white noise was set at 75 dB.

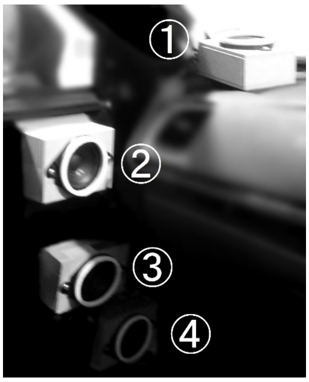
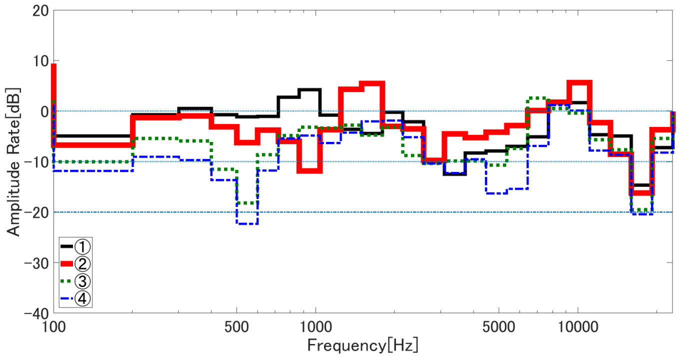


FIGURE 1. Loudspeaker layout in car

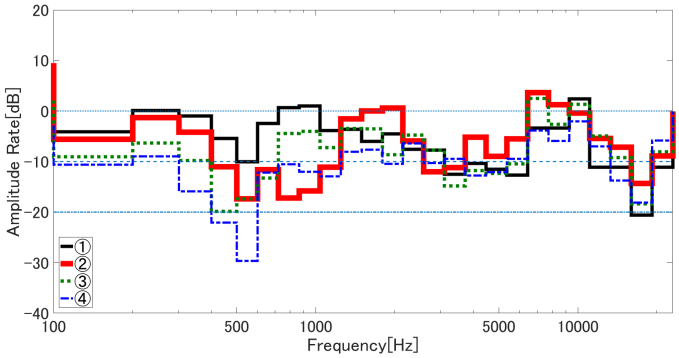
2.2. Comparison of transfer function. The transfer function between each loudspeaker and the ears was calculated using the cross-spectral method [4]. To express the difference between the transfer functions calculated in the car and the anechoic room clearly, the properties of each loudspeaker position were normalized using the properties of the anechoic room. Furthermore, the transfer functions were converted to obtain the maximum value based on the listening resolution for each critical bandwidth [5]. Figure 2 shows the normalized and converted transfer functions in each of the critical bandwidths. The horizontal axis represents the frequency while the vertical axis represents the amplitude rate, making it easy to find the difference. The result of the analysis showed that the amplitude rate was flat between 0.5 and 2 kHz, especially for loudspeaker No. 1. Figure 2 indicates that the amplitude rate for No. 1 was 11 dB larger in the car than in the anechoic room. The amplitude rates for loudspeaker Nos. 2, 3, and 4 were respectively 18 dB, 20 dB, and 30 dB lower in the car than in the anechoic room.

3. Auditory Experiments.

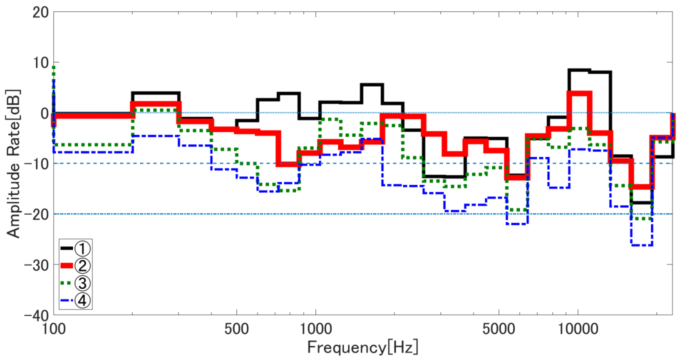
3.1. Auditory experiment with nonprofessional participants. For the auditory experiment, three music stimuli were selected that are widely used in car auditory evaluations by Japanese manufacturers. Table 1 shows the applied music stimuli. Stimulus sounds were created by convolving the transfer functions obtained from the position of each loudspeaker from No. 1 to No. 4. Figure 3 shows the paths between the loudspeakers and the ears. Ten participants (9 males and 1 female, aged 22-24) participated in the experiment that was conducted by applying Scheffe's method [6]. The participants evaluated the sound clarity using five rating steps (from -2 to $+2$). The auditory experiment was performed with a headphone (SENNHEISER/HD595) and a headphone amplifier (SONY/STR-DN850) in an anechoic room.



