

Button-Sound-Quality Evaluation for Car Audio Main Units

Shunsuke Ishimitsu

Abstract Society widely appreciates the idea of sound being a normal part of a product's operation. As a result, much attention has been directed at designing various sounds that are treated as noise, such as automobile acceleration. Car drivers detect variations in the sound characteristics between different buttons of an audio system; e.g., the pitch, tone color, loudness, and duration. These characteristics can affect the desirability of both a car and its audio system. In this study, we evaluated the sound design of transient signals for 11 different button sounds. To accurately represent button sounds, one of the time–frequency representations, wavelet transform, which structure is similar to an auditory time–frequency resolution feature, is used. An impression was extracted using the semantic differential method, and the relationship between the representation of the wavelet transform and its sound impression was investigated.

Keywords Sound design • Button sound • Audio • Car interior • Wavelets

1 Introduction

Society recognizes the idea of sounds being a normal part of product operation. As a result, much attention has been directed at designing various sounds that are treated as noise, such as the sound of an automobile accelerating or a vacuum cleaner [1–3]. The mechanical sounds generated by the buttons of a car audio unit have been found to contribute to a user's perception of the car itself [4]. Drivers detect variations in sound characteristics, such as the pitch, tone color, loudness, and duration, between

S. Ishimitsu (✉)

Hiroshima City University, 3-4-1, Ozuka-Higashi, Asa-Minami, Hiroshima 731-3194, Japan
e-mail: ishimitu@hiroshima-cu.ac.jp

different buttons; these characteristics appear to have a psychological effect on the driver. The restful, lower sound generated by executive cars gives the impression that the button sounds are integral parts of the car's luxury status. As another example of the relationship between the quality of a sound and a person's perception of reality, it is possible that some of the characteristics previously thought to relate to the tactile sensations of striking a golf ball [5] are actually influenced more by the sound of the impact. Kuwano et al. [6] asked 10 subjects to rate sounds using seven-point adjective scales ranging "hard-soft," "sharp-dull," "refreshing-not refreshing," "powerful-weak," and "vivid-dead." Strong correlations were obtained between the subject's perceptions and the psychoacoustic metrics [7] of loudness and sharpness of the measured impact sounds. These metrics are widely employed in sound-quality evaluations and were also used in these two studies despite being developed for steady, continuous sounds. However, impact sounds are not steady and continuous. The features of these sounds are widely extracted via frequency analysis based on the Fourier transform. The analysis also assumes that the signal being analyzed is periodic and stationary. To accurately represent button sounds, it is necessary to analyze time characteristics as well as frequency characteristics; this is known as a time-frequency ($t-f$) representation. Typical tools of analysis include the spectrogram, Wigner distribution [8], and wavelet transform (WT) [9]. The ($t-f$) resolution features of a WT are characterized by a multiple structure with high-frequency resolution in the low-frequency range and high time resolution in the high-frequency range. This structure is similar to an auditory ($t-f$) resolution feature. This method, used in applications in a wide range of fields [10, 11], performs analysis using the affine transformations (similarity transformations and translations) of a base function known as an analyzing wavelet (AW), whose distribution is localized in both time and frequency.

In this chapter, an impression was extracted employing the semantic differential (SD) method, and the relationship between the representation of a WT and its sound impression was revealed.

2 Time-Frequency Analysis

2.1 Wavelet Transform (WT)

A WT is used in multi-resolution analysis to match auditory ($t-f$) resolutions. The WT is obtained by calculating the inner product of the signal $f(t)$ and AW $\psi(t)$ [12, 13]:

$$WT_f(b, a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \psi^* \left(\frac{t - b}{a} \right) f(t) dt \quad (1)$$

Variable a is a scale parameter used in similarity transformations. Variable b is a shift parameter used in the translation of $\psi(t)$. The WT is initially expressed in the $(t - f)$ time–scale plane, but can be regarded as an approximation of the $(t - f)$ distribution using a time- and frequency-localized AW. We selected a Morlet wavelet in a preliminary experiment.

