Hindawi

Retraction

Retracted: Multimedia Recognition of Piano Music Based on the Hidden Markov Model

Advances in Multimedia

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Zhu, "Multimedia Recognition of Piano Music Based on the Hidden Markov Model," *Advances in Multimedia*, vol. 2021, 8 pages, 2021.

Review Article

Multimedia Recognition of Piano Music Based on the Hidden Markov Model

Ying Zhu ®

College Office, Zibo Vocational Institute, Zhangdian, Zibo 255300, Shandong, China

Correspondence should be addressed to Ying Zhu; lvbin@cupes.edu.cn

Piano performance is an art with rich artistic elements and unpredictable performance skills. It is an important carrier for playing beautiful piano sounds. The generation of musical tension and expression of piano performance is a vivid display of piano performance skills. In piano performance, we should pay attention to the cultivation and flexible application of performance skills. In order to ensure the richness and artistry of piano performance, it is fully based on the artistic characteristics of piano performance. Through in-depth analysis of the principle of the hidden Markov model, it is applied to the multimedia recognition process of piano playing music. In the process of obtaining the template, the fundamental frequency of the piano playing music differs greatly, and the piano playing music appears during the performance process. For the problem of low recognition rate, this paper proposes a multimedia recognition method for piano music. Finally, the analysis of experimental results shows that the method proposed in this paper has a 16% higher recognition rate than the traditional method, and it has a certain value in the multimedia recognition of piano music.

1. Introduction

With the continuous development of science and technology and electronic products, piano music recognition technology has also been further developed. The piano performance recognition technology has been transformed from the indoor research process to the market application stage and has developed to a relatively high level [1–3]. At present, the focus of domestic and foreign research is on the continuous performance of a nonspecific player. Multimedia recognition has high efficiency and accuracy, such as the content of music recognition, performance environment, piano used by players, and performance speed. There are certain differences between the existing piano music performance recognition technology and actual needs.

Due to the large differences between the characteristics of piano music and human piano music and the previous piano music recognition technology cannot meet the needs of piano performance music, the research on piano performance recognition technology has become the focus of research by scholars at home and abroad. This paper proposes a multimedia recognition system for piano performance music based on the hidden Markov model. Through continuous improvement of piano performance music recognition ability, it can be widely used in the field of piano performance music recognition.

2. Hidden Markov Model

The hidden Markov model is usually composed of multiple different states. Because of the continuous change of time, the hidden Markov model can stay in a certain state and can also switch between different states. Each observation vector can get different output frequencies for different states. In this paper, a hidden Markov model with four states S1~S4 is used for computational processing. The input observation sequence is represented by $o_1, o_2, \ldots, o_T$, and the transition probability between each state or between the states is represented by a0, where each observation sequence is used as each observation sequence. In this model, the observation

sequence $o_1, o_2, \ldots, o_T$ is used as the observable input sequence, but the state at each time cannot be directly observed.

The pitch part adopts the pitch period judgment method. When the difference between the pitch of the next frame and the average value of the pitches of all previously stored frames is less than the threshold, it is determined that these frames are frames of one category, and the process will continue until the difference between pitch cycles of the next frame is greater than the threshold. All these frames are processed as one frame, and the melting spectrum parameter Mel-ceptral (MCEP) is calculated to obtain the 13th MCP. The threshold defined in the experiment is 1. The schematic diagram is shown in Figure 1.

In order to maintain the balance of the data in the subsequent spectrum average (CMS) and correlated spectrum (PASTA) algorithm, we copy the calculated MCPP data in the corresponding 10 ms frame and perform spectral region filtering on it. The first-order and second-order differences are performed on the filtered parameters to obtain 39-order parameters. Finally, the parameters are Gaussized to improve the recognition rate.

The Markov model uses the Baum–Welch algorithm to solve the Markov parameter estimation problem and solve the hidden Markov training problem. Usually when using a given sequence of observations $O = o_1, o_2, \cdots, o_T$, this method uses the determined $\lambda = (A, B, \pi)$ parameter to ensure that the value of $P(O|\lambda)$ can reach the maximum value.

According to the related definitions of forward probability and backward probability, we know

$$P(O|\lambda) = \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_t(i) a_{ij} b_j(o_{t+1}) \beta_{t+1}(j) \quad (1 \le t \le T-1). \tag{1}$$

When using $P(O|\lambda)$ to reach the maximum value, since the training sequence of each experiment is limited, the best method of estimating parameters cannot be realized. In this case, the Bam–Welch algorithm uses $P(O|\lambda)$ with a recursive idea; the part is very large, and finally, the model parameter $\lambda = (A, B, \pi)$ is obtained.

