Research Article

Research on the Sound Quality Evaluation Method Based on Artificial Neural Network

Xiedong Song[1,2,3] and Wei Yang [4]

[1]*College of Technology, Yantai Nanshan University, Yantai 265713, China*
[2]*Technology Department, Xiaoxing (Shandong) Internet Technology Co.Ltd, Yantai 264000, China*
[3]*College of Computing and Information Technology, National University, Manila 0900, Philippines*
[4]*School of Digital Creativity and Animation, Shenzhen Polytechnic, Shenzhen 518000, China*

Correspondence should be addressed to Wei Yang; yangwei@szpt.edu.cn

For the improvement of the traditional evaluation effect of the automobile sound quality, an evaluation model of automobile sound quality is constructed based on BP neural network. The first is to introduce the basic principle of the BP neural network in detail. The second is to use the MGC parameters to establish the vehicle interior sound conversion model. The converted sound characteristic parameters are taken into the WORLD model to synthesize the new sound signals. Furthermore, the wavelet decomposition method is used to remove noise from the synthesized sound signals. Finally, a sound evaluation model based on BP neural network is established. The sound quality of automobiles can be better evaluated by carrying out the ABX test and MOS test in the field of sound conversion. For the newly synthesized sound signal and the target sound signal, it can be seen that the newly synthesized sound signal is more inclined to the target sound signal, and the sound quality is better. In addition, the sound quality is tested through loudness, roughness, sharpness, and level A in the field of sound quality evaluation. The final results show that the quality of newly synthesized sound is better, and the average errors of sound signals meet the sound standard. Therefore, the constructed sound conversion model and the sound evaluation model are feasible and effective.

1. Related Work

As science and technology continuously develop, people pay more and more attention to the performance of automobiles, so the noise problem has become one of the important indicators of automobile purchase. The noise of automobiles seriously affects the comfort of the person when riding in the vehicle, also interferes with the communication between people, and even seriously damages the auditory characteristics of human ears. At the same time, vibration and noise are also very different in the automobile sound. Therefore, how to better distinguish vibration and noise, and evaluate the automobile noise, has become the focus of current research. In practice, companies such as Nissan, FEV of Germany, and AVL of Austria are all studying the interior sound quality of electric vehicles. The most influential noise sources and special frequencies have been found. Furthermore, some active or passive technologies are used to eliminate noise and compensate for frequency, which achieves good results.

In the academic field, Xie et al. proposed an adaptive neural network sound evaluation method. The characteristics of the proposed method are simple and fast. At the same time, they introduced a mainstream neural network algorithm [1]. Yin et al. calculated the dither noise of automobile side windows by using the large eddy simulation method. As can be seen that the variation rules of loudness, sharpness, roughness, and undulation with wind speed and window opening are obtained [2]. Yang et al. adopted binaural transfer path analysis (BTPA) to measure the vibration and noise transmission paths of automobiles under transient and steady-state conditions. Then, the characteristics and differences of the internal noise of automobiles under different operating conditions are studied, where the loudness, sharpness, roughness, and A-weighted sound

pressure level are adopted [3]. Park and Kang established a sound quality index model, which reflects the reviewers' different styles. Also, the model is constructed by utilizing K-means clustering, factor analysis, and multiple linear regression. This study provides an additional reference for the evaluation of sound quality [4]. Zhao et al. proposed a DBN based on linear regression (LR-DBN). There are 6 psychoacoustic indicators and 26 Mel frequency cepstrum coefficients being taken as input characteristics. At the same time, the ordinary DBN, multiple linear regression (MLR), and backpropagation neural network (BPNN) are used to verify the performance of LR-DBN. What can be seen is that compared with other methods, the correlation coefficient of LR-DBN is higher, and the prediction error is lower. Furthermore, LR-DBN is more stable. So, this method is a reliable method for evaluating EEV sound [5]. Wang et al. proposed an objective evaluation method of interior noise, which is based on the displacement of the human basal membrane [6]. Firstly, noise samples of different seats are collected under different running conditions. Secondly, the comparison method of adaptive grouping pairs is used to obtain the subjective evaluation value of noise samples. Thirdly, the total parameter model of the human ear is adopted to calculate the average value for the basal membrane displacement response (SMVBMDR). So, the characteristic matrix based on SMVBMDR is established. Finally, there are two BP neural network models constructed to evaluate the interior noise sound quality respectively. Also, they took the traditional psychoacoustic indicators and extracted the feature matrix as the input. The SMVBDMDR is very correlated with SEV. Moreover, the sound quality prediction accuracy of the model based on SMVBMDR is higher.

It can be seen from the above research that the qualitative and quantitative methods are used for the present sound quality evaluation. Also, machine learning algorithms, deep learning algorithms, and other algorithms are introduced. However, the above sound evaluation rarely involves the preprocessing of noise data. In this paper, it is proposed that on the basis of parameter extraction of automobile sound feature extraction, BP is used to synthesize automobile sound, and then evaluate sound according to sound quality evaluation parameters so as to provide a new reference for automobile sound quality evaluation.