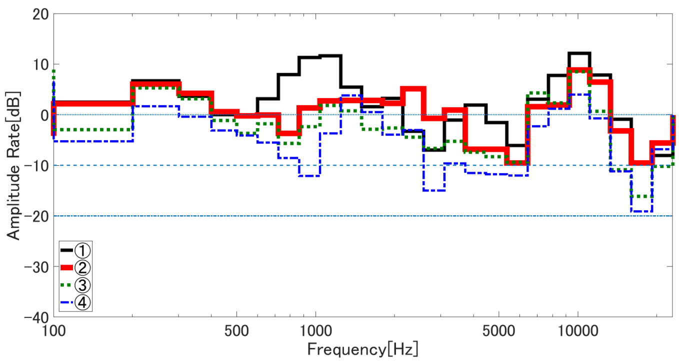
(a) From left loudspeaker to left ear



(b) From left loudspeaker to right ear



(c) From right loudspeaker to left ear



(d) From right loudspeaker to right ear

FIGURE 2. Normalized and converted transfer functions from each loudspeaker to each ear

TABLE 1. Music stimuli

Stimulus	**Genre (Included sound)**
Music A	Rock (Guitar and Conga sound)
Music B	Pops (Drum and Synthesizer sound)
Music C	Ballade (Female vocal)

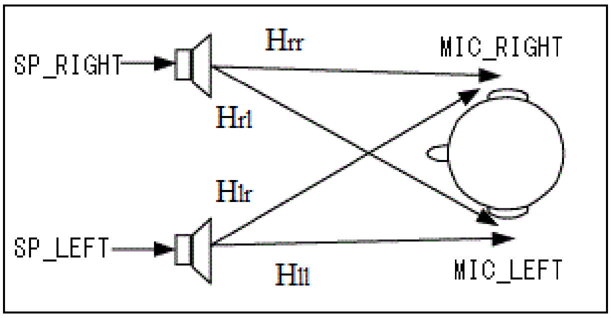
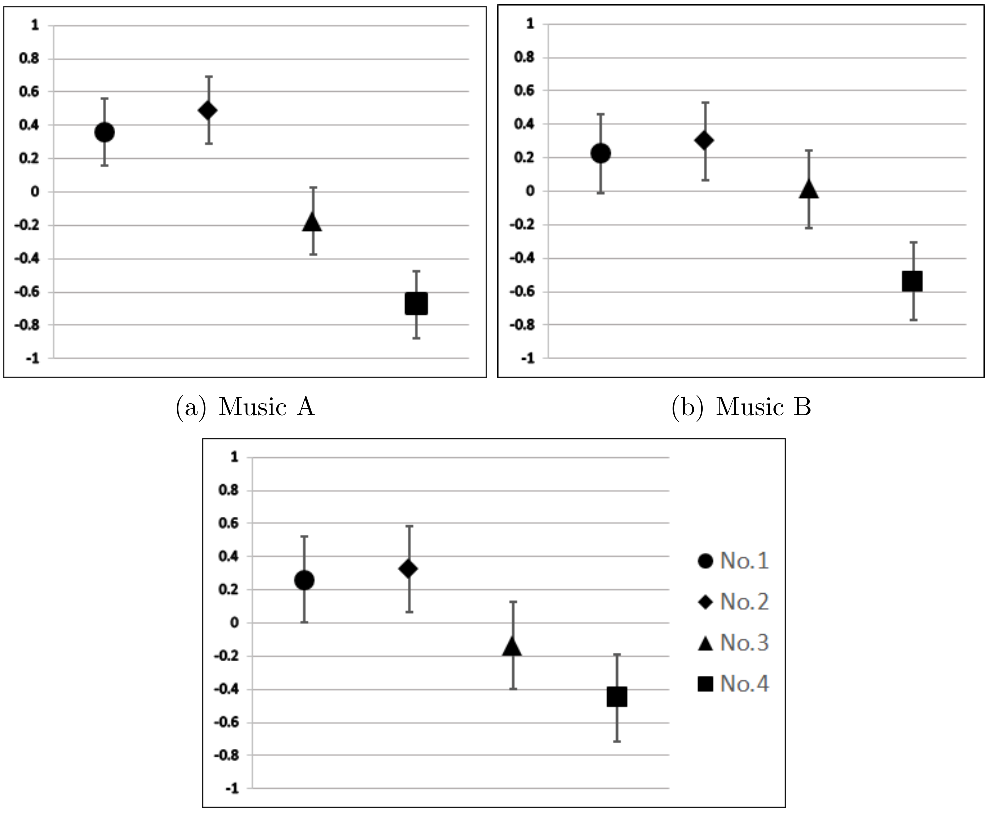


