



Sound Design and Sensitivity

KONISHI Masahide

1 Introduction

In recent years, various industries, including automotive and construction machinery, have intensified initiatives to achieve a carbon-neutral society, accelerating the development of electrification technologies. In line with this trend, the construction machinery sector has introduced fully electrified machines without hydraulic systems, as well as hybrid hydraulic construction machines that are powered by electric motors instead of existing engines. For example, battery-powered hydraulic excavators have been developed to eliminate the need for diesel engines and enable the machines to be driven by electric motors. These machines achieve zero exhaust emissions and substantial noise reduction while delivering operating performance equivalent to conventional models. However, in these electrified construction machines, the elimination of engines, which were the main source of noise, may enable mechanical components to produce relatively more noticeable noise during operation, which was previously masked by engine noise. This means that the operating noise of hydraulic pumps, KYB's core product, would be more noticeable. Consequently, customers may demand higher noise standards, which could lead to complaints if not met. In this situation, the need for further noise reduction increases.

Although KYB has developed various noise reduction technologies, hydraulic pumps inevitably produce noise due to rapid pressure changes in the pump chambers during suction and discharge, as well as pressure pulsations in the piping caused by these changes. Due to inherent limits, it is difficult to fully suppress the sound. This study focuses on "comfortable sound design"^{Note 1)}, which converts noise into comfortable sounds in the auditory sense.

In this paper, we investigated the relationship between sound and human sensitivity in order to incorporate "comfortable sound design" approaches into product development to improve sound quality. Specifically, we conducted a sensory evaluation to determine people's impressions of sounds generated by hydraulic pumps, then analyzed the relationship between these impressions and the pumps' acoustic

characteristics.

This paper provides an overview of these efforts and the insights obtained from them.

2 Overview and Application Examples of Comfortable Sound Design

In modern society, people live surrounded by a variety of products that are becoming quieter year by year as technology advances. However, simply reducing the sound pressure level does not necessarily create a comfortable sound environment for users. In fact, people may perceive certain low-level-sounds as grating or uncomfortable. Such sounds can create a negative impression, resulting in product rejection or complaints, even if the technology meets all quality requirements.

In this context, a new design concept called "comfortable sound design" has recently gained attention. This approach considers human sensitivity. Rather than merely reducing the sound pressure level as in conventional noise reduction measures, comfortable sound design aims to convert product sounds into "audible sounds unlikely to be perceived as uncomfortable" or "sounds perceived as comfortable." By transforming negative noise into positive, comfortable sound, this approach can add value to products.

Many companies have already implemented initiatives for comfortable sound design. One example is the automotive industry. As electric vehicles have grown in popularity, cars have grown quieter due to the absence of engine noise. Traditionally, engine sound has been an important element symbolizing the character and performance of vehicles. For this reason, simulated engine sounds^{Note 2) 1) 2)} are deliberately added. This simulated sound is not mere imitation. Rather, it is a sound design that combines psychological comfort, functionality, and safety. It also serves as feedback for acceleration and accelerator operation.

Note 1) May be referred to by different names by different companies, such as timbre or sound development.

Note 2) May be referred to by different names by different companies.

Furthermore, suppliers are studying the acoustic designs of air conditioners, mufflers, and other components with human sensitivity in mind³⁴⁾. These initiatives demonstrate the effectiveness of comfortable sound design and are gaining attention. The results and technical information of related research are gradually being disclosed. Additionally, comfortable sound design is expected to help companies build their original brand image by leveraging the acoustic characteristics of their products. Sound design differentiation could be a competitive advantage.

3 Human Auditory Characteristics and Sensory Evaluation

Human hearing is more than just the ability to detect the presence or absence of sound. It is a sophisticated sensory organ capable of processing multiple pieces of information simultaneously, including sharpness (frequency), loudness (sound pressure), direction, and temporal changes. The audible range is generally between 20 Hz and 20 kHz, though it varies by individual and age. Even within this range, sensitivity to sound varies by frequency.

