Gesture Recognition Technology of VR Piano Playing Teaching Game based on Hidden Markov Model

Min Zeng
Music College, Hunan Normal University, China
minzeng012@outlook.com

Abstract: *Virtual Reality (VR) is a kind of simulation environment generated by computer simulation. It is a kind of isolation of users from the physical environment, and users are fully engaged in the simulation environment. In music teaching, especially in the learning of piano performance, VR technology can simulate the real playing experience without physical piano, so that learners can learn and practice piano anywhere. At present, the traditional way of learning piano often needs the guidance of professional piano teachers, and learners must study and practice beside the piano, which undoubtedly brings limitations to the time and space of learners. VR technology provides a more convenient way for learners to learn and practice without a physical piano, making learning more flexible and free. Firstly, a piano playing teaching game platform based on VR technology is constructed to provide a virtual piano playing scene. Then, by collecting a large number of piano players' gesture data, they are accurately marked. Finally, the Hidden Markov Model (HMM) is used to train and model the gesture data, and the key features of the gesture sequence are extracted to realize the recognition and real-time feedback of the piano player's gesture. According to the findings, the recognition accuracy of the Principal Component Analysis-HMM (PCA-HMM) without playing correction is very low after playing correction. Compared to the logarithmic likelihood ratio test, the recognition rate of the quadratic statistical model is only 15%. The results show that this method has good iterative convergence and convergence performance, which can effectively avoid falling into local optima. Compared with the traditional hidden Markov algorithm, it reduces 96.05% and 89.48% respectively. Compared with the fuzzy hidden Markov algorithm, it reduces 90.20% and 73.35% respectively. The average time and standard deviation for real-time positioning and map construction are 0.251s and 0.152s, respectively. The fitness of HMM based on fuzzy Particle Swarm Optimization (PSO) is 6.8% and 2.6% higher than that of the hidden Markov algorithm and the fuzzy hidden Markov algorithm, respectively. At the same time, in the comparison between the research model and other algorithm models, the accuracy and other parameter performance of the research model are better, with the highest error value of only 3.5, which is 5.9 lower than the lowest Convolutional Neural Network (CNN) algorithm model. It can be seen that the error value of the model used in the study is lower and the performance is better. Learners can receive real-time guidance and a more immersive playing experience through this technology. This provides useful reference for further exploring the combination of music teaching and VR technology.*

Keywords: *HMM, piano performance, gesture recognition, SMA, VR technology.*

Received January 31, 2024; accepted April 28, 2024
https://doi.org/10.34028/iajit/21/4/15

1. Introduction

With the continuous development of science and technology, Virtual Reality (VR) is increasingly integrated into People's Daily life, bringing users a unique immersive experience. In this context, the application of VR technology in music education has attracted wide attention, especially in piano teaching. VR was chosen over other XR technologies for the study because VR technology provides an all-encompassing immersive experience that allows students to better focus on piano learning and can simulate a real piano playing environment, resulting in better learning outcomes [2, 15]. One of the challenges in achieving accurate and efficient piano playing instruction in a VR environment is how to accurately recognize the learner's gestures. As for the challenges of music teaching methods, traditional music teaching methods often require professional teachers to teach face-to-face, which may lead to a shortage of teachers in many areas. On the other hand, traditional music teaching methods are often rigid and may not meet the learning needs of all students. In response to these problems, VR piano teaching tools can provide a more flexible and personalized way of learning through virtual environments. Therefore, this paper will deeply study and apply gesture recognition technology based on Hidden Markov Model (HMM) in VR piano playing teaching games [9, 16]. HMM model is a statistical model widely used in serial data processing, with excellent time series modeling ability, and has been proved to be an efficient method in the field of gesture recognition. Piano playing is an important form of expression and creative pursuit for many people. Traditional piano teaching methods usually require students and teachers to meet face to face at the same time and place, which may be limited by time and place for many students. Moreover, the traditional piano teaching method requires a physical piano, which is a difficult threshold for some students who do not have a piano or do not have the conditions to buy one. As for why VR tools are needed for piano teaching, traditional piano teaching methods usually require students and teachers to communicate face to face at the same time and place, which may be limited by time and place for many students. Moreover, the traditional piano teaching method requires a physical piano, which is a difficult threshold for some students who do not have a piano or do not have the conditions to buy a piano. However, VR piano teaching tools allow students to learn anytime, anywhere, and do not need a physical piano, students can learn and practice piano performance in a virtual environment, greatly reducing the learning threshold. VR technology offers the possibility of making piano teaching more affordable and accessible. Firstly, the importance of gesture recognition in piano teaching is explained, and the limitations of the existing gesture recognition technology are deeply discussed. Then, the basic principle of HMM model and its application in gesture recognition will be introduced in detail. Several mainstream gesture recognition methods are compared to demonstrate the advantages of HMM in VR piano playing teaching games. Finally, by comparing the performance of different gesture recognition technologies in VR piano playing teaching games, the effectiveness of the proposed gesture recognition method based on HMM model will be obtained [10, 12]. As for the difference between this study and other VR piano teaching tools, this study mainly studies how to use HMM for gesture recognition, provide more accurate and real-time feedback, and help students learn piano playing better. At the same time, according to the characteristics of piano performance, this study also introduces some specific model optimization to make gesture recognition better. The study will be conducted in four parts. Firstly, the research will give an overview of gesture recognition technology for VR piano teaching games based on HMM in the first part. This part will summarize the application of gesture recognition technology in VR piano teaching games, and the basic theory of gesture recognition model based on HMM. Secondly, the second part will deeply study the method of gesture recognition technology of VR piano teaching game based on HMM. This section will elaborate on how to use HMM to build and optimize gesture recognition models, and how to implement this model in a VR environment. Next, in the third part, the method of the second part will be experimentally verified. The study will test the effect of the gesture recognition model through an actual VR piano teaching game to prove the effectiveness of the method. Finally, the fourth part will summarize the research content and point out the shortcomings. In this section, the research results will be reviewed, and point out the problems that may be encountered in the practical application of the current method and the places that need further improvement.

