

<https://doi.org/10.70917/ijcisim-2026-0140>
Article

Optimization Effect of Nonlinear Denoising Algorithm of Audio Signal in Piano Performance on Aesthetic Performance of Piano Timbre

Xiting Yang *

College of Music and Dance, Huaqiao University, Xiamen, Fujian, 361021, China; xiting_music@163.com

Abstract: In recent years, with the rapid development of computer science and information technology, research on the acoustic aspects of audio signals has gradually become a hot topic in the field of music. Addressing the issue of low accuracy in current instrument audio recognition, this paper proposes a set of nonlinear denoising algorithms to optimize the aesthetic performance of piano tones. Using WAV files of Yamaha YA128EX piano performances as experimental data, MATLAB simulation results show that within 10s, 20s, and 30s, the root mean square errors (RMSE) of the signal processing using this method are 4.49%, 3.30%, and 3.36%, respectively, while the RMSE of other methods are 5.80%, 6.88%, and 3.76%, respectively. Therefore, the proposed method achieves more significant results in piano audio signal feature extraction, with higher extraction accuracy, and the beam integration of the recognition output is significantly improved.

Keywords: nonlinear denoising; signal features; MATLAB simulation; piano performance

1. Introduction

The piano is an instrument with exceptional tonal aesthetic expressiveness. Compared to other instruments, the piano has an extremely wide range and a rich variety of tones, which can be manipulated through different playing techniques to create diverse tonal effects [1-4]. For example, when playing a gentle melody, the performer can use a soft playing style and vary the force applied to the keys to convey the emotional nuances of the melody [5-6]. When strong emotional expression is required, the performer can use stronger playing force and a more resolute left-hand rhythm to demonstrate their musical expressiveness [7-9].

However, in traditional piano performance, the expression of piano tonal aesthetics is easily influenced by factors such as the piano's materials, the performer's musical understanding, and their emotional state, leading to suboptimal expressiveness [10-12]. Therefore, optimizing the aesthetic expression of piano tone color is crucial for enhancing the overall performance effect. In this process, nonlinear denoising algorithms for audio signals can play a significant role [13-15]. Nonlinear denoising algorithms for audio signals include variational nonlinear linear frequency modulation mode decomposition, Savitzky-Golay filtering algorithm (SG filtering algorithm), and spectral subtraction, primarily applied in audio recording, audio transmission, audio coding and decoding, and audio enhancement. In the optimization of piano tone aesthetics, these algorithms play an important role in enhancing the aesthetic quality of piano tones [16-19].

Reference [20] describes the implementation of a reliable and reproducible octave tuning method based on spectral entropy minimization and conducts a subjective evaluation of the timbre of tuned pianos. Reference [21] emphasizes the importance of emotional integration in piano performance and proposes a deep learning-based piano music emotion extraction algorithm. Through comparative experiments, it is shown that the algorithm achieves an accuracy rate of approximately 94% on the test set and over 95% on the experimental set. Literature [22] introduces the important role of body



movements in piano performance, which not only produce expected sounds but also achieve purposes such as expressing emotional states or coordinating with collaborators, and discusses the influence of movement characteristics such as finger speed on piano music performance. Literature [23] elaborates on the role of literary cultivation in piano performance, which enhances the deep understanding of musical works and enables the performance of moving melodies through reasonable playing methods, thereby achieving the purpose of refining one's character. Literature [24] emphasizes that artistic expression in piano performance largely depends on the precise control of touch intensity, with the accuracy of intensity determining the quality of sound production and directly shaping the completeness and depth of musical expression. Literature [25] explores piano track programming and creative optimization methods based on big data technology. By discussing the potential applications of big data technology in the field of music composition, it proposes an innovative framework for piano repertoire arrangement and composition optimization, which can mine massive piano repertoire data to extract musical elements and composition patterns.

Addressing the issue of noise significantly affecting the accuracy of sound signal recognition, this paper utilizes a nonlinear denoising algorithm to optimize the aesthetic performance of piano tones. Prior to denoising the piano audio signal, the audio signal undergoes preprocessing through channel conversion, normalization, pre-emphasis, resampling, windowing, and mean value removal. Subsequently, EMD decomposition is applied to the original data to obtain the intrinsic mode functions (IMFs). Components with higher correlation coefficients are retained to achieve the goal of removing noise from the piano audio signal. The experimental data in this paper are sourced from the Yamaha YA128EX piano sound WAV file. The proposed algorithm will be tested and analyzed in a MATLAB simulation environment to validate its effectiveness in optimizing the aesthetic performance of piano tone.

2. Piano audio signal processing based on nonlinear denoising algorithms

In the process of audio signal classification and recognition, there are three key steps: noise reduction of audio signals, extraction of audio signal features, and design of classifiers. Generally, audio signal recognition systems need to classify audio signals with variable signal-to-noise ratios and complex, diverse noise. However, noise has a significant impact on the accuracy of audio signal recognition. Therefore, the key to improving the recognition accuracy of audio signal recognition systems is to extract effective audio signal features from noisy audio signals.