The revaluation formula of the Baum–Welch algorithm is derived by recursion as

$$\overline{\pi} = \xi_1(i),$$

$$\overline{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i, j)}{\sum_{t=1}^{T-1} \xi_t(i)}, \tag{2}$$

$$\overline{b}_{jk} = \frac{\sum_{t=0 \text{ and } O_t = V_k}^{T} \xi_t(j)}{\sum_{t=1}^{T} \xi_t(j)}.$$

Among them, $\xi_t(i, j)$ represents the given training sequence O and the model parameter λ, the Markov chain at time t is in the θ_t state and the probability of the θ_i state at time $t + 1$, and $\sum_{t=1}^{T-1} \xi_t(i)$ represents the expected value of the number of transitions from the state θ_t to the state θ_i.

Define the auxiliary function as

$$Q(\lambda, \overline{\lambda}) = \frac{1}{P(O|\lambda)} \sum_{all \ S} P(O, S|\lambda) \ln P(O, S|\overline{\lambda}). \tag{3}$$

Among them, λ is the original model parameter $\lambda = (A, B, \pi)$, $\overline{\lambda}$ represents the model parameter $\overline{\lambda} = (\overline{A}, \overline{B}, \overline{\pi})$ to be solved, O represents the sequence of observations used in training and $O = o_1, o_2, \ldots, o_T$, and S as a certain state sequence $S = q_1, q_2, \ldots, q_T$.

The Markov model can not only find a good enough state transition path but also quickly calculate the output probability corresponding to the path. At the same time, the amount of calculation required by the method of using the Markov model to calculate the output probability is much less than that in the total probability formula.

Define $\delta_t(i)$ as $q_1, q_2, \cdots, q_t$ along a path at time t and $q_1 = \theta_i$; the maximum probability of $S = q_1, q_2, \ldots, q_T$ being generated, namely,

$$\delta_t(i) = \max_{q_1, q_2, \ldots, q_{t-1}} P(q_1, q_2, \ldots, q_t, q_1 = \theta_i, o_1, o_2, \ldots, o_t | \lambda). \tag{4}$$

The recursive form of the hidden Markov model is as follows.

(1) Initialization:

$$\delta_t(i) = \pi_i b_i(o_1) \quad 1 \le i \le N,$$
$$\psi_1(i) = 0. \tag{5}$$

(2) Recursion:

$$\delta_t(j) = \max_{1 \le i \le N} \left[\delta_{t-1}(i) a_{ij} \right] b_j(o_t) \quad 2 \le t \le T, 1 \le j \le N,$$
$$\psi_t(j) = \arg \max_{1 \le i \le N} \left[\delta_{t-1}(i) a_{ij} \right] \quad 2 \le t \le T, 1 \le j \le N. \tag{6}$$

(3) End:

$$P^* = \max_{1 \le i \le N} \delta_T(i),$$
$$q_T^* = \arg \max_{1 \le i \le N} \delta_T(i). \tag{7}$$

(4) Find the state sequence:

$$q_t^* = \psi_{t+1}(q_{t+1}^*) \quad 1 \le t \le T-1. \tag{8}$$

Among them, $\delta_t(i)$ represents the probability of accumulating the output value of the ith state at time t, $\psi_t(i)$ represents the continuous state parameter of the ith state at time t, q_t^* is the state at time t in the optimal state sequence, and P^* is the final output probability.