2. Related Works

HMM is widely used in Natural language processing, speech recognition, Gesture recognition, bioinformatics and other fields. In these fields, HMM model can be used for tasks such as language modeling, speech recognition, and action recognition. By inferring the observation sequence, hidden states are analyzed and predicted. To study the propagation mode of muscle weakness, Shin *et al*. [17] proposed a method based on HMM by analyzing the spatiotemporal binary muscle strength data. This model can estimate the incidence rate of muscle weakness and the probability of disease state transition. The research results indicate that the spatial network of susceptible muscles follows a Markov process. Therefore, this model can flexibly and effectively monitor muscle condition. Xia and Chen [21] found that in daily life, SMS services are occupied by spammers. Although traditional filtering techniques can improve email filtering performance, algorithms have become increasingly complex. To solve this problem, a discrete HMM is proposed. This model not only greatly reduces the complexity of filtering technology, but also makes short message services more secure and efficient.

Zhiyong [23] have found that there is a significant amount of uncertainty in predicting the remaining service life of mechanical systems at the current stage. Reducing this uncertainty is a very important aspect in service life prediction. To solve this problem, a correct time prediction method is proposed, and an HMM is established to map the entire path. This model can reduce the predictive uncertainty of mechanical systems under unobservable degradation conditions. The research results indicate that this method has higher accuracy than traditional methods. To accurately estimate the health index status of power transformers and predict the early operation faults of transformers, Jiang *et al*. [11] proposed an HMM dynamic fault prediction technology based on dissolved gas analysis. The static characteristics of different health states are extracted from the dissolved gas analysis dataset of power transformers under working conditions. Then the static characteristics between the health status and dissolved gas concentration are established. The experimental results indicate that this power transformer based on HMM model can accurately provide health indicators for power transformers. Don and Khan [8] proposed to combine the HMM-Bayesian network hybrid system with a new prediction technology to predict and isolate multiple identified faults in the Tennessee Eastman process. In the experiment, the HMM normal operating condition data is trained offline. The log likelihood value of the online data string is determined after repeated training. Compared with historical values, the most likely future state of the system is predicted. The experimental results indicate that the system successfully predicts all selected faults in the Tennessee Eastman process, while accurately isolating multiple faults within them.