2.1. Audio Signal Preprocessing

Noise reduction is part of audio signal preprocessing. Before performing noise reduction on instrument audio signals, the following six preprocessing steps must be performed on the instrument audio signals:

(1) Channel conversion. Since the collected instrument audio signals are stereo, they consist of two separate audio signals: the left channel and the right channel. Therefore, if the speech features we wish to extract do not distinguish between left and right channels, the multi-channel instrument audio signals must be converted to mono. This can be achieved by summing the two channels and then taking the average to convert the multi-channel instrument audio signals into mono audio signals.

(2) Normalization. This step is to eliminate the influence of different volumes on the acoustic features that characterize timbre, so the amplitude is normalized.

(3) Pre-emphasis. The purpose of pre-emphasis is to retain only signals within a certain frequency range. In fact, this process acts as a high-pass filter, i.e., the frequency domain of the normalized original audio signal is convolved with the high-pass filter.

(4) Resampling. In practical applications, the audio signals we encounter may come from different devices, and the parameters set during recording are not the same. The most important parameter is the sampling rate. According to the Nyquist sampling theorem, the maximum frequency of the signal must be less than or equal to half the sampling rate to ensure that the sampled data can be used to reconstruct the signal. This paper adopts the default sampling rate of 44.1 kHz from the Cool Edit Pro audio editing software.

(5) Framing windowing. Based on the short-term stationary characteristics of speech, instrument audio signals can be processed in frames. This paper uses a Hamming window to apply windowing to instrument audio signals to reduce the impact of the Gibbs effect. When dividing into frames, this paper sets one frame to 20 ms and the frame shift to 10 ms.

(6) Mean subtraction, with the purpose of removing the DC component from the signal [26].

2.2. Eigenmode Functions and Empirical Mode Decomposition

2.2.1. Eigenmode Function

From a physical perspective, the meaningfulness of instantaneous frequency is contingent upon the function being symmetric, having a local mean of zero, and possessing an equal number of extrema and zero-crossing points. Based on this, the concept of the intrinsic mode function (IMF) was introduced, where the instantaneous frequency is meaningful at all points on the IMF. The intrinsic mode function is also referred to as the eigenmode function. Any signal is composed of several IMFs, and composite signals are formed by the overlap of multiple IMFs. From the concept of IMFs, it is known that within each vibration period of an IMF, only one vibration mode is allowed, and no other complex harmonics are permitted. An IMF is not restricted to being a narrowband signal; it can also be non-stationary or modulated in frequency and amplitude. Signals with only amplitude modulation or frequency modulation can also be IMFs [27].

2.2.2. Empirical Mode Decomposition

The Empirical Mode Decomposition (EMD) algorithm was proposed based on the Hilbert-Huang transform. By performing EMD decomposition on the original data, the intrinsic mode functions can be obtained. EMD achieves noise removal by retaining IMF components with high correlation coefficients with the original signal [28]. The specific decomposition process of the EMD algorithm is as follows:

(1) Identify all the maximum points in the original data sequence $X(t)$, and use a cubic spline interpolation function to fit the upper envelope of the original data sequence based on all the maximum points. Similarly, identify all the minimum points and use a cubic spline interpolation function to fit the lower envelope of the original data sequence based on all the minimum points.

(2) Let the mean of the upper and lower envelopes of the original data be denoted as $mean(t)$. A new data sequence $c(t)$ is obtained by subtracting $mean(t)$ from the original data sequence $x(t)$, as shown in Formula (1):

$$c(t) = x(t) - mean(t) \quad (1)$$

(3) The $c(t)$ obtained from formula (1) generally does not satisfy the conditions for stopping the selection of intrinsic mode functions, so $c(t)$ is used as the new $x(t)$, and steps (1) and (2) are repeated until the conditions for stopping the selection of IMFs are satisfied. At this point, the first-order difference IMF is obtained, denoted as $c_1(t)$.

(4) Subtract $c_1(t)$ from the original audio signal $x(t)$ to obtain the sequence of residual components $r_1(t)$ as shown in formula (2):

$$r_1(t) = x(t) - c_1(t) \quad (2)$$

(5) Use $r_1(t)$ as the new original audio signal, and repeat all the above steps to obtain the next-order difference IMF component. Continue this filtering process until the residual audio signal $r_n(t)$ at the n th order is smaller than the expected value or becomes a monotonic function. At this point, you obtain several MF components $c_1(t), c_2(t), \dots, c_n(t)$, and the residual term $r_n(t)$. Thus, the original audio signal $x(t)$ is decomposed into a series of IMF components and a residual term $r_n(t)$, as shown in equation (3):

$$x(t) = \sum_{i=1}^n c_i(t) + r_n(t) \quad (3)$$

Although the EMD algorithm is particularly suitable for adaptive denoising of nonlinear, non-stationary signals, when similar feature time scales are distributed across different eigenmode functions or a single eigenmode function contains feature time scales with extremely large differences, adjacent eigenmode function waveforms may overlap, making them difficult to distinguish and causing mutual interference. This phenomenon is referred to as mode overlap. When performing EMD decomposition for noise reduction on signals with abrupt changes in time scales, modal overlap occurs, thereby affecting the effectiveness of noise reduction for the signal.