2.2 Experimental Condition

Eleven types of button for use in six car audio unit models were evaluated. Data were recorded in an anechoic chamber. Each button was pushed three times, and the sound was recorded with a microphone placed about 30 cm from the car audio unit. The sound made by pushing the button (push sound) and the sound made by releasing the button (back sound) are depicted.

2.3 Analysis Conditions and Results

Psychoacoustic metrics (such as the measurements of loudness and sharpness) are used to numerically represent the psychoacoustical features of hearing [7]. The ISO standard 532B relates to loudness for stationary sound. The metrics are widely employed in sound-quality evaluations and were also used in these studies [5–7] despite their having been developed for steady, continuous sounds. However, loudness is used to evaluate stationary sounds and cannot be adequately used to evaluate non-stationary sounds such as button sounds.

As a result, we decided to compare the time–frequency structure and the jury test score.

WT results for each main unit are shown in Figs. 1, 2, 3, 4, 5, and 6. Each figure includes the discrete Fourier transform (DFT) magnitude and the WT of push sounds and back sounds. Morlet wavelets were used as the AW.

Low-frequency button sounds tended to receive a high score in the jury test. As the sound frequency increased, the evaluation score decreased. Evaluation scores were also affected by the duration of the energy burst.

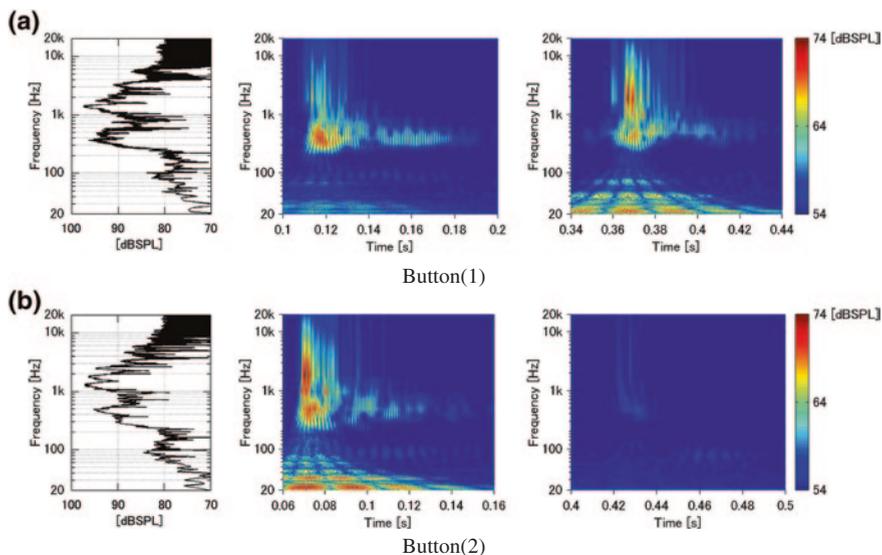


Fig. 1 Magnitude response of the main unit 1: (left), DFT of push and back sound; (center), WT of push sound; (right), WT of back sound. **a** Button (1). **b** Button (2)