FIGURE 3. Experimental setting



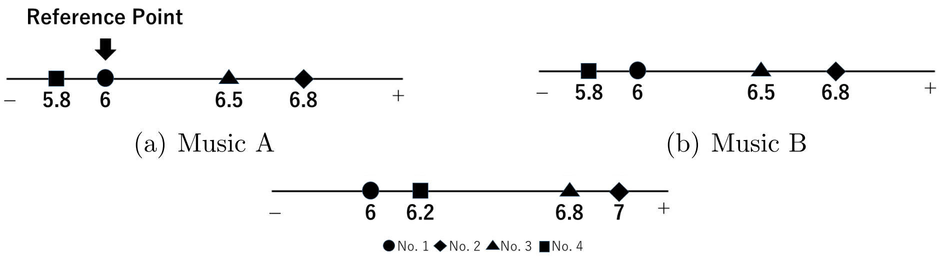
(a) Music A (b) Music B

(c) Music C

FIGURE 4. Mean scores for four loudspeaker layouts for each stimulus

3.2. **Results.** Figure 4 shows the results of the auditory experiment. The vertical axis represents the mean scores for four loudspeaker layouts for each stimulus. The error bar indicates the confidence interval at a significance level of 1%. According to the experimental results, loudspeaker No. 2 produced the best sound clarity. The ranking of the positions of the loudspeakers, starting with the highest quality, are as follows: No. 2, No. 1, No. 3, and No. 4.

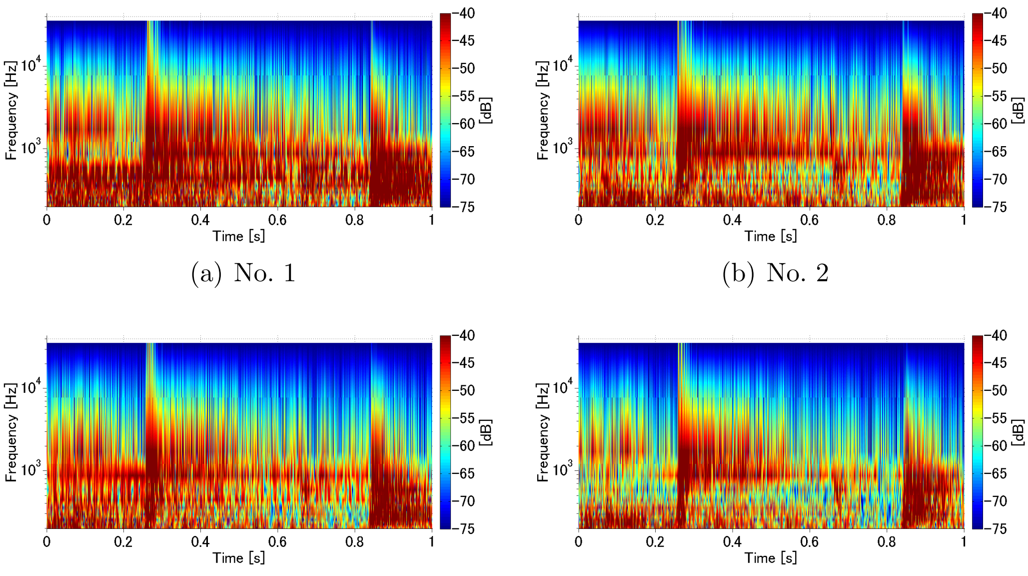
3.3. **Evaluation by professional audio estimator.** In addition, a professional audio estimator qualitatively evaluated the stimuli. Figure 5 shows the result of the evaluation. Professional audio estimator defined No. 1 as 6 points that was used as a reference. It is easy to find the auditory difference if the evaluated difference was greater than one point. Compared to the sound of loudspeakers Nos. 1 and 4, other loudspeaker positions indicated higher sound qualities. According to the results of both nonprofessional participants and the professional audio estimator, No. 2 provided the highest sound quality. The amplitude of the transfer function was lower in No. 2 than in No. 1, as seen from Section 2. However, No. 2 showed high sound clarity. Therefore, we believe that sound clarity is not only correlated with the magnitude properties of the transfer function.