3. Multimedia Recognition Process of Piano Music

Piano performance music recognition has been successfully applied to smart devices such as mobile phones and TVs, which will have a profound impact on the future
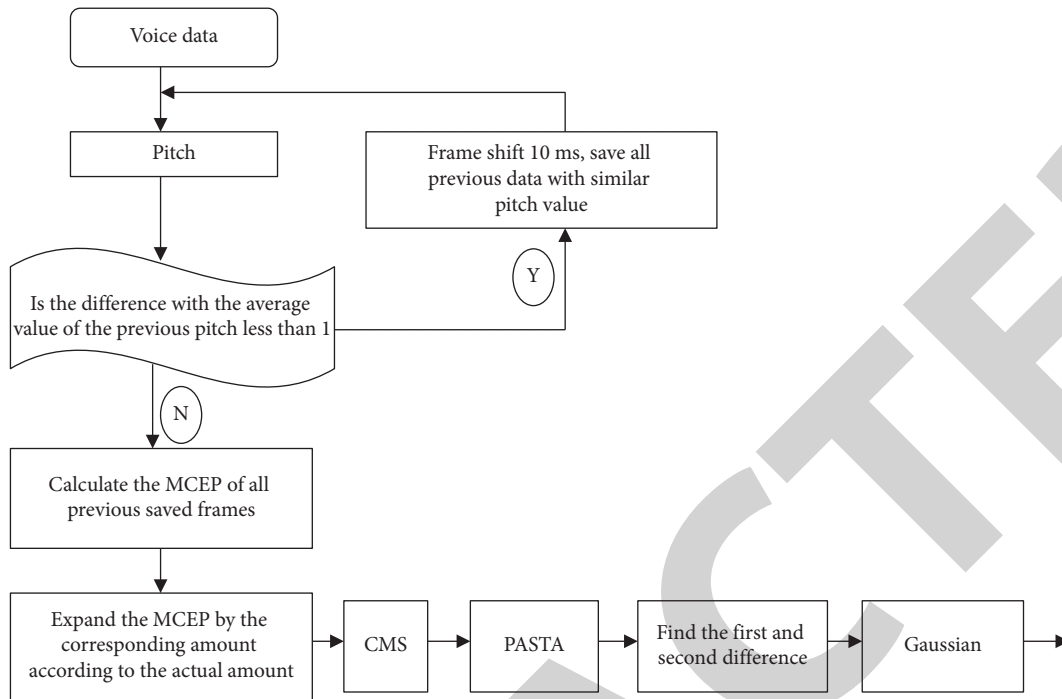
Figure 1: Pitch classification and recognition model.

lifestyle of mankind. The multimedia recognition of piano performances converts all the piano performance data into text form, breaks through the differences in language and intonation, causes communication barriers between machines and people, and uses the interactive system of piano performances as an important tool for human-computer dialogue. The construction of the piano performance music recognition system is carried out on a certain hardware condition and experimental platform. The piano performance music multimedia recognition is essentially a pattern recognition process. It mainly includes the pre-processing of the piano playing music signal, and its basic principle is shown in Figure 2.

It can be seen from Figure 2 that, in addition to the core recognition program, the piano performance music multimedia recognition system also includes piano performance music input, parameter analysis, and grammar language model construction. The piano performance music recognition system is mainly composed of three parts: piano performance music signal preprocessing, core calculation, and recognition basic data [4].

4. System Hardware Structure Design

The multimedia recognition of piano performance music based on the hidden Markov model correctly converts the received piano performance music signal, as shown in Figure 3, into a text form.

It can be seen from Figure 3 that the piano is a time-varying signal and has stability. Therefore, when the piano processes the music signal, it is necessary to use a function to distinguish the music signal with the piano. Each segment is called a frame, and there is a certain amount between
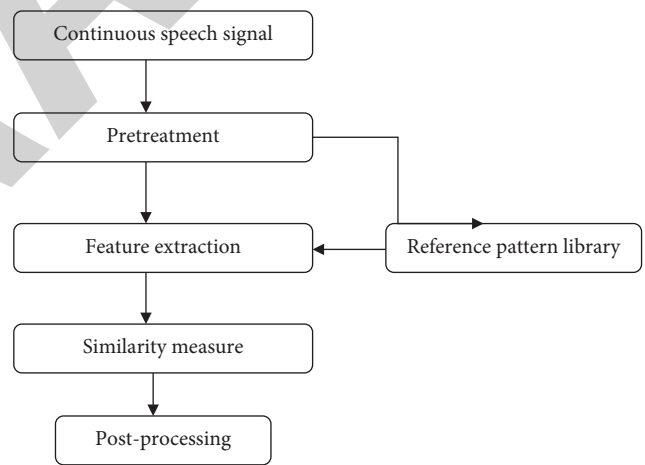

Figure 2: The basic principle of piano music recognition.

adjacent frames. The overlap can reduce jump changes. The robustness features of piano music signals are extracted from each frame, which can be used for noise removal and feature extraction [5, 6].

4.1. Piano Music Signal Processing Module. When the piano plays the music signal, it will change with time, but once the aliasing noise is generated, the piano performance signal processing will be invalid. Therefore, before the multimedia recognition, a low-pass filter must be used for aliasing prevention processing. Figure 4 shows the low-pass filter design of the piano music signal processing module.