Li *et al*. [13] proposed the "Magnetic resonance Imaging of Piano" technique, which is the first parametric skeletal model of the hand based on human magnetic resonance imaging data. It is biologically accurate, easy to animate, and enables more accurate anatomical modeling than traditional hand models. This model can be used at the neural network layer to achieve fine-grained semantic loss training, driving new tasks of hand bone anatomy and semantic understanding from MRI or RGB images. Briggs *et al*. [3] proposed radio frequency identification technology based on the current issue of real-time and remote wildlife monitoring. This technology can remotely identify individual animals implanted with passive integrated transponder tags. Through experiments, the number of animal individuals recorded using radio frequency identification technology was compared with the number of animal individuals based on on-site capture records. The two data were relatively close. Therefore, the emergence of this technology has greatly reduced the invasiveness of direct contact with research animals, while greatly improving the accuracy and remote efficiency of monitoring results. Deng *et al*. [7] found that it is still difficult to accurately determine whether small diameter nodules are hepatocellular carcinoma or cirrhosis using current diagnostic methods. A new recognition technique was proposed to distinguish the two types of liver nodules mentioned above. The photoacoustic imaging method was used to project small non-invasive modules into the image. Then a probe was developed to specifically target hepatocellular carcinoma by identifying biomarkers on the cell membrane of hepatocellular carcinoma cells. The experimental results indicate that the imaging of this recognition technology can improve tissue resolution and sensitivity. In addition, the accuracy and precision of this specific probe targeting hepatocellular carcinoma were previously unattainable.

To sum up, many researchers and scientists have obtained a large number of phased experimental results in the field of HMM or recognition technology. These remarkable research results have laid a good foundation for the combination of HMM and recognition technology. This research on Gesture recognition technology of VR piano playing teaching game based on HMM aims to achieve accurate recognition of piano playing gestures by using HMM model, so as to provide better interactive experience and teaching effect.

3. Gesture Recognition Technology Model of VR Piano Playing Teaching Game Based on HMM

With the rapid development of VR technology, providing students with a more immersive and interactive learning experience has been widely applied in fields such as teaching and training. With the rapid development of VR technology, providing students with more immersive and interactive learning experience has been widely used in teaching, training and other fields. "Statistical models of time series based on Hidden Markov Models (HMM) algorithm" will introduce the fundamental theory of HMM, including the definition of HMM, its basic assumptions, theoretical principles, and its application in time series analysis. "Virtual Reality (VR) piano performance design based on amazon lumberyard" is likely to focus on the use of Amazon Lumberyard game engine to design and build VR piano performance system, including explaining the design concept, system module construction and specific implementation technology. "Application of HMM model and gesture recognition Technology in VR piano design" describes the application of combining HMM algorithm and gesture recognition technology in VR piano design, may elaborate the basic concept and implementation of gesture recognition technology, and discuss how to combine HMM model and gesture recognition technology together. Finally, it may be proposed how to optimize and improve the experience of VR piano playing by this combined method.

3.1. Research on Piano Teaching Recognition Based on HMM

The development of gesture recognition technology opens up new possibilities for piano teaching. Recognition techniques provide feedback on skill and style by analyzing and understanding hand movements. The key to this technique is to understand the generation of random sequences and how to generate observable random sequences from unobservable random sequences. In order to integrate these two fields effectively, it is necessary to conduct in-depth research on the time probability model. Such models infer hidden conditions or variables by looking at the time of the data. In this model, the generated random sequence of states is called an "array of states," each state has an observation value, and the resulting result is called an observation sequence [18, 19]. By applying this theory to the intersection of piano teaching and gesture recognition technology, learners can be given more intuitive and personalized feedback. The initial probability of HMM is shown in Equation (1).

$$\begin{cases} \lambda = (A, B, \pi) \\ Q = \{q_1, q_2, \ldots, q_N\} \\ V = \{v_1, v_2, \ldots, v_M\} \end{cases} \tag{1}$$

In Equation (1), A and π are state sequences. B is the determining observation sequence. Q is the collection of all possible states. N is the number of possible states. V represents all possible observed sets. D M is the possible number of observations. π is the initial state probability vector. The state sequence and observation sequence of HMM are shown in Equation (2).

$$\begin{cases} O = \{o_1, o_2, \ldots, o_T\} \\ I = \{i_1, i_2, \ldots, i_T\} \end{cases} \tag{2}$$

In Equation (2), I is the state sequence with a length of T. O is the corresponding observation sequence. The operation of the observation matrix and state matrix is shown in Equation (3).

$$\begin{cases} A = \left[a_{ij}\right]_{N \times N} \\ B = \left[b_{ik}\right]_{N \times M} \end{cases} \tag{3}$$