Furthermore, hearing is affected not only by physical sound characteristics, but also by psychological and perceptual factors. In other words, people may have different subjective impressions of the same sound pressure level depending on the timbre, duration, or timing of sound onset. For example, a sound may be perceived as “noisy” or “pleasant.”

One quantitative approach to evaluating sensitivity is to use sound quality evaluation indices based on psychoacoustics. These indices quantify human psychological responses to sound and play an important role in designing and evaluating product sounds. Table 1 shows typical sound quality evaluation indices.

Table 1 Typical sound quality evaluation indices

Item	Description
Loudness	An indicator of sound volume as it is perceived by humans. Unlike physical sound pressure levels, this index considers frequency characteristics to evaluate perceived loudness.
Sharpness	An indicator of the acuteness or intensity of sound. Sounds with more high-frequency components are perceived as sharper.
Roughness	An indicator of the coarseness or uncomfortableness of sound. Sounds with greater temporal changes due to amplitude modulation show higher roughness.

Additionally, acoustic signal processing methods that consider human auditory characteristics related to sound perception can also be used as an

effective approach for comfortable sound design. Table 2 shows typical methods.

Table 2 Typical acoustic analysis methods

Item	Description
Mel-frequency cepstral coefficients (MFCC)	A method of extracting speech features using a frequency scale based on human auditory characteristics. MFCC is widely used in speech recognition.
Octave band analysis	A method of analyzing sound components by dividing the frequency components of sound into octave bands. It is effective for evaluating noise and designing acoustics.
Formant analysis	A method of extracting the resonant frequencies that characterize vowels from the speech spectrum. It is widely used in speech processing and acoustic analysis.

Using these sound quality evaluation indices and acoustic signal processing methods allows for more than just measuring physical sound pressure levels. It also allows for sound evaluation that aligns with human sensitivity. However, because human sensitivity varies from person to person, subjective evaluations always exhibit variability. Therefore, it is difficult to quantify perceptive impressions, even when using these indices. This study utilized the results of sound quality evaluation index and acoustic signal processing methods multifacetedly to discuss the potential of predicting sensory evaluation. Specifically, we attempted to predict sensitivities with high accuracy by comprehensively analyzing multiple indices. When using multifaceted indices, it is important to accurately determine the relationship between them. This study introduces machine learning as a method to capture these relationships and perform multidimensional analysis. Machine learning can handle nonlinear, high-dimensional data structures, leading us to expect improved accuracy in sensory evaluation.

4 Product Sound Recording

As illustrated in Fig. 1, acoustic noise testing of piston pumps was conducted in our anechoic room. During the test, the target pump was secured to a fixture, and its noise was measured and recorded. Conventional recording uses a single microphone installed one meter away from the pump at the same height, as shown in Fig. 2. However, monaural recording cannot accurately reproduce the actual sound environment because the human ears perceive the direction, distance, and depth of a sound source

based on time and sound pressure differences and frequency characteristics reaching the left and right ears. Therefore, this study adopted binaural recording to create a more natural listening environment for sensory evaluation. Binaural recording uses a dummy head that simulates the shape of the human head and the structure of the auricle. Microphones in the left and right ear positions faithfully reproduce the sound transmission path (Fig. 3).

The KYB piston pump was selected as the subject for acoustic recording, and the experiment was conducted accordingly. To simulate a real-life evaluation situation in which a person would observe the test, we installed the dummy head directly in front of the pump at the height of a person's ears, assuming an average-height adult standing in front of the pump. The effect on spatial localization (the sense of sound direction and distance) is limited because the suction and discharge ports are positioned on the left and right sides of the pump. However, we chose this configuration to prioritize the reproducibility of the actual hearing experience.