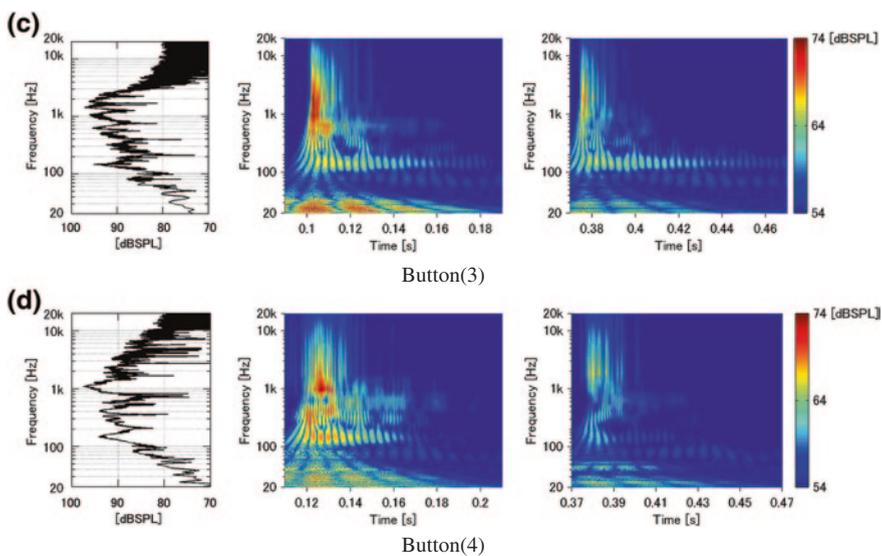


Fig. 2 Magnitude response of the main unit 2: (left), DFT of push and back sound; (center), WT of push sound; (right), WT of back sound. **c** Button (3). **d** Button (4)

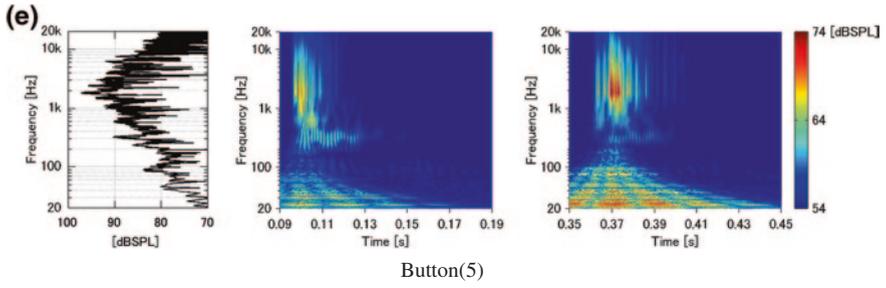


Fig. 3 Magnitude response of the main unit 3: (left), DFT of push and back sound; (center), WT of push sound; (right), WT of back sound. e Button (5)

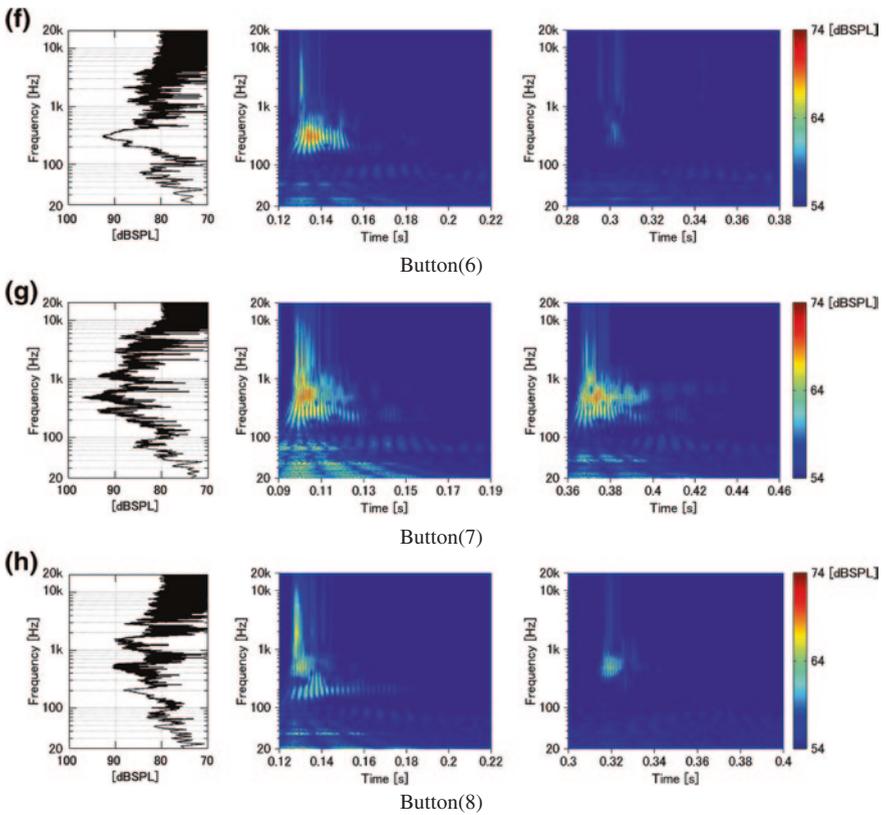


Fig. 4 Magnitude response of the main unit 4: (left), DFT of push and back sound; (center), WT of push sound; (right), WT of back sound. f Button (6). g Button (7). h Button (8)