In Equation (3), a_{ij} is the probability of state locking from the condition at time t to time t+1. B is the observation probability matrix. b_{ik} is also the probability of the state being locked at a later moment. For the probability calculation of HMM, the output probability of HMM's state sequence is shown in Equation (4).

$$P(O|Q,\lambda) = \prod_{t=1}^{T} P(o_t|q_t,\lambda) = b_{q1}(o_1) b_{q2}(o_2) \cdots b_{qr}(o_r) \tag{4}$$

In Equation (4), $P(O|Q,\lambda)$ includes the possible occurrence of all hidden state sequences of observation sequence O. The complex calculation is approximately in $2TN^{-T}$. The related direct algorithms include brute force algorithm, Forward algorithm and backward algorithm. In the related Dynamic programming algorithm, the complexity of this calculation reduces the weight order. The forward probability $a_t(i)$ and observation sequence probability that can be recursively calculated are shown in Equation (5).

$$a_{t+1}(i) = \left[\sum_{j=1}^{N} \alpha_t(j) \alpha_{ji}\right] b_i(o_{t+1}), i = 1, 2, \ldots, N \tag{5}$$

In Equation (5), $a_{t+1}(i)$ is the value of the next state under the forward probability. $a_t(i)$ is the forward probability. The final termination formula can be obtained, as shown in Equation (6).

$$P(O|\lambda) = \sum_{i=1}^{N} \alpha_T(i) \tag{6}$$

In Equation (6), $P(O|\lambda)$ is the probability of the observation sequence. Based on recursive calculation, the backward algorithm of Equation (7) can be obtained.

$$\begin{cases} \beta_t(i) = \sum_{j=1}^{N} a_{ij} b_i(o_{t+1}) \beta_{t+1}(j) \\ P(O|\lambda) = \sum_{t=1}^{N} \pi_i b_i(o_1) \beta_1(i) \end{cases} \tag{7}$$

In Equation (7), the backward probability is $B_t(i)$. The probability of the observation sequence is $P(O|\lambda)$. Furthermore, the relationship between forward and backward probabilities is shown in Figure 1.

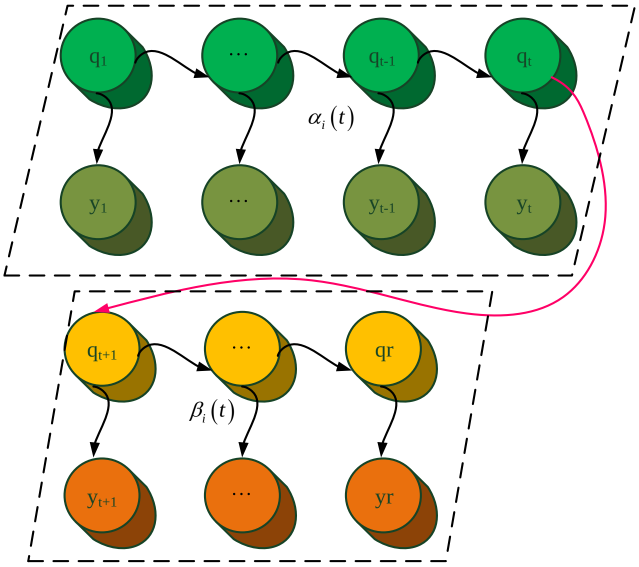


Figure1. Graph of the relationship between forward and backward probabilities.

In Figure 1, the relationship between forward and reverse probabilities is very close. In fact, in a probability graph, the calculations of other probabilities are often combined with each other. The forward and backward probabilities can be connected by a method called forward and backward algorithms. On this basis, a new method based on the multiplication of forward and reverse probabilities is proposed. Algorithms in both positive and negative directions are widely used for modeling many sequence models [5, 20]. Viterbi algorithm is an unsupervised learning approach. Model parameter learning can be achieved using the EM algorithm. The E step of the EM algorithm is shown in Equation (8).

$$Q(\lambda, \bar{\lambda}) = \sum_{I} \log P(O, I|\bar{\lambda}) \tag{8}$$

In Equation (8), $\bar{\lambda}$ is the current estimated value. λ is the maximized HMM parameter. The M step of the EM algorithm combines the Lagrange function to calculate the parameter π_i, as shown in Equation (9).