It can be seen from Figure 4 that the high-fidelity OPA 604 low-pass filter is used to input the JFET, which has the characteristics of high carrier impedance and low distortion.
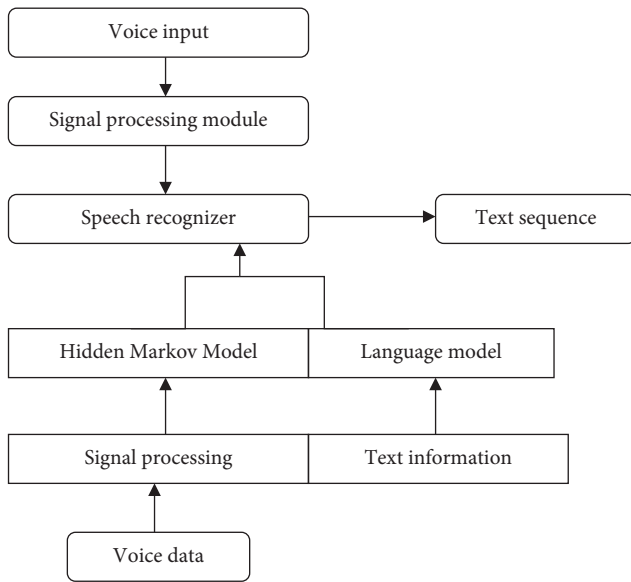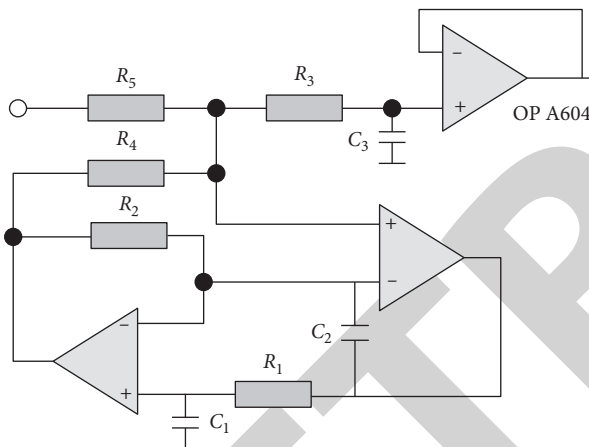
Figure 3: System hardware structure design.

Figure 4: Low-pass filter.

The signal processing process of piano music is not affected by aliasing noise, and correct and effective signal processing results are obtained. It can be guaranteed to provide accurate data for the multimedia recognition of piano performance music [7, 8].

4.2. Multimedia Recognition Module for Piano Playing Music. The music multimedia recognition module, which uses the result of the signal processing obtained by playing the piano to perform a large number of calculations, can use a DSP chip to process digital signals, and has the function of being compact and suitable for installation. DSP chips have strong online interaction capabilities. Choose OMP AP 5912ZZG model DSP chip to form a variety of development tools and multimedia database. You can use the system for free. The design of the multimedia recognition module for piano performances is shown in Figure 5.

It can be seen from Figure 5 that the OMUAP 5912ZZG type chip storage processor specification is 300 KB random access memory, and the piano performance data will be

buffered on the LCD. Use the memory card to expand the system memory, use the vector diagram to buffer the audio, and transplant the relevant piano performance music recognition sequence through the Ethernet interface [9, 10].

Design the hardware structure of the system in accordance with the principle of music recognition for piano performance. If you use the function to process the signal of the piano performance in segments, the jump change will be smaller. The process of piano playing music signal is affected by alias noise, and a low-pass filter to prevent aliasing is designed to ensure the accuracy and effectiveness of the signal processing result. The result of the signal processing is to use the multimedia recognition module of piano music to perform a large number of calculations. If the DSP chip of the OMP AP 5912ZG model is selected, the system design cost is greatly reduced. The recognition sequence of the relevant piano music is transplanted through the Ethernet interface to complete the system hardware structural design [11, 12].

4.3. System Software Function Design. According to the above-designed piano performance music multimedia recognition module, we design its software function [13, 14]. The specific design process is shown in Figure 6.

In the recognition of piano music, the sound has nonlinear characteristics, consistent with the transmission and reception of human auditory nerve signals, and the recognition efficiency is high. The characteristics of the piano music are used to process the filter samples, and the piano music is separated according to the frame. In order to blindly process the cell matching, the hidden Markov model can be used to smooth the signal transmission between adjacent frames of the piano music [15].

The multimedia recognition system based on the hidden Markov model to play music on the piano automatically selects the window function form according to the characteristics of the music played by the user on the piano. Both part-of-speech decoding and grammatical analysis are carried out under the Markov model to obtain the signal frequency of the piano music, and we use the hidden Markov model to transform the frame sequence and analyze and delete some invalid data in the frame sequence.