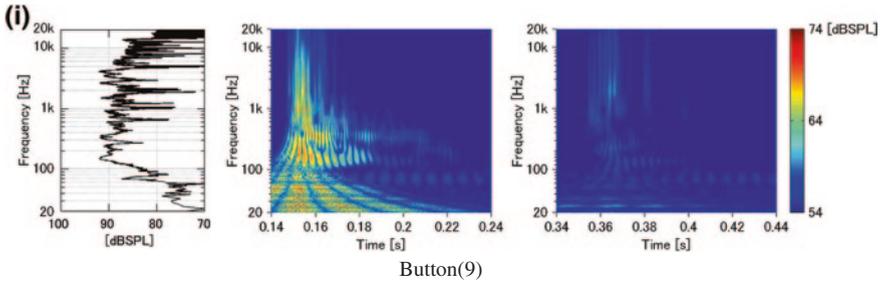


Fig. 5 Magnitude response of the main unit 5: (left), DFT of push and back sound; (center), WT of push sound; (right), WT of back sound. **i** Button (9)

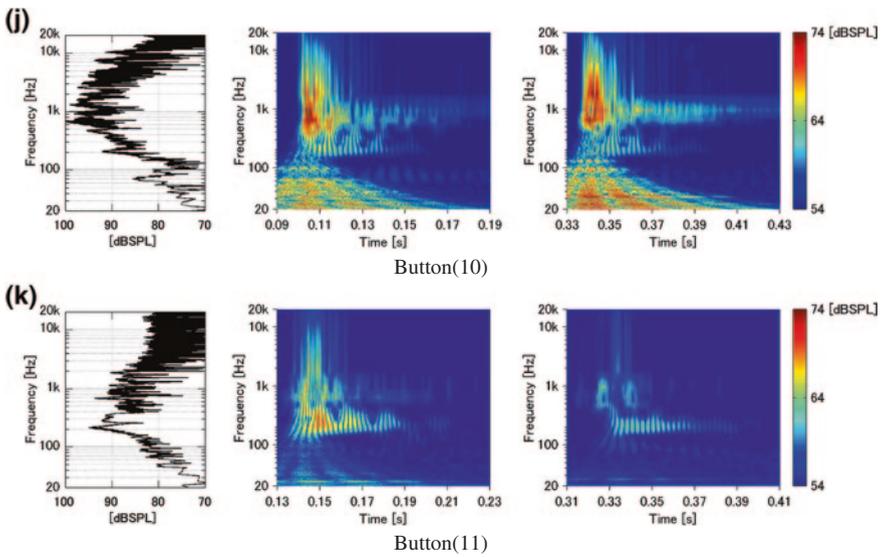


Fig. 6 Magnitude response of the main unit 6: (left), DFT of push and back sound; (center), WT of push sound; (right), WT of back sound. **j** Button (10). **k** Button (11)

3 Jury Test

3.1 Quantification of Psychoacoustics

“Loudness,” “pitch,” and “sound quality and tone” were used for psychoacoustic quantification [7]. Loudness is equivalent to a physical quantity, known as the sound pressure level. Sound frequency is related to pitch, and the time-varying

structure and spectrum are related to sound quality and tone. These are called the sensation dimensions. However, when sound quality is examined, dimensions such as “brightness” and “hardness” can be found. Quantifying the number of dimensions involved in producing the magnitude and pitch of sound needs to be simple. The SD method allowed us to quantify the sensation dimensions in the experiment. The method employs many adjective scales expressing sound quality and tone, and it measures sound using these scales. Factor analysis was carried out to evaluate the common factors from these results, determine, and quantify dimensionality.