(a) Music A (b) Music B

(c) Music C

Figure 5. Evaluation by professional audio estimator for each position



(a) No. 1 (b) No. 2

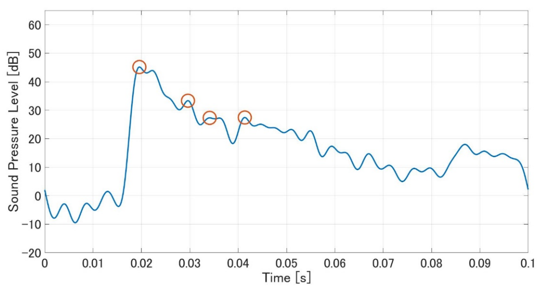
(c) No. 3 (d) No. 4

Figure 6. Results of time-frequency analysis for each loudspeaker position

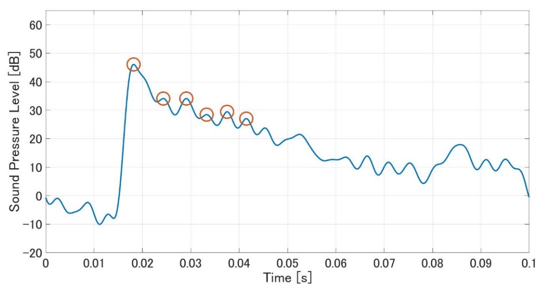
4. Analysis of Attack Property.

4.1. **Time-frequency analysis.** We performed time-frequency analysis on the stimuli for a sound sample named "Music A" [7]. The stimuli were conga sound created with each position from No. 1 to No. 4 that are the listening points defined by the professional audio estimator. Figure 6 shows the results of the time-frequency analysis for the right channel sound. The horizontal axis represents the time while the vertical axis represents the frequency. The results indicated that Nos. 1 and 2 had attack properties with high sound pressure level at high frequencies. However, for the layout of No. 1, the energy of the sound was not separated, compared to the energy of other positions (the peaks were observed at 35 kHz). Therefore, we believe that sound clarity is correlated to the attack property and sound separation.

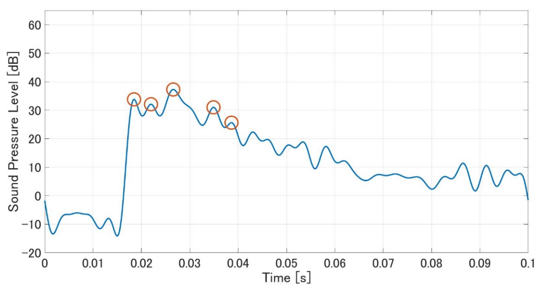
4.2. **Envelope analysis.** It is difficult to compare numerical values from the results of the time-frequency analysis, as Figure 6 showed only qualitative properties. Therefore, time waveforms were extracted at 35 kHz to perform envelope analysis. It was clarified in the preceding paragraph that high frequency band showed sound separation. Figure 7 shows the results of the envelope analysis, where the round marks indicate detected peaks. The horizontal axis represents the time while the vertical axis represents the sound pressure level. To express the attack property and sound separation, peaks over 25 dB were detected. As a result, the attack property and the sound separation were simplified. Furthermore, the attack property was expressed at high sound pressure levels, with many peaks detected at high clarity.



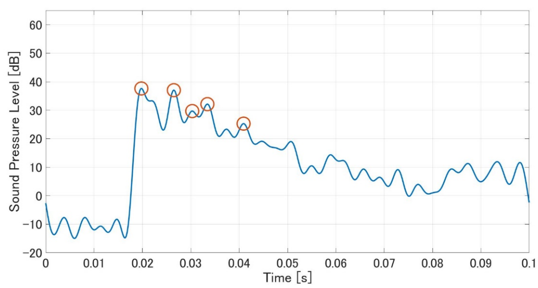
(a) No. 1



(b) No. 2



(c) No. 3



(d) No. 4

FIGURE 7. Results of envelope analysis for each loudspeaker position

5. Multiple Regression Analysis.

5.1. **Prediction model of evaluation by nonprofessional participants.** From the results of the transfer function and envelope analysis, prediction models were calculated using multiple regression analysis [8]. First, to calculate the prediction model of evaluation by nonprofessional participants, the following parameters were employed:

Dependent variable

- Evaluation of "Music A" by nonprofessional participants

Explanatory variable

- Sound pressure level of the attack quality
- Number of detected peaks

- Minimum time interval between peaks
- Average time interval between peaks
- Characteristic difference

The peaks mean detected peaks in Section 4.2. The characteristic difference represents the difference in the transfer properties of each loudspeaker position between the car and the anechoic room. Data between 0.5 and 2 kHz were used as maximum values because the frequency band clearly shows the differences in the position of each loudspeaker. The data with the highest correlation with the dependent variable were selected as the explanatory variable using the stepwise method. As a result, a regression equation can be expressed by the "number of detected peaks" and the "characteristic difference," as shown in Equation (1).

$$Y = 0.347 + 0.137X_1 - 0.037X_2 \tag{1}$$

where Y represents the evaluation point, X_1 shows the number of detected peaks, and X_2 represents the characteristic difference. In this case, the adjusted coefficient of determination shows a high value of 0.825. Furthermore, the F value was obtained as 0.006 that is quite significant ($F < 0.01$). The results suggest that if there were many peaks, there would be many sound separations. A small characteristic difference implies that the property is flat and sound clarity evaluation by nonprofessional participants is high.

5.2. **Prediction model of evaluation by professional audio estimator.** The prediction model of evaluation by a professional audio estimator was calculated using multiple regression analysis. Using the stepwise method, a regression equation could be expressed by the "average time interval between peaks" and the "characteristic difference", as shown in Equation (2).

$$Y = 8.060 + 206.066X_1 - 0.022X_2 \tag{2}$$

where Y represents the evaluation point, X_1 average time interval between peaks, and X_2 the characteristic difference. In this case, the adjusted coefficient of determination showed a value of 0.381. Furthermore, the F value was obtained as 0.130 that is not significant ($F > 0.1$).

6. **Conclusion.** In this study, we investigated and evaluated the sound quality of loudspeakers by changing their layout in a car. Four pairs of loudspeakers were mounted around the driver's seat at different positions. The sound clarity was analyzed using transfer functions based on the critical bandwidth, auditory experiments, and time-frequency analysis. An auditory experiment was conducted to investigate the relationship between physical factors and auditory impression. Finally, a prediction model was calculated using multiple regression analysis, with the following clarified:

- The sound clarity is not only correlated with the magnitude property of the transfer function.
- The sound clarity is correlated with the attack property and sound separation.
- If there are many sound separation and the transfer property is flat, the sound clarity evaluation by nonprofessional participants become high.

From the above, it is believed that the proposed method enables an objective evaluation of sound clarity in audio systems. However, nonlinear distortions were not taken into account because the auditory experiment and time-frequency analysis were conducted using convolved sounds. Therefore, further research on the prediction model is needed using reproduced sound.

In future studies, we intend to analyze reproduced sound and consider the influence of nonlinear distortions.

REFERENCES

[1] A. Miskiewicz, Concert hall sound clarity: A comparison of auditory judgments and objective measures, *Archives of Acoustics*, vol.37, no.1, pp.41-46, 2012.

[2] S. Fajt, Method of evaluating the quality of room acoustics based on energy relations of sound, *Technical Journal*, vol.8, no.3, pp.222-228, 2014.

[3] J. S. Bradley, Review of objective room acoustics measures and future needs, *Applied Acoustics*, vol.72, no.10, pp.713-720, 2011.

[4] F. Satoh, Measurement techniques of room impulse response, *The Journal of the Acoustical Society of Japan*, vol.58, no.10, pp.669-676, 2002.

[5] S.-I. Iwamiya, *Science of Tone Color*, CORONA Publishing Co., Ltd., Japan, 2010.

[6] T. Hideyuki, Practical statistical tests and machine learning-III: Significance tests for human subjective tests, *Systems, Control and Information*, vol.58, no.12, pp.514-520, 2014.

[7] S. Shin and K. Nakano, *The Illustrated Wavelet Transform Handbook*, Asakura Publishing Co., Japan, 2005.

[8] G.-I. Satake and K. Noguchi, *Introduction Statistical Theory*, Chuokeizai-Sha Holdings, Inc., Japan, 1994.