2.3. Frame Estimation of Piano Audio Signal Features

During signal processing, it is necessary to calculate the short-term energy changes of the signal. When using short-term amplitude to represent signal energy, it is difficult to ensure the accuracy of energy calculations due to the uneven effective length of audio data. Therefore, it is necessary to first divide the audio data into frames, transforming it into uniform audio segments. During the framing process, it is necessary to first define a rectangular window:

$$d(x) = \begin{cases} 1, x \in [0, X-1] \\ 0, x \in (-\infty, 0) \cup (X-1, +\infty) \end{cases} \quad (4)$$

In the formula, $d(x)$ represents the effective amplitude of the audio data subframe; X represents the number of sampling points. At this point, the average amplitude of each frame of audio signal can be expressed as:

$$\begin{cases} G_i = \frac{\sum_{i=1}^n |d_i(x)|}{X} \\ G = \frac{\sum_{i=1}^n G_i}{T_n} \end{cases} \quad (5)$$

In the equation, G represents the average amplitude value of the audio data segment, G_i represents the average amplitude value of the i th sub-frame of audio after segmentation; $d_i(x)$ represents the effective amplitude of the i th sub-frame; T_n represents the number of sub-frames into which the audio data segment is divided. Through effective segmentation, the average amplitude can be used as an endpoint detection method, and the adaptive correlation function of the audio signal can be expressed as:

$$h_i(p) = \sum_{i=1}^n T_i(p) \cdot T_i(u_n + q_n) \quad (6)$$

In the equation, $h_i(p)$ denotes the adaptive correlation function of the i th subframe of the audio signal; $T_i(p)$ denotes the frame length of the i th frame of the audio signal; u_n denotes the period length of the signal; and q_n denotes the number of sample points in the subframe audio. At this point, the relevant coefficients can be directly calculated using the recursive formula. In the cepstrum coefficients, the formula for the linear prediction indicator is:

$$u_n = \begin{cases} f_n + \sum_{k=1}^{n-1} \frac{\varphi_p a_{s+k}}{n}, n \leq p_i \\ f_n + \sum_{k=n-1}^{n-1} \frac{\varphi_p a_{s-k}}{n}, n > p_i \end{cases} \quad (7)$$

In the equation, u_n represents the cepstral coefficient of the linear prediction index of the piano audio signal; a_{s+k} and a_{s-k} represent the feature extraction parameters of the piano audio signal, respectively; φ_p represents the linear prediction coefficient index; and p_i represents the order of the cepstral coefficient. The threshold estimation expression at this point is:

$$\gamma_t = \mu_n \sqrt{2 \log 10(p_k)} \quad (8)$$

In the equation, γ_t represents the indicator amplitude of the linear threshold in the piano audio signal; μ

represents the noise filtering coefficient; p_k represents the signal threshold for white noise. Using

the above formula, the error coefficient during the frame division process can be calculated, and the frame-based estimate of the threshold for the piano audio signal can be obtained.

2.4. Construction of Piano Audio Signal Denoising Algorithm

In the noise model, the noise variance can be obtained using the above threshold estimation method. The calculation formula is as follows:

$$\theta_{pk} = \frac{\sum_{i=1}^n MAD_i}{0.544} \quad (9)$$

In the equation, θ_{pk} represents the variance coefficient of noise in the process of double noise elimination; MAD_i represents the median estimate index of the i th layer coefficient of the signal. In the noise elimination algorithm, the threshold of the signal usually changes. At this time, the calculation formula for the signal threshold can be expressed as:

$$\gamma_t = \frac{\mu_n \sqrt{2 \log 10(p_k \log p_k)}}{\sum_{i=1}^n p_k} \quad (10)$$

The unknown variables in the equation are defined as shown in formula (10). At this point, an initial mean error estimate can be set, and its maximum and minimum thresholds can be calculated. While generating the mean square error, the likelihood estimate can be used as an adaptive selection coefficient to improve the noise reduction effect of the signal-to-noise ratio. The formulas for calculating the maximum and minimum thresholds are as follows:

$$\mathcal{G}_u^{i-1}(d_i) = \begin{cases} 0, & |d_i| \leq u \\ 1, & |d_i| > u \end{cases} \quad (11)$$