2. BP Neural Network

2.1. Introduction. BP neural network is a multilayer feed-forward neural network. It is trained by error back-propagation. In addition, the characteristics of the BP neural network are that the model is clear and the structure is simple. The basic ideas for training the BP neural network model are as follows [7]:

Herein, training sample is defined as (x_1, y_1), (x_2, y_2), ..., (x_n, y_n), where x represents the input vector, y represents the output vector, and n represents the number of training samples. Supposing that layer $l-1$ and layer l contain m and k nodes, respectively, the output a_j^l of the j th node of layer l can be expressed as formula (1), and the vector a_j^l composed of the output of layer l can be expressed as formula (2).

$$a_j^i = f\left(z_j^l\right) = f \sum_{i=1}^{m} w_{ji}^l a_i^{l-1} + b_j^l, \tag{1}$$

$$a^l = f\left(z^l\right) = f\left(W^l a^{l-1} + b^l\right). \tag{2}$$

Among them, w_{ji}^l and b_j^l are the weight and bias of node j from $l-1$ to l layer, respectively; $f(z)$ is the activation function; w^l represents $k \times m$ matrix composed of the lth layer weight; b^l represents $k \times 1$ vector composed of the bias outputs of the l layer.

As can be seen from the above formula that the output vector of each layer in the network can be calculated. Then, the loss function is used to find the appropriate weight matrix and bias vector corresponding to all hidden layers and output layers. The gradient descent method is adopted to continuously update the weight matrix and bias vector. Thus, the best weight matrix and bias vector are obtained. Finally, the optimal weight matrix and bias vector are utilized to predict the predicted value which is closest to the actual value. According to Pan et al. [8], the mean square error function is selected as the loss function. The expressions are as follows:

$$\text{MSE} = \left(\frac{\sum_{i=1}^{N} \left(d_i - o_i\right)^2}{N}\right)^{1/2},$$

$$R = \left(1 - \frac{\sum_{i=1}^{N} \left(d_i - o_i\right)^2}{\sum_{i=1}^{N} o_i^2}\right)^{1/2}, \tag{3}$$

where d_i and o_i are the target output and actual output vector of the network respectively, and N is the number of test samples.

The structure of the BP neural network is shown in Figure 1 [9–11]. It includes the input layer, the hidden layer, and the output layer. Neurons at all levels of the network are connected by the way of a full connection layer. The hidden layer is in the middle. It includes multiple layers, but there is no connection relationship between neurons at the same level.

2.2. BP Neural Network Structure. In BP neural network, the number of neurons per layer is associated with the actual number of input and output data. In the hidden layer, too large or too small neurons may prolong the training time. It cannot be fitted well. Therefore, continuous search and testing are required to determine the optimal neurons, and available formulas are expressed as

$$m = \sqrt{n+1} + a.$$
$$m = \sqrt{nl}, \tag{4}$$
$$m = \log_z n,$$

Here, m, n, and l are the neurons in the hidden layer, input layer, and output layer, respectively; a is a constant, which ranges from 1 to 10.
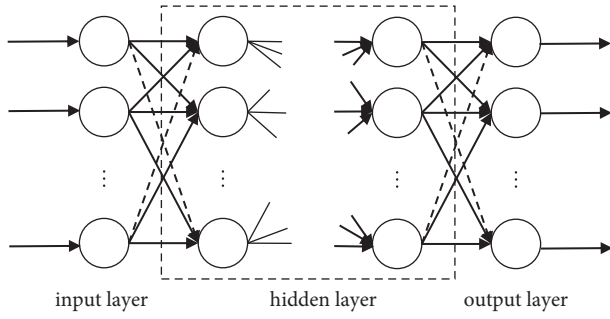
Figure 1: Structure of BP neural network.

The input vector and output vector in the BP neural network should be normalized to the range −1.0~1.0 so as to effectively avoid the overfitting of output vectors [12–14]. The calculation formulas are as follows:

$$x_{mid} = \frac{x_{max} + x_{mix}}{2}.$$
$$\overline{x}_i = \frac{x_i - x_{mid}}{2(x_{max} - x_{min})}, \qquad (5)$$

The abovementioned formulas show that x_i is the data to be normalized; x_{min} is the minimum value; x_{max} is the maximum value; x_{mid} represents the average value; and $\overline{x}_i$ is the normalized data.