Through the above steps, the processing result of the piano performance music frame can be obtained, but the processed result is affected by sudden noise. The short-term average energy of some piano performance music frames increases sharply, and the recognition result obtained is incorrect. The design is shown in Figure 7.

The specific implementation steps are as follows:

(1) When playing the music signal with the piano, when it is in the silent stage, status = 0 increases the signal frame of the piano playing music signal. If a certain frame is short and the energy is too high, this frame is the starting point of the piano song signal. At this time, status = 1; the signal of piano performance has entered a transitional period.
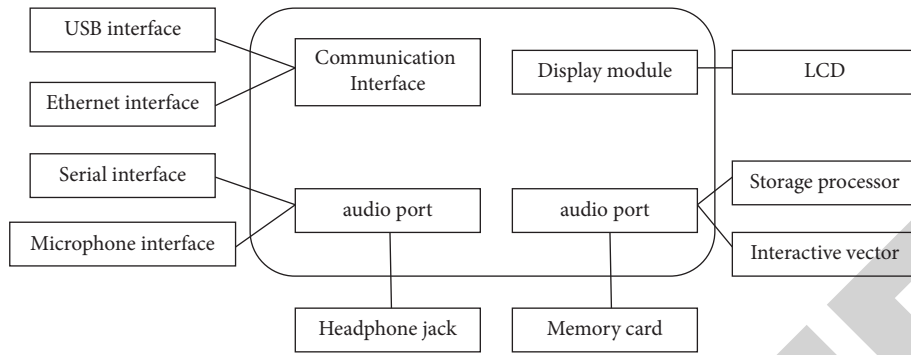
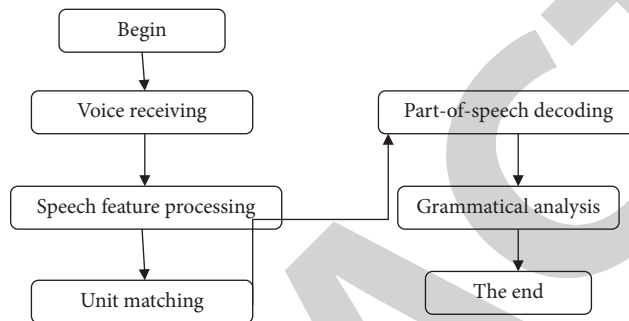Figure 5: Multimedia recognition module of piano playing music.
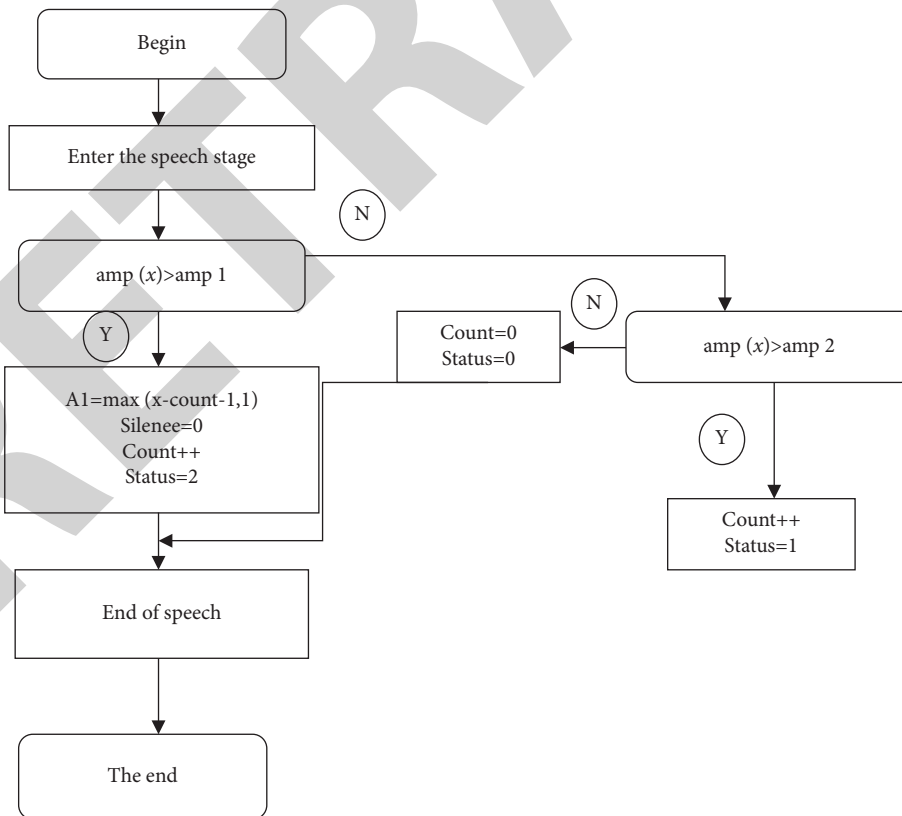

Figure 6: Software function design process.