Developing a machine learning model requires large volumes of diverse data. However, it is difficult to change the design specifications of actual machines to collect samples with different acoustic characteristics. Therefore, we used audio editing software to manipulate the pump's harmonic components and frequency bands, creating derivative

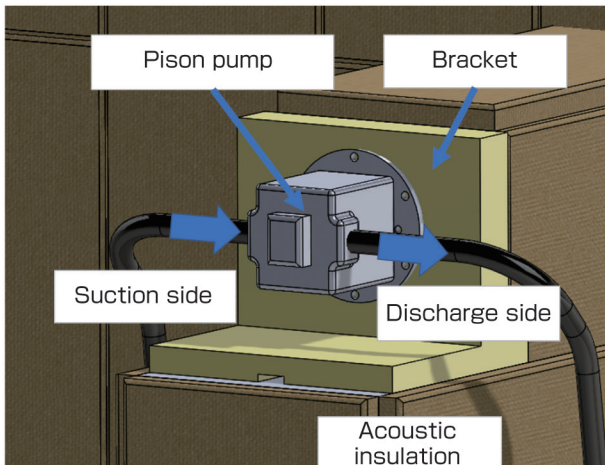


Fig. 1 Test apparatus

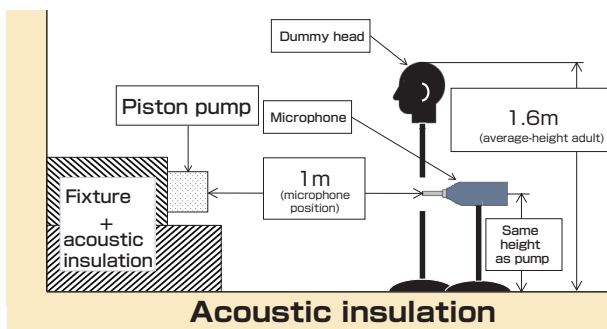


Fig. 2 Installation position of recording equipment (side view)

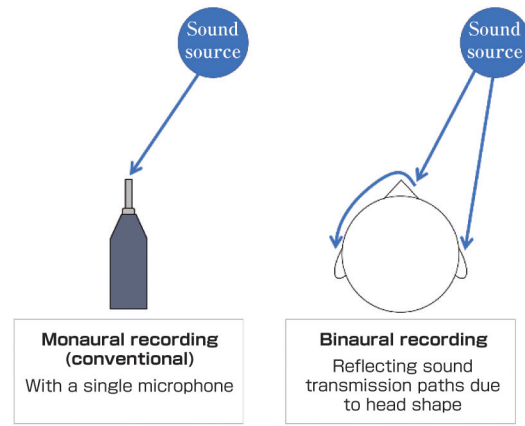


Fig. 3 Difference between monaural and binaural recording methods

sample sounds with intentionally adjusted timbre.

5 Sensory Evaluation Methods

This study used the Semantic Differential (SD) method to assess sensitivity. The SD method is a multi-stage questionnaire that uses paired adjectives, such as “comfortable-uncomfortable” or “sharp-dull,” to enable the quantitative measurement of subjects’ impressions and sensations of an object. When applying this method, it is essential to select evaluation items appropriate for the object in question. Referencing a prior study⁵⁾ on the sensory evaluation of construction machinery, this study conducted a seven-point evaluation of the ten items shown in Fig. 4.

Sound sensory evaluation poses two significant challenges: variability of sensitivity and lack of consistency in subjective assessments.⁶⁾ That is, even for the same sound, the same subject may vary in their assessment, posing a risk of poor reproducibility of assessment results. To address these challenges, we implemented the following two measures for sensory evaluation.

5.1 Providing a Reference Tone

For evaluation purposes, subjects first heard the operating sound of an actual piston pump as a reference tone before hearing derivative sample sounds.

5.2 Verifying Through Reevaluation

The subjects performed a reevaluation using their initial assessment results as a reference to check for contradictions or misunderstandings, thereby improving the accuracy and reliability of the assessments. Although presenting the initial assessment results could introduce bias, reevaluation was introduced to prevent misperceptions and inconsistencies, and confirm evaluator self-consistency.

Fig. 5 shows the sensory evaluation flow incorporating these measures. Based on past

experience with sound sensory evaluation using the SD method, it is empirically recognized that multiple evaluations of a single sound sample by the same subject produce evaluation values that vary by approximately ± 1 point. Therefore, this variation must be considered when using machine learning to predict sensory evaluation in this study. Based on this, an error range of ± 1 point was defined as the tolerance for evaluating prediction accuracy. Robust prediction was considered to be achieved if prediction within this range was possible.