3. Construction of Sound Conversion Model Based on BP Neural Network

3.1. BP Neural Network Sound Conversion Model Based on MGC Parameters.
On the basis of the BP neural network, sound characteristic parameters are extracted from the original sound signal through the world sound analysis synthesis system, including one-dimension $lf0$, 60-dimensional MGC parameters, and 5-dimensional bap parameters. Considering that the above three parameters are obtained through the world sound synthesis system and $lf0$ is consistent with the numerical changes of bap parameters, only the 60-dimensional spectrum envelope MGC parameters are performed in modeling training. The BP neural network sound conversion model based on MGC parameters is constructed in this paper, as shown in Figure 2. The neuron number of input and output is 60, the hidden layers are 2, and the neurons of hidden layers is 59 [15–18]. Parameters select the target sound signal value and then place the generated parameters into the WORLD system to synthesize a new sound signal.

Among them, a set of sound signals has 60-dimensional MGC parameters, and one dimension has 117 data, which is represented as $60 * 117$. Here, 45 groups of portable and artificial head data are selected for training, and 5 groups of data are randomly selected as test sets. As the large amount of training data, the parameters are divided into input vector and output vector. For the input vector, there are 45 $60 * 117$ MGC parameters collected by the portable. For output vector, there are 45 $60 * 117$ MGC parameters collected by the artificial head, and the amount of data is large.

Therefore, all of the neurons are 60. In addition, the learning rate of the model algorithm is 0.01; the number of maximum errors is 10; the number of maximum training is 10000, and the accuracy of learning and training is 0.001.

After the transformation model is established, new MGC parameters can be obtained for synthesizing new sounds so as to provide sound parameters for the establishment of subsequent sound evaluation model.

3.2. Sound Signal Denoising Based on Wavelet Decomposition.
In the extraction and synthesis of sound features, the algorithm calculation leads to the sound spectrum decline, frequency offset, resulting in a large amount of noise. These noises will seriously interfere with the quality and evaluation effect of sound signals. Therefore, it is necessary to denoise the transformed and synthesized sound. In this paper, the most widely used wavelet denoising method is adopted. The flow of this method is as follows:

(1) Decompose the original sound signal $s(n)$ by wavelet. So, the real target sound signal $a(n)$ and noise signal $d(n)$ can be separated;

(2) Calculate the wavelet coefficients of real sound signals, and remove the wavelet coefficients of noise signals;

(3) Adopt the inverse transformation method to obtain the new sound signal after wavelet transform so as to complete the denoising of the sound signal.

The formula of an original sound signal with noise signal is

$$s(n) = a(n) + d(n). \qquad (6)$$

The formula (6) shows that $s(n)$ is the original sound signal with noise; $a(n)$ is the normal target sound signal; $d(n)$ is the unwanted noise signal. Furthermore, $d(n)$ is subjected to $N(0, \sigma^2)$ distribution, while $a(n)$ is the non-stationary signal subjecting to non-Gaussian distribution.

4. Experimental Results and Analysis

After the sound conversion and synthesis, the sound evaluation model based on the field of sound conversion and the automobile sound quality will be established so as to evaluate the new synthetic sound signal better. Then, the subjective evaluation was performed for the field of sound conversion, and the objective parameters of the field of automotive sound quality were used to evaluate the synthetic sound signal.

4.1. Evaluation in the Field of Sound Conversion

4.1.1. Experimental Environment and Data.
To achieve better experimental results, this paper selects a meeting room with a strong sound insulation effect in a university for testing, with an indoor temperature of 22°C and the humidity of 45%. Also, the better sound signal can be obtained by using a high-fidelity headphone for sound playback Sennheiser HD650.
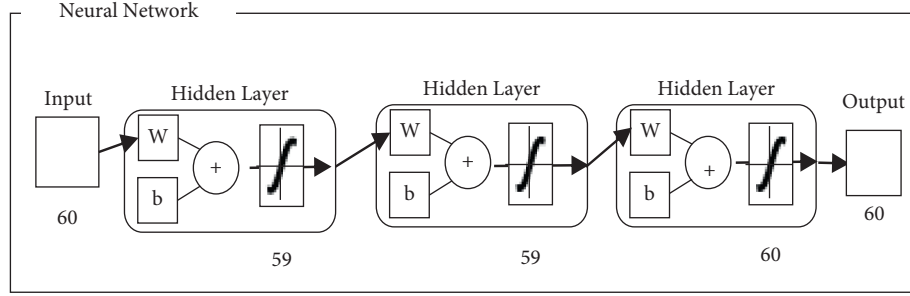
FIGURE 2: The structure of the BP neural network sound conversion model based on the MGC parameters.

The experimental data were derived from 5 healthy and normal hearing postgraduates in a university, including 3 boys and 2 girls.

4.1.2. Test Method. The ABX test and MOS test are adopted to evaluate synthetic sound signals so as to ensure the accuracy of subjective evaluation in the field of sound conversion.

The ABX test is a common method for subjective evaluation of sound conversion. A and B represent the original and target sound signals, respectively, and *X* represents the converted sound signal. The testing process is to distinguish the similarity between the converted sound signal with the original or target sound signal by different people's subjective auditory feelings. Finally, the probability statistics are used to obtain the ABX score of the system so as to evaluate the conversion performance effect of the system.