3.2 Auditory Experiment and Result

A jury test was conducted employing the SD method with 67 healthy people forming the jury. Sounds were reproduced through headphones. The evaluation paper used in the jury test is shown in Fig. 7. The age and gender distribution of subjects in the jury is shown in Fig. 8. The experimental results are shown in Fig. 9.

First, each adjective was matched with a button.

Button (6)—“dark”—“deep”—“simple”—“soft”—“heavy”—“like”—“high-class”—“low”—“charming”

Button (11)—“round”—“warm”—“fresh”—“natural”

Button (7)—“beautiful”—“pleasant”—“relaxed”—“heart”

Button (3)—“simple”

Button (1)—“common”—“simple”

Button (8)—“small”—“weak”—“fine”—“thin”—“unsatisfactory”—“delicate”—“short”

Button (2)—nothing

Button (9)—nothing

Button (5)—“light”—“thin”—“high”—“cold”

Button (10)—“large”—“strong”—“clear”—“force”—“bold”—“showy”—“hard”—“long”—“artificial”—“dry”—“coarse”—“thick”—“bright”—“sharp”—“cheap”—“jarring”

Button (4)—“loose”—“uneasy”—“dislike”—“complicated”—“dirty”—“boring”—“blurred”

WT showed that the back sound of button (4) produced a sweep sound at 100–600 Hz. The adjectives “loose,” “uneasy,” and “dirty” were associated with this sweep sound distribution. Continuous low-pitched sounds were associated with button (3), but these sounds did not affect the auditory impression. A continual high-pitched sound was associated with button (10), and this sound was described as “long” and “cheap.” However, sound-quality matching was insufficient because the correlation was between adjectives.

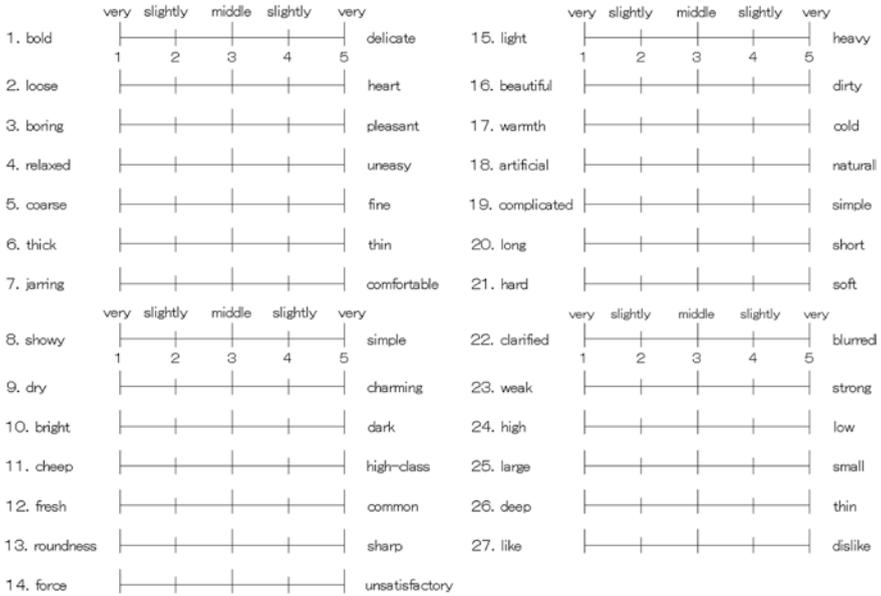
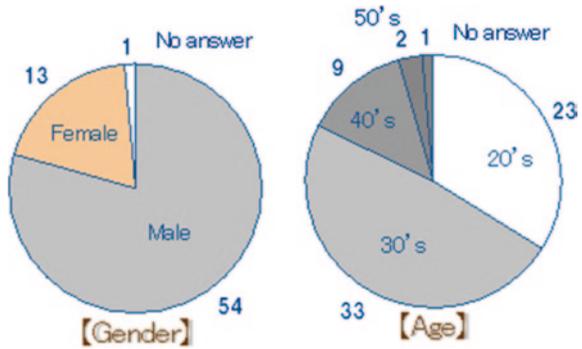


Fig. 7 Evaluation paper

Fig. 8 Subjects



3.3 Factor Analysis

Factor analysis and WT matching analysis of the experimental results were carried out. The relationship between each button sound and factor was investigated by selecting 10 pairs of significantly different adjectives from a set of 27 pairs of adjectives. Factor loadings and factor scores are shown in Table 1 and Fig. 10, respectively. The principal divisor method and the varimax rotation method were used for factor extraction.