In the equation, $\mathcal{G}_u^{i-1}(d_i)$ denotes the contraction coefficient after setting the threshold to zero in the wavelet coefficients; u denotes the global decomposition noise comparison result caused by signal discontinuity during the process of avoiding discontinuity; d_i denotes the i th discontinuous convergent signal curve. During the Fourier transform, since the time-frequency plane of the signal can be equivalently analyzed through the spectral analysis of the kernel function, the rotation angle α of the Fourier transform axis can be calculated. This allows point N to be transformed into the distribution pattern, yielding the coordinate transformation formula:

$$\begin{cases} b_1 = b \cos \alpha - c \sin \alpha \\ c_1 = b \sin \alpha - c \cos \alpha \end{cases} \quad (12)$$

In the equation, the coordinates of point N_1 are (b_1, c_1) at the position of point N after undergoing an angle α transformation; similarly, the position of point N_2 is at the position of point N after undergoing an angle β transformation, with specific coordinates (b_2, c_2) ; after the angle θ transformation, point N appears at the coordinates (b_3, c_3) and is marked as N_3 . Using these transformed points, a denoising model can be established, and the formula can be simplified as:

$$g(i) = d(i) + \varepsilon_n \cdot q(i) \quad (13)$$

In the equation, $g(i)$ represents the signal after pure denoising; $d(i)$ represents the signal containing noise; ε_n represents the intensity of noise in the signal; $q(i)$ represents the threshold of white noise. Through the double denoising method, a denoising model can be obtained and simplified to the formula shown in (10).

3. Experimental Results and Analysis

3.1. Experimental Data

The experimental data is derived from WAV files of the Yamaha YA128EX piano sound. This data was collected using 13 microphones and 18 microphones to capture the timbre and spatial characteristics of the mechanical key signal of the piano's 65th key. The sampling frequency was 45 kHz, and the quantization precision was 16 bits. A total of 10,000 audio spectra of the piano's mechanical key operation were collected, which were divided into a training set and a test set in an 8:2 ratio. The original time-domain data is shown in Figure 1.

In the figure, the audio signals from left to right represent the duration of short strokes, the duration of long strokes, the duration when the sustain pedal is pressed, the intensity of forceful strokes, the intensity of soft strokes, the actual frequency, and the theoretical frequency. To achieve better experimental results, the proposed algorithm will be tested and analyzed in the MATLAB simulation environment, and the frequency will be extracted from the test files.

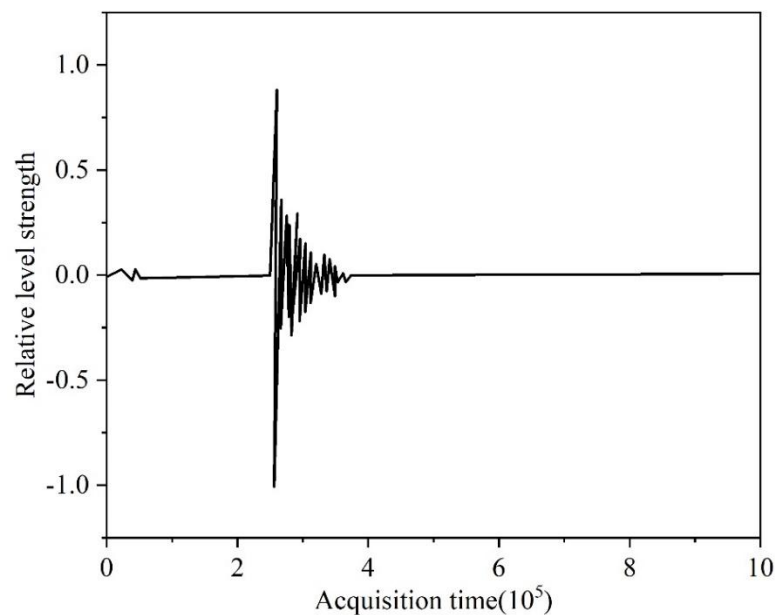


Figure 1. Time domain diagram of signal acquisition

3.2. Selection of the Number of Wavelet Decomposition Layers

When performing wavelet decomposition on different noisy speech signals, there is an optimal number of decomposition levels. Selecting the optimal number of decomposition levels for wavelet denoising yields the best denoising results. Therefore, the selection of the number of decomposition levels significantly impacts the denoising effect. In practical applications, when using wavelet transformation for noise reduction, setting too few decomposition levels may result in no significant improvement in the signal-to-noise ratio after noise reduction, failing to achieve satisfactory noise reduction effects. However, this does not mean that setting more decomposition levels will necessarily lead to better noise reduction results. On the contrary, setting too many decomposition levels may cause severe loss of useful speech signals, significantly reducing the signal-to-noise ratio after noise reduction and greatly increasing computational complexity. Therefore, when using the wavelet threshold method for noise reduction, it is essential to select an appropriate number of decomposition levels to achieve optimal noise reduction results. The noisy speech signals collected in the experiment were decomposed using the `waveDec()` function in MATLAB, and the signal shapes at different decomposition levels were observed. The noisy speech signals are shown in Figure 2.