Figure 7: Process of the processing stage.

(2) Continue to add piano performance signal frames. When a certain frame has too low energy in a short period of time, the frame indicates that the transition level returns to the silent stage, and status = 0 at this time.

(3) If the short-term energy of this frame is higher than amp1 and the frame number continues to increase, the signal can be determined to enter the piano playing music stage; at this time, status = 2; the current piano playing music frame frame number is the initial point of the piano playing music.

(4) When the current frame is used to play music on the piano, status = 2, and if the short-term energy of the frame for playing the music on the piano is lower than amp2, this segment is noise.

(5) Continue to increase the frame number. When the duration is longer than the mute phase, it means that the end point of the piano playing music signal is normal, and valid piano playing music can be output.

According to the system software design process, the implicit Markov model can be used to process the window operation of unit matching, so as to make the signal transmission between adjacent frames of the piano performance smoother. The window function form is automatically selected, and the frame sequence converted by the implicit Markov model is obtained. The result obtained is that due to the existence of part of invalid data, it is necessary to delete part of the data and design the process of the processing stage to complete the design of the system software part.

5. Experiment

In order to experimentally analyze the effectiveness of the piano performance music multimedia recognition system based on the hidden Markov model, it is necessary to extract a part of the piano performance music training group from the standard pattern recognition database.

5.1. Experimental Parameter Settings. The experimental parameter settings are shown in Table 1.

5.2. Experimental Environment Settings. In order to prevent the piano performance music multimedia recognition system installed on the computer from being affected by the hardware performance and not being able to exert all the performance, it is necessary to use the performance of the high-end system on the computer for experimental verification and analysis.

5.3. Multimedia Recognition Process of Piano Music. The first one is to obtain sound samples and conduct simulation experiments with a microcomputer. The four Chinese phrases of music, radio, GPS navigation, and air-conditioning are used as the experimental objects. Two samples of

each phrase are recorded through multiple four-phrase recordings to generate sample data.

Secondly, the initial model is trained through the Bam–Welch algorithm. Training is the most complicated problem in Hidden Markov, and it is the most important problem. Collect the N-path sound signal of a phrase and save it with the characteristic parameters of each frame signal. These parameters represent the characteristics of short-term speech fragments. Analyze the collected MFCC parameter vectors of all sample sounds through the clustering method to form a codebook vector set. Use the observation symbol instead. The training part is mainly based on the observation sequence to determine the two parameter matrices A and B. The specific steps of training are to first determine the initial values of A and B and use forward and backward algorithms to calculate forward and backward probabilities and output probabilities. λ, finally, confirms the convergence. The results can be repeated many times. Usually, we set the end times of 20 repetitions.

Third, we input the sound existing in the sample template, perform endpoint detection, and detect the start point and end point of the sound signal. In endpoint detection, the frame length is selected as 240 points. When the frame moves to 80 points, the for loop must be used to detect the endpoint of each word.

Fourth, we call the feature parameter extraction function, extract the feature parameters, process the sound signal frame by frame, and extract the frame parameters of the sound signal.

Fifth, we call the Viterbi function, input and recognize the voice and the template's voice, and find the maximum probability of the output template is the recognition result. In Figure 8, the training and recognition process of the speech recognition system based on the hidden Markov model is shown.

5.4. Experimental Results and Analysis. Based on the above experimental parameters and experimental environment, the recognition effect of the traditional system and the Markov model are compared and analyzed under the influence of sudden noise.

The piano music signal and short-term energy of these two systems were verified, and the results are shown in Figure 9.

It can be seen from Figure 9 that the signal of the previous system is 1000~2000 Hz and 6300~6900 Hz, and interruption occurs at 8900~9200 Hz, and the energy fails in a short time. On the contrary, based on the implicit Markov model, there is no interruption phenomenon, and short-term energy can be accurately obtained.

Based on the above comparison, the recognition effects of the two systems are compared under the influence of sudden noise, and the results are shown in Table 2.

From the comparison results in Table 2, it can be seen that the recognition effect based on the hidden Markov model is better than that of the traditional system.

According to the above content, when the noise is 20 dB, the following experimental results can be obtained. At 40 dB,

Table 1: Experimental parameter settings.