ABX calculation process is expressed as [19–21]:

$$\text{ABX} = \frac{\sum_{m=1}^{M} p_m}{M} \times 100\%, \tag{7}$$

where M represents the number of testers and p_m represents the test result. When p_m value is 0, it indicates that the converted sound signal is more similar to the original sound signal. When the p_m value is 1, it indicates that the converted sound signal is more similar to the target sound signal.

The MOS test is also known as average opinion score test. Its main function is to evaluate the tester's overall satisfaction with the converted sound signal. If the MOS test score is high, it indicates that the converted sound signal is up to the standard, and its naturalness and intelligibility are better. The expression is as follows:

$$\text{MOS} = \frac{1}{M} \sum_{n=1}^{M} \text{score}_m. \tag{8}$$

Here, M represents the total participants in evaluation, N represents the sound signals in the test, and $\text{score}_{n,m}$ represents the evaluation score of the m-th individual on the nth sound signal.

When conducting subjective evaluation experiments, it is necessary to be in a room with a better environment. The temperature of the selected automobile is 22°C, and the humidity is about 45%. In addition, the sound playback selects high-fidelity headphones to obtain a better sound signal. The sound signals evaluated by the test are mainly four sets of signals obtained through the world sound synthesis system, and the four sets of sound signals correspond to the operating speed of an automobile at 60 km/h, 100 km/h, 30 km/h, and 80 km/h. The four groups correspond to the numbers 1 to 4 in Tables 1 and 2.

The specific evaluation effect is as follows:

(1) ABX test

The results of ABX are shown in Table 1.

According to the above table above, the highest ABX score of the fourth group, reaching 85%, indicates the best effect of this experiment, the lowest ABX score, only 63%; the main reason may be due to too little experimental data. Therefore, the next step is to further expand the amount of data properly. However, comprehensive analysis found that the ABX test score averaged 74%, reaching the experimental standards, with preliminary proof that the newly synthesized sound signal meets the experimental requirements.

(2) MOS test

MOS test results are shown in Table 2.

According to the table, of the four sound signals, the first MOS test score was 4.2, indicating that the experiment was good; the first MOS test score compared with the other three, the lowest score reached only 3, the main reason may be still that the number of data sets is relatively small. The comprehensive analysis shows that the MOS test score was averaged into 3.65, which met the experimental requirements, and preliminarily proved that the newly synthesized sound signal met the experimental requirements. The scoring standards of the MOS test are shown in Table 3.

4.2. Evaluation in the Field of Automobile Sound Quality. First of all, the synthetic sound signal is evaluated by using the objective parameters such as SPL, roughness, loudness, and sharpness [22–24].

Furthermore, the error evaluation is performed between the five groups of synthetic sound signals with the corresponding five groups of portable and five groups of artificial head sound signals. Also, the objective parameters are calculated by LMS Test.Lab.

TABLE 1: ABX test scores.

No.	1	2	3	4
1	0	1	1	1
2	1	0	0	1
3	0	1	1	1
. . .	. . .	. . .	. . .	. . .
30	1	0	1	1
Total	63%	74%	79%	85%

TABLE 2: MOS test scores.

No	1	2	3	4
1	2	3	2	3
2	4	3	4	5
3	1	2	5	4
. . .	. . .	. . .	. . .	. . .
30	5	4	3	5
Total	3	3.4	3.8	4.4

TABLE 3: Scoring standards of MOS test.

Score	Quality evaluation	Distortion degree
5	Excellent	Basically impossible to detect distortion
4	Good	Slightly detect distortion without discomfort
3	Medium	Possible to detect distortion with discomfort
2	Poor	Obnoxious but tolerable
1	Inferior	Obnoxious and intolerable

4.2.1. Roughness. Roughness is a parameter that shows the modulation degree of the sound signal. The unit of roughness value is asper. When the roughness is 1 asper, the sound signal is a sinusoidal pure tone signal with a SPL of 60 dB. Also, beyond that, the frequency is 1 kHz, the modulation amplitude is 1, and the modulation frequency is 70 Hz. It can be seen that the roughness calculation formula is [25]

$$R = 0 \cdot 3 f_{\mathrm{mod}} \int_0^{24\mathrm{bank}} \Delta L_E(z) \mathrm{d}z \mathrm{asper}. \quad (9)$$

Formula (9) shows that f_{mod} represents the modulation frequency, and the unit is kHz. Here, G represents the change value of the excitation stage of a sound signal, expressed as

$$\Delta L_E = 20 \log_{10} \left(\frac{N_{\max}(z)}{N_{\min}(z)} \right), \quad (10)$$

where z represents the critical band Bark number and $N_{\max}$ and $N_{\min}$ represent the maximum and minimum values of the specified loudness in the feature frequency band for a sound signal, respectively.