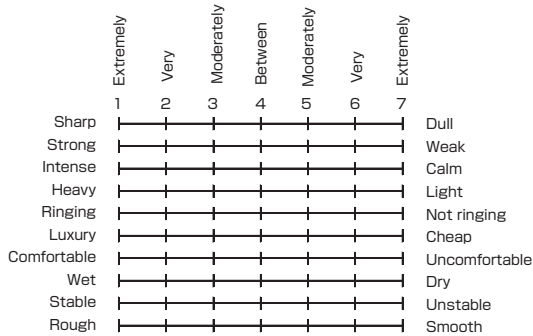


Fig. 4 SD method evaluation chart

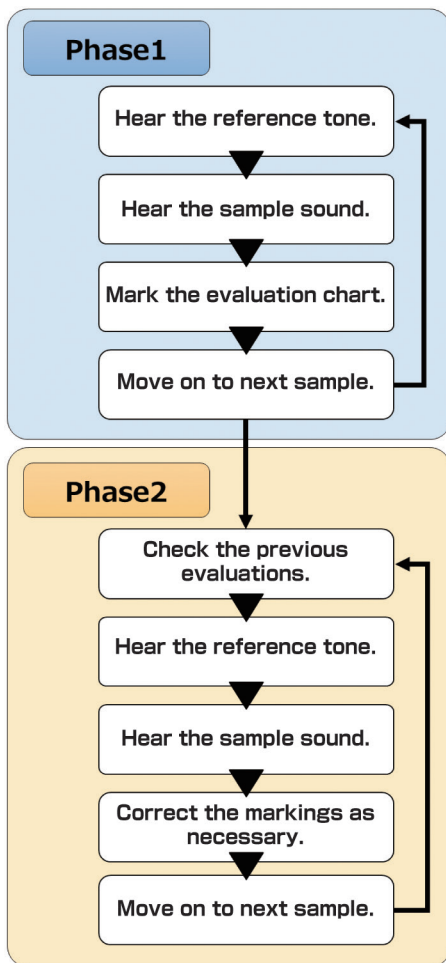


Fig. 5 Sensory evaluation flow

6 Building Prediction Models Using Machine Learning

This study employed machine learning with multiple features based on acoustic analysis to develop a model for predicting sensory evaluation. The following acoustic indicators were used:

- (1) Sound quality evaluation indices (loudness, sharpness, and roughness)
- (2) Mel-frequency cepstral coefficients (MFCC)
- (3) Octave band analysis
- (4) Formant analysis

Using the features obtained through these analyses and the sensory evaluation results obtained by the SD method as teaching data, we built a prediction model with two types of machine learning algorithms. The algorithms were Random Forest (for classification) and Bayesian estimation.

Fig. 6 shows an overview of each algorithm. Random Forest is a machine learning method that creates multiple decision trees from the training data and integrates their prediction results based on the majority rule. This method demonstrated good predictive accuracy in our internal sensory evaluation. This study also adopted it to verify its effectiveness.

In contrast, Bayesian estimation is a method of making predictions by updating a preset probability distribution with learning data. Based on probability theory, this method offers the advantage of flexible data handling, even when the data contains missing values or variations.

By using these two approaches, we aim to build a prediction model that can provide highly accurate and robust estimations, even for data with subjective variations, which are inherent in sensory evaluation.