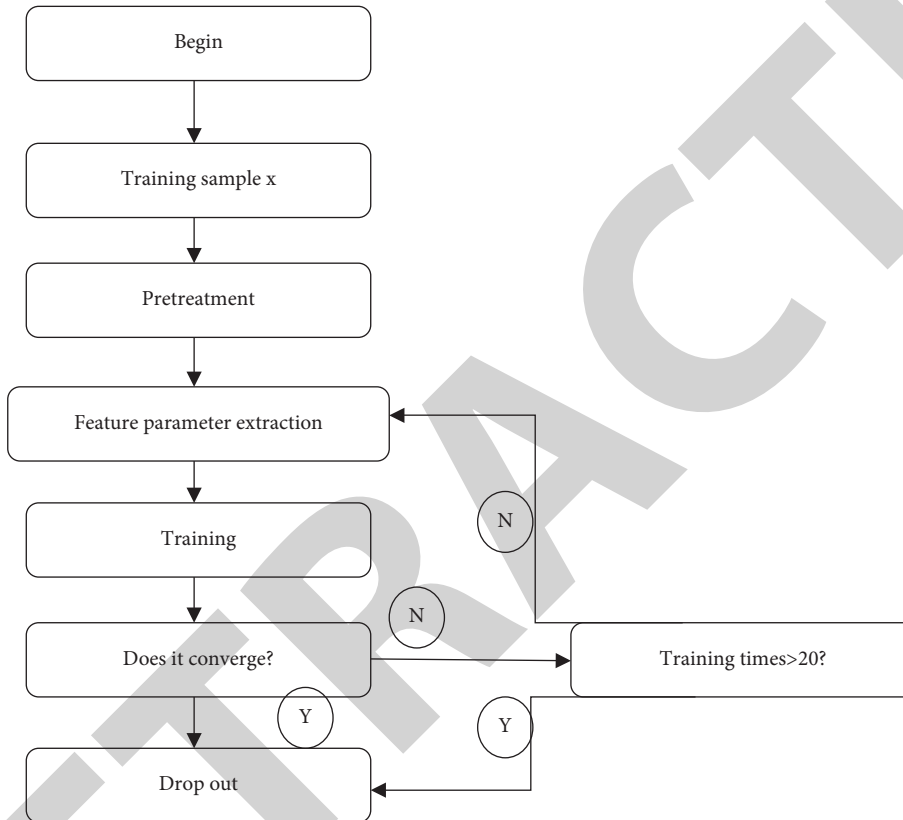|	Parameter	Numerical value
	Sampling frequency	25 kHz
	Hamming window	25 dimensions
Piano playing music signal window	Length	20 ms
	Framing	260 points
	Frame shift	90points
	Parametric vector quantization	60code

Figure 8: The training and recognition process of the speech recognition system based on the hidden Markov model.
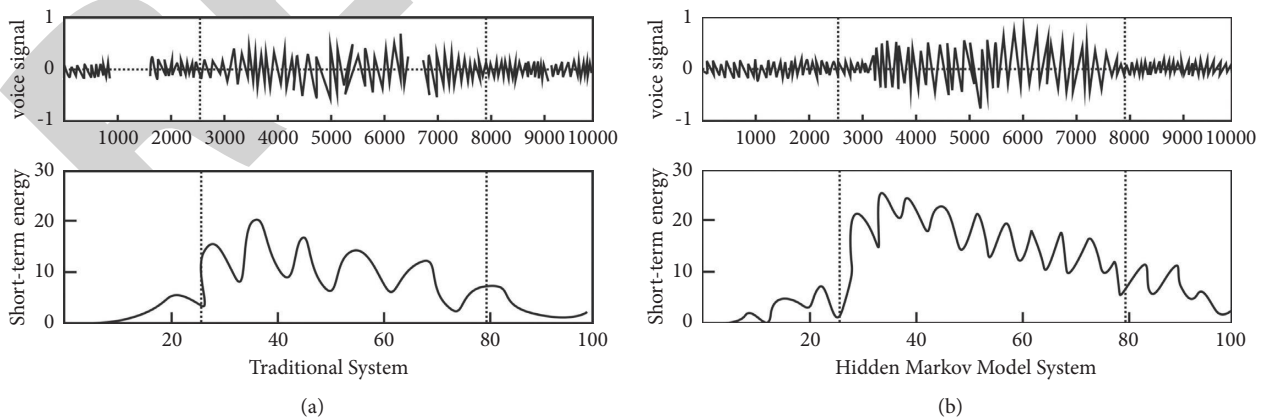
Figure 9: Two types of system endpoint detection.