As can be seen from the above Figure 3, B, R, and X represent portable, artificial head, and newly synthesized sound signals, respectively. Here, the newly synthesized sound signals were close to the artificial head signals. Using
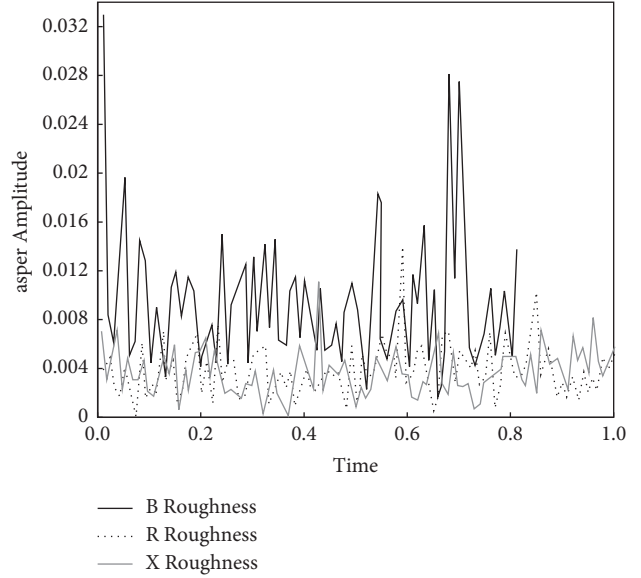

FIGURE 3: Roughness comparison.

LMS Test.Lab to calculate the RMS value, which is found that the portable, artificial head and newly synthesized RMS values are 0.00995 asper, 0.00530 asper, and 0.00494 asper, respectively. The difference between portable and artificial head is 87.74%, while the difference between new synthesized and artificial head is 6.79%. Therefore, the roughness of the newly synthesized sound signal is obviously improved and meets the experimental requirements.

4.2.2. Loudness. Loudness is a parameter proposed to show how the human ear feels about the strength of the sound signal. The magnitude of loudness is determined by the size of the original sound amplitude, which is also associated with the frequency size. The unit of loudness value is expressed as song (sone). When the loudness value of the sound signal is 1 sone, which means it is a pure sound signal with a SPL of 40 dB, and the frequency size is 1 kHz.

The Zwicker algorithm is commonly used to calculate the loudness, which is expressed as

$$N = 0 \cdot 08 \left(\frac{E_{TQ}}{E_O} \right) \left[\left(0 \cdot 5 + 0 \cdot 5 \frac{E}{ETQ} \right)^{0.23} \right] - 1. \quad (11)$$

Formula (13) shows that E_{TQ} represents the excitation generated by the listening valve when it is relatively quiet, E_O represents the corresponding excitation under the reference sound intensity $I_0 = 10^{-12}$ W/m^2, and E represents the corresponding excitation of the sound signal calculated by test. At the characteristic frequency band Bark from 0 to 24, performing the integral operation for the feature loudness to obtain the value of the total loudness, and the expression is as follows:

$$N = \int_0^{24_{\mathrm{bark}}} N'(z) \mathrm{d}z. \quad (12)$$

Here, N is the total loudness, namely the calculation model of steady-state sound signal loudness.

As can be seen from Figure 4, B, R, and X represent portable, artificial head, and newly synthesized sound signals, respectively. Here, the newly synthesized sound signals are between the portable signals and the artificial head signals. Also, using LMS Test.Lab to calculate the RMS value finds that the portable, artificial head and newly synthesized RMS values are 0.63 sone, 1.24 sone, and 0.95 sone, respectively. In addition, the error of portable and artificial head is 49.19%, and the error of new synthesized and the artificial head is 23.39%, indicating that the loudness of the new synthesized sound signal is significantly improved and reaches the experimental standard.

4.2.3. Sound Pressure Level (SPL).

The sound pressure level is one of the important methods to express subjective feelings of loudness. In ANSI S1.8 (1989) and ANSI S1.13 (1995), the sound pressure level is calculated as

$$ \text{SPL} = 20 \, \log_{10}\left(\frac{p_e}{p_{\text{ref}}}\right)\text{d}B. \tag{13} $$

The formula (13) shows that in SPL, p_e represents the effective sound pressure value for testing sound signals, and p_{ref} represents the sound pressure size, that can be heard by human ears in a stable condition, whose sound signal is 1 KHZ. In addition, this sound pressure value is also the audible threshold value. The reference sound pressure refers to the minimum root mean square sound pressure. In other words, at standard atmospheric pressure, and when $p_{\text{ref}} = 2 \times 10^{-5}$ pa, the subjective perception of sound intensity by human ear is not very closely related to the sound pressure level itself. If the loudness increases, the sound pressure level will increase in a logarithmic speed. The expression of Stevens extracting the power relation of both is

$$ L = kI^{0.3}, \tag{14} $$

where L represents the perceived loudness of the ear, k represents the coefficient of each subject, and I represents sound intensity.