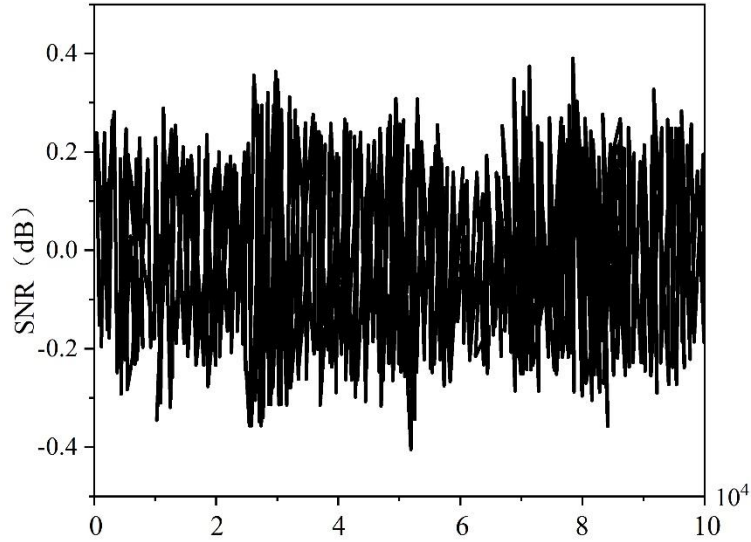


Figure 2. denoting speech signal

After subjecting the noisy speech signal to wavelet decomposition and reconstruction across different numbers of layers, and then applying wavelet denoising processing at various decomposition levels, the resulting signal-to-noise ratio (SNR) after denoising is shown in Table 1. By analyzing and comparing the SNR values in the table, it was found that for the audio data collected in the experiment, performing four-layer wavelet decomposition followed by denoising processing yielded the highest SNR, achieving optimal denoising results. Therefore, in this paper, when performing noise reduction on the noisy speech signals collected in the experiment, the number of wavelet decomposition layers is set to four.

Table 1. Signal signal-to-noise ratio (dB).

Signal	Voice signal	A layer of decomposition signals	Two layers of decomposition signal	Three layer decomposition signal	The four-layer decomposition signal	The five-layer decomposition signal
SNR(dB)	8.0062	8.2415	12.4571	16.3875	20.0479	15.4933

3.3 Verification of Feature Extraction Results

To validate the effectiveness of the proposed signal feature extraction method, the experiment input the mechanical key audio signal of piano No. 65, collected as described above, into the GRNN neural network model for feature extraction. Since the onset point of this signal is close to the 13th input channel, the input from the 13th microphone was used as a reference, and the input from the 18th microphone was subjected to correlation processing, yielding the actual correlation data for the 18th microphone relative to the onset point of the 65th piano key as follows: -0.2278, -0.1540, -0.1979, -0.2745, -0.2847, -0.2276, -0.2984, -0.2280, -0.2098, -0.1893, -0.2111, -0.2096, 1.0004, 0.7125, 0.8676, 0.0667, 0.1241, and 0.1035. Based on the above test sample data, it was input into a nonlinear denoising method for tone and spatial sense feature extraction and analysis, yielding the test results shown in Table 2. As shown in the table, after applying the nonlinear denoising method to extract features from the piano mechanical key audio signal, the predicted timbre result was 78, while the human evaluation result was 84, with an error of 6. The predicted spatial sense result was 66, while the human evaluation result was 70. Comprehensive analysis indicates that the proposed nonlinear denoising method achieves high accuracy in feature extraction for piano mechanical key audio signals, with minimal extraction error, which can provide effective parameters for subsequent signal processing.

Table 2. Piano mechanical key feature extraction results.

Piano brand	Trait group	Predd value	Prediction station fruit	Manual evaluation	Error
Ya128ex	Timbre	0.2	78	84	6
Yaizaex	Sense of space	0.3	66	70	7

3.4 Verification of Signal Processing Results

To validate the performance of the proposed piano mechanical key signal processing method, the experiment will use this method to process the signal of piano mechanical key No. 65, resulting in the signal processing results shown in Figure 3. As can be seen from the results in the figure, after processing the piano mechanical key audio signal using the proposed piano mechanical key signal processing method based on nonlinear denoising feature extraction, the audio signal feature extraction effect is more significant, and the beam integration of the feature recognition output is significantly improved.