Table 2: Comparison of the recognition effect of the two systems (%).

Noise (dB)	Traditional system	Based on hidden Markov model
20	72	87
40	66	85
60	63	61
80	45	95
100	41	90

60 dB, 80 dB, and 100 dB, based on the implicit Markov model, the recognition effect is 15% higher than that of the traditional system: 20%, 26%, 22%, and 48%. Therefore, based on the hidden Markov model, it can be seen that the design of the multimedia recognition system for playing music on the piano is effective.

6. Conclusions

Because the existing piano performance music recognition system is relatively complicated and restricted by time conditions, the traditional recognition methods are easily affected by sudden noise, and the piano performance music recognition is relatively poor. Based on the hidden Markov model, this paper proposes a multimedia recognition method for piano performances. This method is aimed at sudden noise, plus a low-pass filter. However, due to environmental factors, the recognition performance of piano music is enhanced, the signal-to-noise ratio is improved, and the accuracy of piano music recognition is improved.

Data Availability

The data used to support the findings of this study are available upon request to the author.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] J. M. Orjuela-Rojas and P. Montaés, "Recognition of musical emotions in the behavioral variant of frontotemporal dementia," *Revista Colombiana de Psiquiatria*, vol. 64, no. 11, pp. 620–623, 2021.

[2] X. Wang, "Research on the piano teaching mode based on the computer platform," *Revista de la Facultad de Ingenieria*, vol. 32, no. 12, pp. 1095–1099, 2017.

[3] None, "Music-making piano droid [the big picture]," *IEEE Spectrum*, vol. 54, no. 8, pp. 16-17, 2017.

[4] E. Togootogtokh, T. K. Shih, W. G. C. W. Kumara, S.-J. Wu, S.-W. Sun, and H.-H. Chang, "3d finger tracking and recognition image processing for real-time music playing with depth sensors," *Multimedia Tools and Applications*, vol. 77, no. 8, pp. 9233–9248, 2018.

[5] C. Traube, "Piano tone control through variation of "weight" applied on the keys," *Journal of the Acoustical Society of America*, vol. 141, no. 5, p. 3874, 2017.

[6] K. Mcneely-Whtie and A. M. Cleary, "Music recognition without identification and its relation to déjà entendu: a study using "piano puzzlers"," *New Ideas in Psychology*, vol. 55, pp. 50–57, 2019.

[7] L. Chen, "An optimization analysis of modern piano playing mode based on ultimedia system," *Boletin Tecnico/Technical Bulletin*, vol. 55, no. 11, pp. 519–525, 2017.

[8] D. Johnson, D. Damian, and G. Tzanetakis, "Detecting hand posture in piano playing using depth data," *Computer Music Journal*, vol. 43, no. 1, pp. 59–78, 2020.

[9] A. Cosimato, R. D. Prisco, A. Guarino, D. Malandrino, and R. Zaccagnino, "The conundrum of success in music: playing it or talking about it?" *IEEE Access*, vol. 7, pp. 1–5, 2019.

[10] A. Adamyan and A. Anna, "Music education issues for adult beginners in Armenia: specifically the analysis of the difficulties of piano playing," *British Journal of Music Education*, vol. 33, no. 8, pp. 66–79, 2018.

[11] C.-J. Chau and Y. Hong, "An analysis of low-arousal piano music ratings to uncover what makes calm and sad music so difficult to distinguish in music emotion recognition," *Journal of the Audio Engineering Society Audio Acoustics Applications*, vol. 33, no. 2, pp. 133–145, 2017.

[12] F. Vugt and E. Altenmüller, "On the one hand or on the other: trade-off in timing precision in bimanual musical scale playing," *Advances in Cognitive Psychology*, vol. 15, no. 1, pp. 197–210, 2019.

[13] N. Hajj, M. Filo, and M. Awad, "Automated composer recognition for multi-voice piano compositions using rhythmic features, *n*-grams and modified cortical algorithms," *Complex & Intelligent Systems*, vol. 4, no. 1, pp. 55–65, 2018.

[14] H. D. Ecer and S. Saritaş, "The effects of music on the life signs of patients in the reanimation unit/recovery room after laparoscopic cholecystectomy," *Holistic Nursing Practice*, vol. 33, no. 5, pp. 295–302, 2019.

[15] A. Campayo-Muoz, D. Cabedo-Mas, and Hargreaves, "Intrapersonal skills and music performance in elementary piano students in Spanish conservatories: three case studies," *International Journal of Music Education*, vol. 38, no. 1, pp. 93–112, 2019.