As can be seen from the comparison diagram, B, R, and X also represent portable, artificial head, and newly synthesized sound signals, respectively. It can be found from Figure 5 that the new synthesis is between the portable signals and the artificial head signals. The RMS values of the three are 47.86 Pa, 45.32 Pa, and 44.98 Pa, respectively. The error of portable and artificial head is 5.60%, and the error of new synthesis and human foreman is 0.75%. It indicates that the A sound level of the newly synthesized sound signal is further improved to meet the requirements.

4.2.4. Sharpness.

Sharpness is an objective parameter. It emphasizes the sharpness of sound signal and mainly shows the proportion of high-frequency signal in the total sound spectrum. In a sound signal, the larger the proportion of high-frequency components, the greater the loudness value, and the greater the sharpness of this sound signal.

The unit of sharpness value is acum, and the sound signals with a sharpness of 1 acum are a narrow band noise signal with
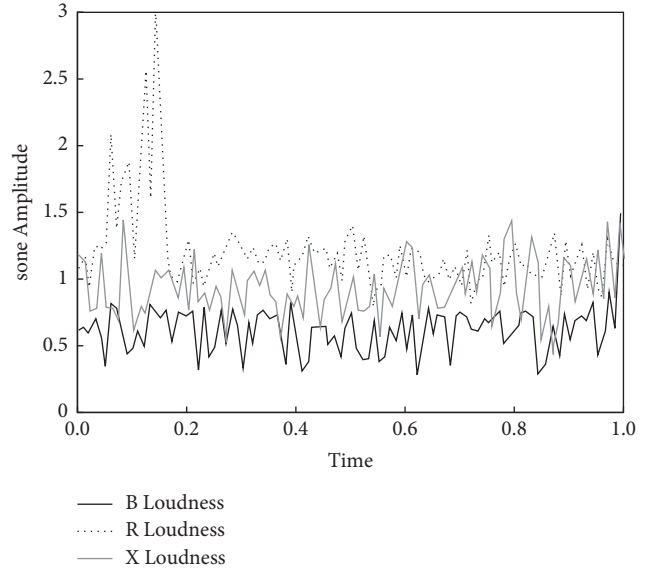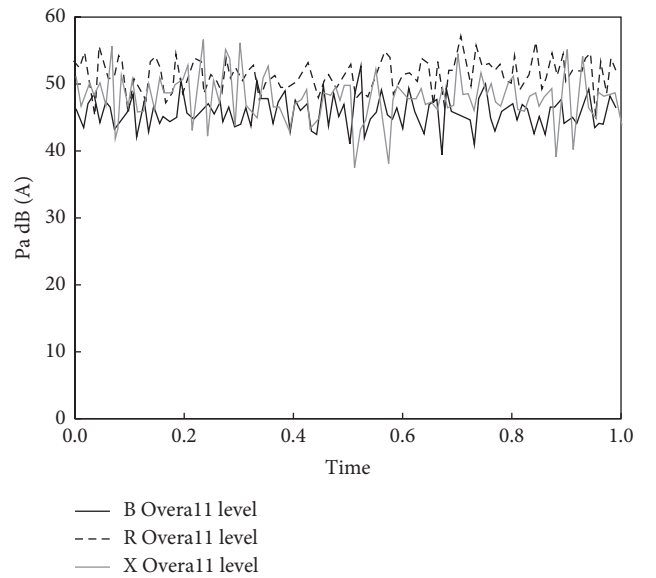

Figure 4: Loudness comparison.

Figure 5: Sound pressure level comparison.

a sound pressure level 60 dB. The center frequency size is 1 kHz, and the bandwidth size is 160 Hz. The solution method of sharpness is to conduct the weighted integral calculation for the total loudness and the spectrum response of the critical frequency band, and the expression formula is

$$ S = 0 \cdot 11 \frac{\int_0^{24\text{bark}} N'(Z) \cdot Z \cdot g(Z)\text{d}Z}{\int_0^{24\text{bark}} N'(z)\text{d}z} \text{acum}, $$

$$ g(z) = \begin{cases} 1, & z \leq 16, \\ 0.06e^{017z}, & 16 < z \leq 24. \end{cases} \tag{15} $$

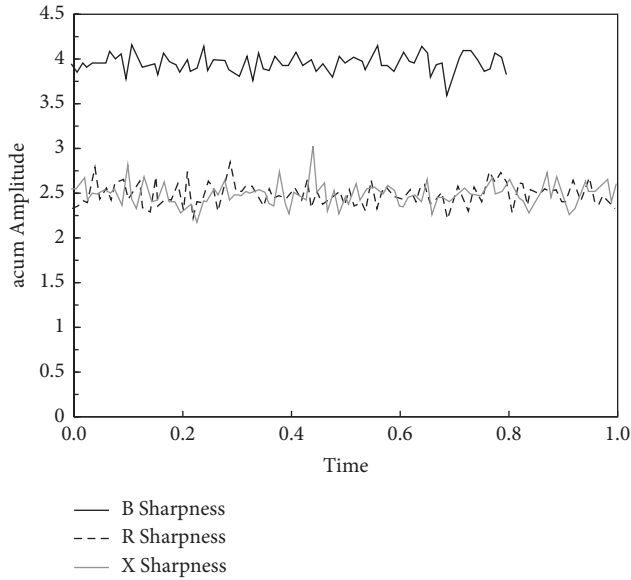Among them, z represents the critical frequency bands number; $N'(z)$ represents the characteristic loudness

Figure 6: Sharpness comparison.