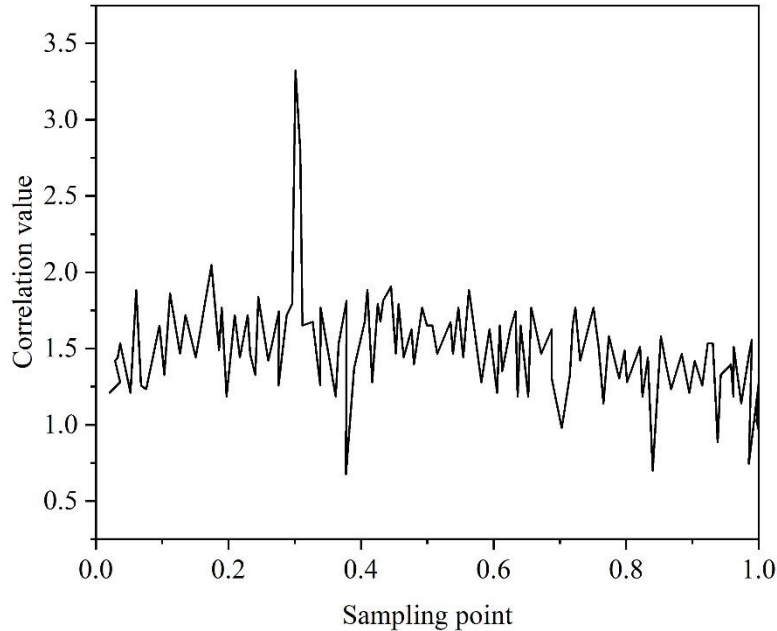


Figure 3. Signal processing results.

To validate the performance advantages of this method, experiments will be conducted using this method, traditional RBF neural networks, and Fourier transform FFT methods to process piano mechanical key signals. The signal processing accuracy of the three methods is shown in Figure 4. As can be seen from the comparison results in the figure, compared to the signal processing effects of traditional RBF neural networks and FFT methods, the signal processing accuracy of the method proposed in this paper can reach nearly 100%. This indicates that the proposed method achieves higher signal processing accuracy than other signal processing methods. This demonstrates that the proposed method has higher precision in extracting features from piano mechanical key signals and better signal processing performance, making it suitable for piano mechanical key signal processing tasks. Based on the comparison results in the figure, to more intuitively illustrate the signal processing performance of the proposed method, the experiment will plot histograms of the signal processing results from the three methods in the figure and compare the root mean square errors (RMSE) of the three methods.

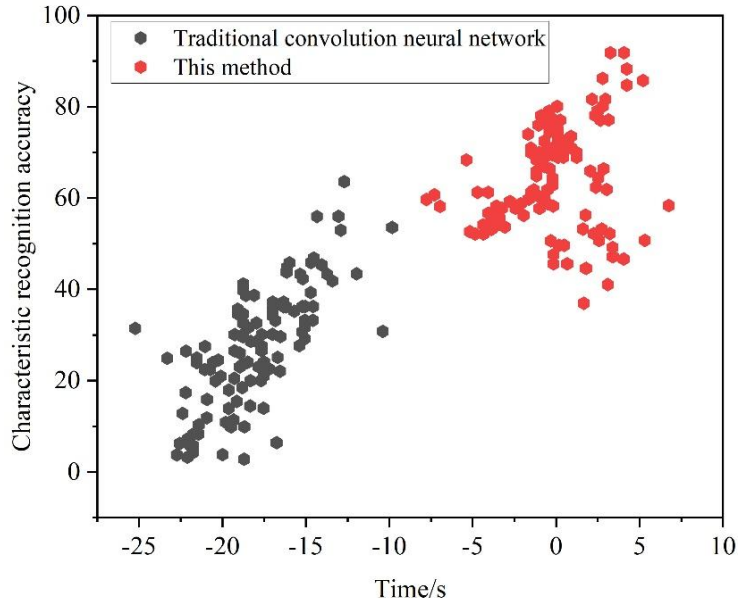


Figure 4. Audio signal processing accuracy.

The comparison results are shown in Figure 5. Within 10 s, 20 s, and 30 s, the root mean square error (RMSE) of this method was 4.49%, 3.30%, and 3.36%, respectively, while the RMSE of the traditional method was 5.80%, 6.88%, and 3.76%, respectively, all of which were higher than that of this method. A comprehensive analysis reveals that this method has a smaller root mean square error, achieving higher accuracy in feature extraction for piano mechanical key signals, thereby effectively enhancing the processing performance of piano mechanical key signals.