A metallic factor, an esthetic factor, and a force factor were extracted sequentially from the first factor. These are “hard,” “comfortable,” and “force” sounds;

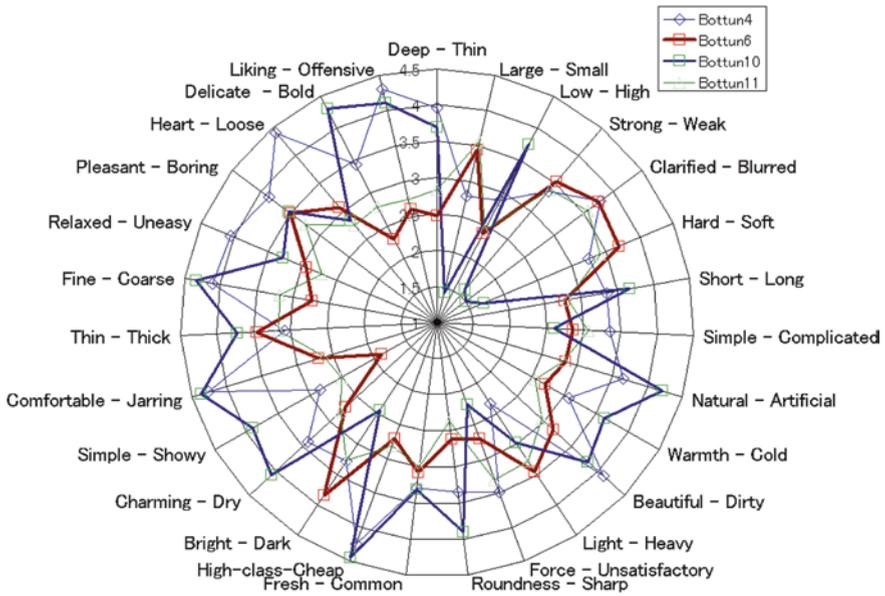


Fig. 9 Subjective evaluation

Table 1 Factor loadings

	Metallic	Esthetic	Force	Similarly
Hard-soft	0.91	-0.20	0.33	0.97
Clear-blurred	0.90	-0.10	0.41	0.98
Low-high	0.86	-0.39	-0.01	0.90
Simple-showy	-0.63	0.52	-0.46	0.88
Comfortable-jarring	-0.22	0.95	-0.14	0.96
High-class-cheap	-0.36	0.91	0.12	0.96
Fine-coarse	-0.09	0.88	-0.42	0.95
Force-unsatisfactory	0.15	-0.02	0.95	0.92
Strong-weak	0.65	-0.20	0.74	1.00
Large-small	0.50	-0.45	0.71	0.95
Factor contribution	36.41	32.02	26.48	94.90

therefore, their jury test scores were high. Moreover, the accumulation contribution was fully satisfied.

“Like” button sounds had low metallic factor and force factor scores and had a high esthetic factor score, as shown in Fig. 10. When a button sound was assigned the “dislike” property, the metallic factor and force factor scores were high, but the esthetic factor score was low. These results correspond with the results of WT experiments shown in Figs. 1, 2, 3, 4, 5, and 6.