Table 4: Grade 10 evaluation scale.

Adjective	Review score
Remarkable	10
Great	9
Very good	8
Satisfied	7
Acceptable	6
Need to improve	5
Offensive	4
Very disgusting	3
Disturbing	2
Intolerable	1

Table 5: Evaluation results of synthetic sound.

Evaluating indicator	Average score
Roughness	7
Loudness	8
Sound pressure level	9
Sharpness	8
Overall average score	8

At the same time, the test is conducted in a closed environment to avoid interference from the external environment. The number of tests is 10, and the final results are averaged. Based on the abovementioned tests, the obtained results are shown in Table 5.

As can be seen from the above scores, the overall average score of the four indicators is 8, which belongs to the great level. It shows that the synthesized sound can be accepted by the evaluator and get a better evaluation.

5. Conclusion

To sum up, the sound quality evaluation model based on BP neural network constructed in this paper is feasible. The model can effectively evaluate the sound signals of new synthetic, portable, and artificial heads. Also, after testing through subjective evaluation and objective parameters, it is found that the newly synthesized sound signal is very close to the target sound signal, with an average error of less than 10%. As can be seen that in the field of sound quality evaluation, the newly synthesized sound is up to the design standards of this paper. However, due to the limitation of conditions, there are still some shortcomings. The main reason is that there are too few data sets elected in the conversion model, which results in that the final test data may have a low monitoring score. It provides a more comprehensive evaluation method for the research and synthesis of automobile sound.

Data Availability

The experimental data are available from the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

References

[1] X. Xie, Z. Ma, J. Ye, F. Zeng, W. Fan, and B. Chen, "Research and development of sound quality in portable testing and evaluation system based on self-adaptive neural network," *Applied Acoustics*, vol. 154, pp. 138–147, 2019.

[2] S. Yin, Z. Gu, Y. Zong, L. Zheng, Z. Yang, and T. Huang, "Sound quality evaluation of automobile side-window buffeting noise based on large-eddy simulation," *Journal of Low Frequency Noise, Vibration and Active Control*, vol. 38, no. 2, pp. 207–223, 2019.

[3] Z. Yang, H. Feng, and S. Lu, "Sound quality evaluation of automobile interior noise under transient and steady-state

function and $g(z)$ represents the weighting function when the critical frequency band of sound is high. Its numerical change is mainly affected by the critical frequency band.

As can be seen from the sharpness comparison in Figure 6, the newly synthesized sound signal is adjacent to the artificial head signal, and the RMS value can be calculated by LMS Test.Lab. The portable, artificial head and newly synthesized RMS values are 6.54 acum, 2.47 acum, and 2.47 acum, respectively. The error of portable and artificial head is 43.32%, while the error of new synthesized and artificial head is 0.00%, indicating that the sharpness of the new synthesized sound signal is further improved and reaches the experimental standard.

The above evaluation is based on a single index to evaluate the sound inside the automobile, which is obviously not objective enough. On the basis of single index evaluation, a comprehensive evaluation method is proposed to evaluate the new synthetic automobile sound. As the units of above indicators are different and belong to different levels, the following ideas are adopted in the comprehensive evaluation: Firstly, 32 different evaluators are invited to evaluate the new synthetic interior sound signals according to the grade standards in Table 4 and scored them one by one so as to obtain the overall evaluation of new synthetic sound.

running conditions," *Journal of the Acoustical Society of America*, vol. 145, no. 3, p. 1882, 2019.

[4] J. H. Park and Y. J. Kang, "Evaluation index for sporty engine sound reflecting evaluators' tastes, developed using K-means cluster Analysis," *International Journal of Automotive Technology*, vol. 21, no. 6, pp. 1379–1389, 2020.

[5] B. Zhao and C. J. Wu, "Sound quality evaluation of electronic expansion valve using Gaussian restricted Boltzmann machines based DBN," *Applied Acoustics*, vol. 170, Article ID 107493, 2020.

[6] Z. Wang, P. Li, H. Liu, J. Yang, S. Liu, and L. Xue, "Objective sound quality evaluation for the vehicle interior noise based on responses of the basilar membrane in the human ear," *Applied Acoustics*, vol. 172, Article ID 107619, 2021.