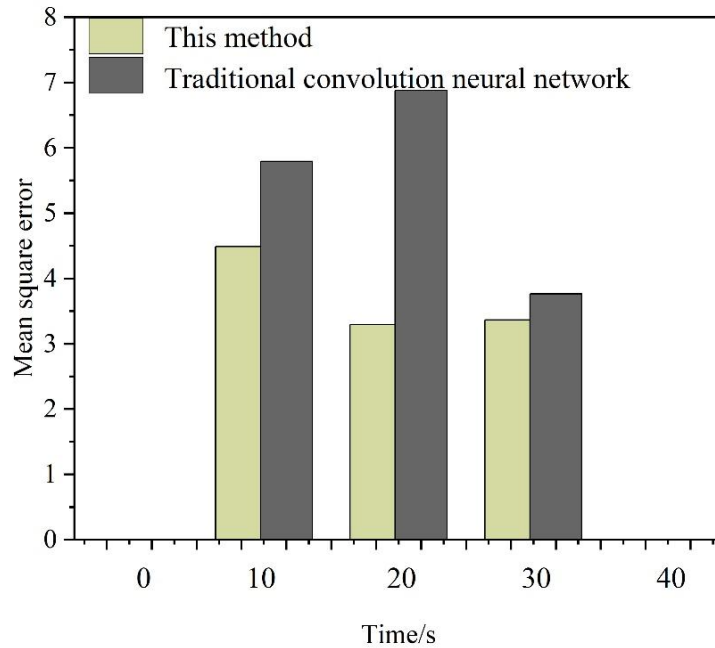


Figure 5. The average square root error of signal processing.

4. Optimization of Tone Quality through Noise Reduction in Piano Audio Signals

4.1. Dynamic Changes and Note Lengthening

The sound of the piano is rich and varied, with a wide range of expressive possibilities. In performance, dynamic changes and note prolongation are two important aspects that can bring rich emotion and realistic expression to music. Dynamic changes refer to fluctuations in volume within music.

By controlling the force and speed of the keys, piano players can freely adjust the volume, thereby creating various dynamic effects. For example, strong notes can convey passion and power, while soft notes can create a gentle atmosphere. Dynamic changes make the music more lively and dynamic, and also highlight the structure of the music and the turning points in its emotional expression.

4.2. Rhythm and Musicality

The piano is a versatile instrument capable of expressing rich rhythmic and musical qualities through various performance techniques. In piano performance, rhythmic sense refers to the combination of note duration and dynamics, while musicality refers to the performer's understanding and expression of the music. First, rhythmic sense plays a crucial role in piano performance. By selecting appropriate note durations and dynamics, performers can create various rhythmic patterns and rhythmic qualities.

4.3. Understanding Chord Progressions and Musical Contours

Piano tone is a diverse and expressive sound that helps us gain a deeper understanding of chord progressions and musical contours. When we discuss piano tone, we refer to aspects such as tonal texture, volume, tonal variations, and the manner in which notes are articulated. First, the richness of piano tone makes it an ideal choice for expressing chord progressions. The piano has a wide range, spanning from deep bass to bright treble, covering nearly the entire musical spectrum. This allows the piano to clearly present the pitch, sound characteristics, and volume changes of each note within a chord. By using different hand positions, volume control, and tone adjustment, piano players can express the unique characteristics and relative importance of each note within a chord progression. Secondly, the piano's timbre plays a crucial role in revealing the contours of music. The piano's timbre contains rich resonance and harmonics, enabling it to produce sustained reverberation and maintain a certain sonic presence even after the notes have been released.

4.4. Focus on the Balance between Melody and Accompaniment

The expression of piano tone is achieved through the balance between the melody line and the accompaniment. When playing the piano, the balance between the melody line and the accompaniment is crucial; they intertwine to create a beautiful musical landscape. First, the melody line is the most prominent and captivating part of the music. It represents the main theme and should be highlighted in the performance to capture the audience's attention. Performers must master excellent finger technique and musical expression to ensure the melody line is clear, flexible, and emotionally compelling. Each note must have distinct dynamics and articulation to highlight the melody line's distinctive characteristics.

5. Conclusion

This paper proposes a piano audio signal processing method based on a nonlinear denoising algorithm, primarily utilizing eigenmode functions and empirical mode decomposition to denoise the audio signal, thereby optimizing the aesthetic expression of the piano tone. Using the Yamaha YA128EX piano sound WAV file as the research sample, the proposed method was tested and analyzed via MATLAB simulation. The results show:

(1) The error in the tone prediction results of the proposed method is only 6. Therefore, the nonlinear denoising method proposed in this paper achieves high accuracy in feature extraction for audio signals, with minimal extraction error.

(2) When processing the mechanical key signal of piano No. 65, the proposed method demonstrates more significant effectiveness in audio signal feature extraction, with a notable improvement in the beam integration of feature recognition output.

(3) Compared to other methods, the feature recognition and extraction accuracy of piano audio signal processing based on the nonlinear denoising algorithm has been significantly improved.

Funding

This paper is a project of the Excellent Case Library of Professional Degree Graduate Students of Huaqiao University, "Original Micro-class Teaching Case Library" (22YJG017).