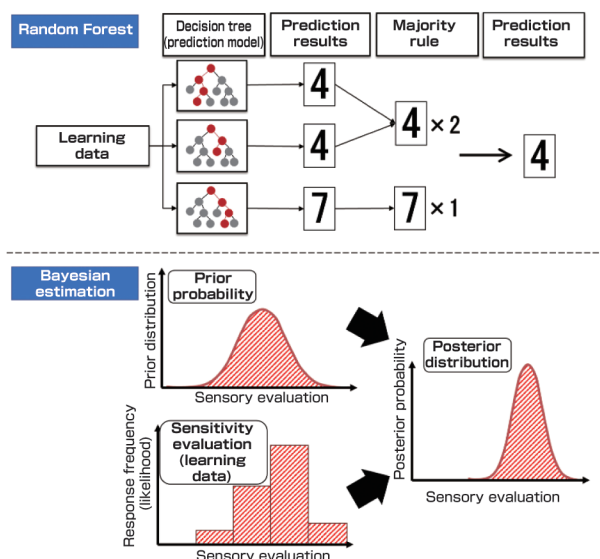


Fig. 6 Overview of algorithms

7 Model Prediction Results and Discussion

Fig. 7 shows the results of the sensory evaluation prediction using the two machine learning algorithms: Random Forest and Bayesian estimation. The horizontal axis shows the actual sensory evaluation values of subjects, and the vertical axis shows the predicted values through machine learning. The red line in the figure is the regression line. Both algorithms presented a correlation coefficient above 0.9, indicating a strong correlation. However, the trend of prediction errors depended on the algorithm used. The Random Forest algorithm's prediction results, shown in Fig. 7(a), suggest the possibility of partial misclassification. This is likely due to the fact that the Random Forest algorithm uses multiple decision trees based on the majority rule to make predictions and only provides the results of each decision tree, which results in wider errors. Additionally, Random Forest might not sufficiently represent continual evaluation variations when used as a classification algorithm because the predicted values are limited to discrete integers (1, 2, 3, etc.).

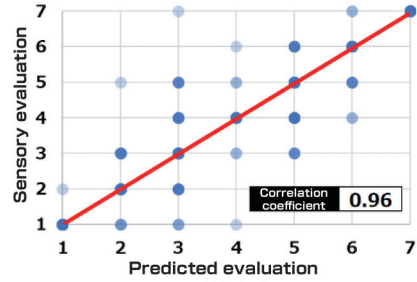
The prediction results of Bayesian estimation, shown in Fig. 7(b), present no large errors, unlike Random Forest. The predicted values generally align with the actual evaluation values. This is because Bayesian estimation reflects evaluation trend data by integrating prior distributions and observation data into the model. Furthermore, Bayesian estimation can also handle the uncertainty of the predicted values, making it possible to respond flexibly to sensory evaluation with large variations. These advantages likely contributed to the positive outcome.

Fig. 8 shows the distribution of the prediction errors (the difference between the evaluation and the predicted values) for each algorithm. Despite the low frequency of occurrence of errors, Random Forest involved cases with maximum errors reaching 5 points, demonstrating wide prediction errors. Bayesian estimation, on the other hand, showed an error range of ± 2 points, demonstrating its ability to make more consistent predictions.

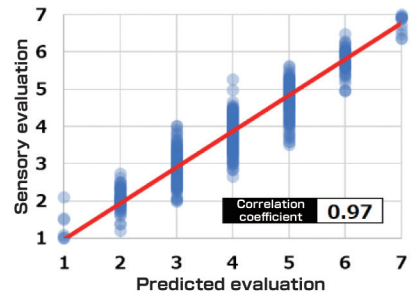
Table 3 shows the accuracy rates for each algorithm. The table contains two accuracy rates: one calculated with an evaluation error tolerance of ± 1 point, as described in Chapter 5, exact-match accuracy. According to Table 3, Bayesian estimation has a prediction accuracy of 99% with a tolerance of ± 1 point, while Random Forest has an accuracy of 84%. Bayesian estimation offers a lower misprediction risk due to its small errors, indicating its effectiveness in predicting sensory evaluation.

These results indicate that sensory evaluation can be predicted, albeit within the scope of using actual piston pump sounds and derivative sample

sounds generated by acoustic editing. However, it is important to note that these derivative samples used in this study were generated by acoustic editing and it remains unverified whether similar acoustic characteristics can be reproduced in actual products. Therefore, it will be necessary to verify the reproducibility of sounds in actual machines and design adjustments during implementation.