[7] S. D. Loreto, F. Serpilli, V. Lori, and S. Squartini, "Sound quality evaluation of kitchen hoods," *Applied Acoustics*, vol. 168, Article ID 107415, 2020.

[8] J. Pan, X. Cao, D. Wang, and J. Chen, "Vehicle interior sound quality evaluation index selection scheme based on grey relational analysis," *Fluctuation and Noise Letters*, vol. 19, no. 3, Article ID 2050031, 2020.

[9] T. Feng, Y. Sun, Y. Wang, P. Zhou, H. Guo, and N. Liu, "Sound feature space effects on the performance of annoyance evaluation model based on support vector machine," *Applied Acoustics*, vol. 154, pp. 99–113, 2019.

[10] Y. Lu, Y. Zuo, H. Wang, and C. Wu, "Sound quality prediction for power coupling mechanism of HEV based on CEEMD-HT and RVM," *Neural Computing & Applications*, vol. 33, no. prepublish, pp. 1–16, 2020.

[11] S. Xu, Z. Lin, G. Zhang, T. Liu, and X. Yang, "A fast yet reliable noise level estimation algorithm using shallow CNN-based noise separator and BP network," *Signal, Image and Video Processing*, vol. 14, no. prepublish, pp. 1–8, 2019.

[12] M. Chen, B. Peng, S. Huang, and P. K. L. Chan, "Understanding the meniscus-guided coating parameters in organic field-effect-transistor fabrications," *Advanced Functional Materials*, vol. 30, no. 1, 2020.

[13] Image and Video Processing, "New findings in image and video processing described from School of Information Engineering (A fast yet reliable noise level estimation algorithm using shallow CNN-based noise separator and BP network)," *Journal of Technology*, vol. 14, 2020.

[14] S. Xiaojun and L. Yalin, "Research on financial early warning of mining listed companies based on BP neural network model," *Resources Policy*, vol. 73, Article ID 102223, 2021.

[15] Y. Jiang, Y. Liu, D. Yu, and F. Li, "Defect signal detection of station process pipelines based on the BP neural network," *IOP Conference Series: Earth and Environmental Science*, vol. 804, no. 4, 2021.

[16] Y. X. Xie, Y. J. Yan, X. Li, T. S. Ding, and C. Ma, "Fault diagnosis method for scintillation detector based on BP neural network," *Journal of Instrumentation*, vol. 16, no. 07, 2021.

[17] S. Skoda, A. Fiebig, F. B. Schulte, and J. B. Schweitzer, "How does the focus of attention influence the sound quality evaluation of a complex auditory scene?" *INTER - NOISE and NOISE - CON Congress and Conference Proceedings*, vol. 255, no. 2, p. 5457, 2017.

[18] C. Ma, C. Chen, Q. Liu et al., "Sound quality evaluation of the interior noise of pure electric vehicle based on neural network model," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 12, pp. 9442–9450, 2017.

[19] M. Moravec, G. Ižaríková, P. Liptai, M. Badida, and A. Badidová, "Development of psychoacoustic model based on the correlation of the subjective and objective sound quality assessment of automatic washing machines," *Applied Acoustics*, vol. 140, pp. 178–182, 2018.

[20] T. Masayuki, K. Iida, T. Shoki, and S. I. Iwamiya, "Sound quality evaluation of noise emitted from brush cutters," *INTER - NOISE and NOISE - CON Congress and Conference Proceedings*, vol. 258, no. 6, pp. 1238–1248, 2004.

[21] J. Zhang, S. Xia, S. Ye et al., "Sound quality evaluation and prediction for the emitted noise of axial piston pumps," *Applied Acoustics*, vol. 145, pp. 27–40, 2019.

[22] Y. Date, N. Yamagiwa, Y. Morimoto, N. Tanimoto, and S. Ishimitsu, "A study on sound quality evaluation method for vehicle using vital information," in *Proceedings of the INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, pp. 3369–3378, Madrid, Spain, September 2019.

[23] W. Sung, P B. Davies, and J. Stuart, "Sound quality evaluation of residential HVAC&R equipment," *INTER - NOISE and NOISE - CON Congress and Conference Proceedings*, vol. 258, no. 6, pp. 1066–1074, 2018.

[24] A. Pereda, Y. Miguel, and E. Stumpf, "Parametric aircraft configuration optimization according to ICAO annex 16 certification standards and sound quality evaluation within conceptual aircraft design," *INTER - NOISE and NOISE-CON Congress and Conference Proceedings*, vol. 258, no. 7, pp. 307–316, 2018.

[25] H. B. Huang, R. X. Li, M. L. Yang, T. C. Lim, and W. P. Ding, "Evaluation of vehicle interior sound quality using a continuous restricted Boltzmann machine-based DBN," *Mechanical Systems and Signal Processing*, vol. 84, pp. 245–267, 2017.