References

1. Haas, M. (2017). Tone colours on the piano: acoustic predisposition or contextual illusion. *Piano J Eur Piano Teach Assoc*, 112, 18-22.
2. Lv, J. (2014, June). Artistic aesthetics significance of piano improvising accompaniment. In 2014 International Conference on Education, Management and Computing Technology (ICEMCT-14) (pp. 107-110). Atlantis Press.
3. Sarac, J. (2024). Evaluation of Tone Quality in Piano Performance by Sight and Sound: A Performer's Perspective. *Music & Science*, 7, 20592043241302193.
4. Roudet, J., & Torbianelli, E. (2021, June). The Aesthetics of Singing in Romantic Piano Playing. In *Stimme–Instrument–Vokalität* (pp. 185-224). Schwabe Verlag.
5. Liddle, J. (2017). The sublime as a topos in nineteenth-century piano music. *Min-Ad: Israel Studies in Musicology*, 14, 37-64.
6. Kuang, J., & Guo, Q. (2024). Exploring the Creative Style and Aesthetic Features of Romantic Period Piano Music. *Cultura: International Journal of Philosophy of Culture and Axiology*, 21(4), 213-230.
7. Huang, D. (2022). Discussion on the Teaching Mode of Piano Performance from the Perspective of Aesthetics. *Frontiers in Educational Research*, 5(10).
8. Jia, H. (2023). Piano performance techniques and musical expressiveness. *Pacific International Journal*, 6(4), 34-37.
9. Riabukha, N. (2020). Sound Poetics in Claude Debussy's art of playing the piano. *musical art: historical and theoretical discourse*, 61.
10. Van Elferen, I. (2021). The Vibrant Aesthetics of Tone Color. *The Oxford Handbook of Timbre*, 69.
11. Phanichraksaphong, V., & Tsai, W. H. (2023). Automatic Assessment of Piano Performances Using Timbre and Pitch Features. *Electronics*, 12(8), 1791.
12. Li, S., & Timmers, R. (2020). Exploring pianists' embodied concepts of piano timbre: An interview study. *Journal of New Music Research*, 49(5), 477-492.
13. Liangguo, X. (2024). A Study Related to the Theory and Practice of Piano Tuning. *Frontiers in Art Research*, 6(8).
14. Zhang, X., & Kang, Q. (2022). The aesthetic emotional expression of piano music art in the background of Internet of things. *Frontiers in psychology*, 13, 974586.
15. Peng, Y. (2025). The analysis of optimization in music aesthetic education under artificial intelligence. *Scientific Reports*, 15(1), 11545.
16. Ali, M. N., El-Dahshan, E. S. A., & Yahia, A. H. (2017). Denoising of heart sound signals using discrete wavelet transform. *Circuits, Systems, and Signal Processing*, 36, 4482-4497.
17. Nigam, H. K., & Srivastava, H. M. (2023). Filtering of audio signals using discrete wavelet transforms. *Mathematics*, 11(19), 4117.
18. Yang, H., Li, L., & Li, G. (2020). A new denoising method for underwater acoustic signal. *IEEE Access*, 8, 201874-201888.
19. Wang, J., Li, J., Yan, S., Shi, W., Yang, X., Guo, Y., & Gulliver, T. A. (2020). A novel underwater acoustic signal denoising algorithm for Gaussian/non-Gaussian impulsive noise. *IEEE Transactions on Vehicular Technology*, 70(1), 429-445.
20. Szwajcowski, A., & Pilch, A. (2020). Optimization of piano tuning by means of spectral entropy minimization. *Applied Acoustics*, 166, 107359.
21. Liu, Q. (2023, July). Research on Aesthetic Emotion Expression of Piano Music Based on Deep Learning. In 2023 5th International Conference on Applied Machine Learning (ICAML) (pp. 420-425). IEEE.
22. Goebel, W. (2018). Movement and touch in piano performance. In *Handbook of human motion* (pp. 1821-1838). Springer, Cham.
23. Luo, G., Ning, Y., & Zong, L. (2023). The expression of sound art in piano performance. *Journal of Education and Educational Technologies*, 3(11), 17-22.
24. Bai, X. (2025). A Study on the Inaccuracy of Keystroke Force in Piano Performance and Its Optimization Strategies. *Journal of Humanities, Arts and Social Science*, 9(5).
25. Liu, L. (2025). Research on Piano Program Programming and Creative Optimization Methods Based on Big Data Technology. *International Journal of High Speed Electronics and Systems*, 34(02), 2440090.
26. Xuan-sen He & Tao Zhu. (2014). ICA of Noisy Music Audio Mixtures Based on Iterative Shrinkage Denoising and FastICA Using Rational Nonlinearities. *CSSP*, 33(6), 1917-1956.
27. Jiale Long, Chuisong Meng, Yi Ding & Jiangtao Xi. (2023). Suppressing the Zero-Frequency Components in Single Quantitative Phase Imaging by Filtering the Low-Frequency Intrinsic Mode Function Components. *Photonics*, 10(7).
28. Zhang Kuo. (2021). A denoising algorithm for linear frequency modulation signal based on window function and empirical mode decomposition. *Journal of Physics: Conference Series*, 1971(1).