(a) Random Forest (classification)



(b) Bayesian estimation

Fig. 7 Sensitivity prediction through machine learning

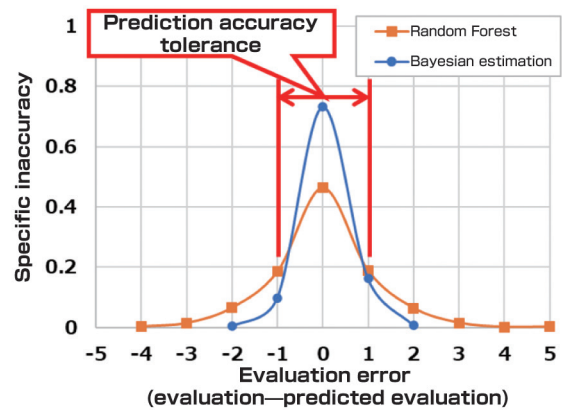


Fig. 8 Evaluation errors

Table 3 Algorithm accuracy rates

Algorithm	Accuracy rate (± 1 point error tolerance)	Accuracy rate (100% match)
Random Forest	84%	46%
Bayesian estimation	99%	74%

8 Future Challenges and Outlook

This study suggests that predicting human sensory evaluation is possible by combining acoustic analysis indices with machine learning. In particular, the prediction model using Bayesian estimation proved highly accurate and stable, demonstrating its effectiveness in quantitatively predicting sensory evaluation. However, several challenges remain. First, the detailed mechanisms by which each acoustic index contributes to or impacts sensory evaluation have not been fully determined. Clarifying the causal relationships through feature selection and significance analysis will be necessary. Second, although it is relatively easy to reproduce comfortable sounds using acoustic editing software, there are no specific guidelines regarding design changes or technical points necessary to implement such acoustic features in actual products. In other words, the technological foundation to ensure consistency between sensitivity-based acoustic design and product design has not been properly established.

Therefore, the future outlook involves the following two important areas of research:

8.1 Quantitatively Elucidating the Relationship Between Acoustic Indices and Sensory Evaluation

We will identify the acoustic elements that contribute to sensory evaluation and visualize their impact by introducing a future contribution analysis and explainable AI (XAI) technology.

8.2 Establishing Techniques for Implementing Acoustic Characteristics in Actual Machines

To reflect optimal, comfortable sounds in actual product design, we aim to commercialize sensitivity-based acoustic design. This involves advancing research on the relationship between acoustic design, structural design, material

selection, and control technology.

Addressing these challenges will facilitate the integration of sensitivity engineering and product development, resulting in more user-oriented product design.

9 In Closing

Finally, I would like to take this opportunity to express my deepest gratitude to the staff of the relevant departments for their tremendous support and cooperation in advancing this activity.

Thanks to dedicated efforts of everyone in the relevant departments, we were able to proceed smoothly with this activity. I am sincerely grateful.

References

- 1) KUBO Norio: "Signal Processing in sound designing of additive engine sound for hybrid and electric vehicle," Pleasant Sound Technology, S&T Publishing inc. (2012).
- 2) YAMAUCHI Katsuya: "Overview of recent issues in sound design for next-generation vehicles due to their quietness," THE JOURNAL OF THE ACOUSTICAL SOCIETY OF JAPAN (2017).
- 3) SANGO Co., Ltd., <https://sango.jp/>
- 4) TAKAO Hideo: "Sound Quality Evaluation by SD Method for Passenger HVAC Noise," 2025 Spring/ JSAE Congress (Technical Sessions), Society of Automotive Engineers of Japan (JSAE).
- 5) KAWAGUCHI Masataka: "Study on Sound Quality Evaluation of Exterior Tonal Sound for Construction Machinery," Transactions of the Japan Society of Mechanical Engineers (1992).
- 6) NAMBA, KUWANO: "Psychological measurement method for sound evaluation," CORONA PUBLISHING CO., LTD. (1998).

Authors



KONISHI Masahide

Joined the company in 2012.
Elemental Technology Research
Sect., Basic Technology R&D
Center, Engineering Div.
Engaged in research on new
technologies and vibration and
noise-related studies.