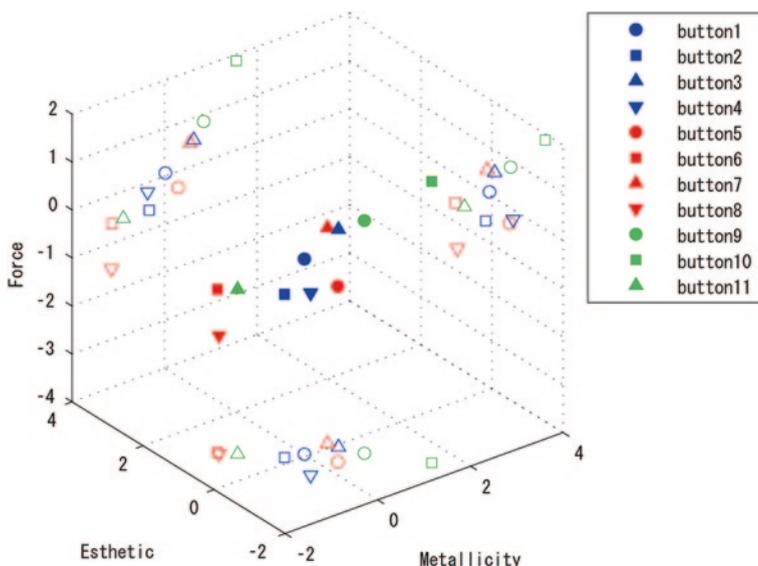


Fig. 10 Factor score

4 Conclusions

We evaluated sound design using 11 different button sounds. Although psychoacoustic metrics are widely employed in sound-quality evaluations, loudness is used to evaluate stationary sound and cannot be adequately used in the evaluation of non-stationary sounds such as button sounds. As a result, we decided here to compare time–frequency structures and jury test scores to evaluate sound quality. First, an impression was extracted employing the SD method, and the relationship between the representation of the WT and its sound impression was investigated. After matching WT characteristics and auditory impressions, the results showed that a low-frequency button sound made a favorable impression, and a high-frequency button sound made a negative impression. The auditory impression of both button sounds was classified into esthetic, metallic, and force factors on the basis of these results. “Like” button sounds had low metallic factor and force factor scores and had high esthetic factor scores. When a button sound was assigned the “dislike” property, the metallic factor and force factor scores were high, but the esthetic factor score was low. These results should assist button sound design in the future.

References

1. Ishimitsu S, Kobayashi H (2006) Study on instantaneous correlation analyses of acceleration car interior noise using wavelets and its subjective evaluation. *Trans Jpn Soc Mech Eng* 72(719) C:2094–2100
2. Toi T (2004) Recommendation of a comfortable sound design and its procedure. *Mech Des* 48(2):36–45
3. Kuwano S (2007) Design of sound environment. Corona, Japan
4. Ishimitsu S, Sakamoto K, Arai T et al (2008) Study on the visualization of the impression of button sounds. In: The proceedings of 3rd international conference on innovative computing, information and control
5. Roberts JR, Jones R, Mansfield NJ, Rothberg SJ (2005) Evaluation of vibrotactile sensations in the ‘feel’ of a golf shot. *J Sound Vib* 285(1–2):303–319
6. Kuwano S, Namba S, Ohta K et al (1999) Relation between envelope pattern and sound quality of impulsive sounds. *J Acoust Soc Jpn* 20(2):153–155
7. Zwicker E (1982) *Psychoakustik*. Springer, Berlin Heidelberg
8. Claasen TACM, Mecklenbrauker WFG (1980) The Wigner distribution—a tool for time-frequency signal analysis, part 1: continuous-time signals. *Phillips J Res* 35(3):217–250
9. Daubechies I (1992) Ten lectures on wavelets. Society for Industrial and Applied Mathematics, Philadelphia
10. Rioul O, Vetterli M (1991) Wavelets and signal processing. *IEEE Signal Process Mag* 10:14–38
11. Mustapha O, Lefebvre D, Khalil M, Hoblos G, Chafouk H (2009) Fault detection algorithm using DCS method combined with filters bank derived from the wavelet transform. *Int J Innovative Comput Inf Control* 5(5):1313–1327
12. Daubechies I (1992) Ten lectures on wavelets. Society for Industrial and Applied Mathematics, Philadelphia
13. Zhang Z, Ikeuchi H, Toda H, Miyake T, Imamura T et al (2008) Designing average complex real signal mother wavelet and applying it in abnormal signal detection. *Int J Innovative Comput Inf Control* 4(4):1009–1022