$$\pi_i = \frac{P(O, i_I = i|\bar{\lambda})}{P(O|\bar{\lambda})} \tag{9}$$

In Equation (9), P is the conditional probability distribution of hidden variables. The prediction problem, namely the Verterbi algorithm, defines a variable, as shown in Equation (10).

$$\delta_t(i) = \max P\left(i_t = i, i_{t-1}, \ldots, i_1, i_t = I_t, o_t, \ldots, o_1 \mid \lambda\right) \tag{10}$$

In Equation (10), when $\delta_t(i)$ is time t, the state of the system is I_t. The recursive process of the Verterbi algorithm is shown in Equation (11).

$$\begin{cases} \delta_t(i) = \max\limits_{1 \le j \le N}\left[\delta_{t-1}(j) a_{ji}\right] b_i(o_t), i = 1, 2, \ldots, N \\ \psi_t(i) = \arg\max\limits_{1 \le j \le N}\left[\delta_{t-1}(j) a_{ji}\right], i = 1, 2, \ldots, N \end{cases} \tag{11}$$

According to the calculation process of Equation (11), the probability matrix can be calculated to achieve the decoding process of the HMM algorithm, as shown in Figure 2.

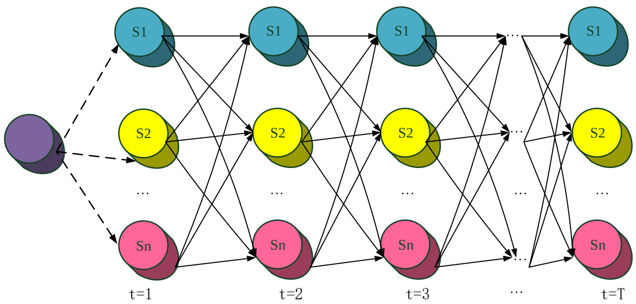


Figure 2. The decoding process of HMM algorithm.

In the HMM of Figure 2, a series of observed data can be seen, but a hidden series of states cannot be seen. The purpose of decoding problems is to identify the most likely hidden state sequence from a known observation data. On this basis, a probability matrix is established to record the possibility of each hidden state occurring at each moment. The path to achieve this possibility is provided. By using recursive methods, the probability matrix is continuously updated to obtain the maximum hidden state sequence.

3.2. VR Piano Performance Design Based on Amazon Lumberyard

VR technology is changing various fields at an astonishing speed. Education and entertainment are one of the areas that have benefited the most. By using VR technology, people can experience the joy of playing the piano in a virtual music world without the need for actual piano instruments. Based on VR technology, players can wear VR headsets. Then the player simulates the action of playing the piano through a controller or customized hand tracker. VR technology can also provide interactive learning experiences. According to the teaching mode, the skills of playing the piano can be learned. Through virtual guidance and instant feedback, performance skills have been improved. Interacting with other virtual piano enthusiasts, forming bands, or participating in piano competitions can be achieved. Through VR, the joy of playing the piano can become more convenient, interactive, and creative [4, 6]. The model equipment parameters are shown in Table 1.

Table 1. Parameters of model equipment.

GPU	NVIDIA GeForce GTX 2080
CPU	Intel(R)Xeon(R)I7-3080
RAM	16.0 GB
Video output	HDMA1.4
USB Port	USB3.0 Port
Operating system	Windows 10 operating system

As shown in Table 1, the device utilizes advanced positioning techniques. The oculus quest pro and apple vision pro have more than 680 built-in sensors that scan the location space on the headset and controller. The user can move freely within the detection range, which enhances the sense of depth and immersion. At the same time, these devices are equipped with powerful built-in graphics cards to provide a smooth virtual experience. Its functional modules include hand modeling module, piano modeling module, Musical Instrument Digital Interface (MIDI) audio module and gesture recognition module.

In the piano design module, realistic piano design is achieved using 3D modeling techniques. This includes creating the appearance of the piano, the keys, the body, the playing area, and so on. Every detail and dimension should match the real piano. Physical simulation is an important part in the piano modeling module. A new simulation method is investigated for physical properties such as string vibration and key response. By scanning and collecting the human body posture, data closer to the innate perception of the human body is obtained. The piano design is based on the Unity3D professional game engine. It has two biggest advantages over other game development tools. One is to provide extremely high visual workflows, and the other is multi-dimensional cross platform support. Depending on the visual processes, it is easy to edit scene layout, bind resources, and write interactive Scripting language. To enrich the performance effect, when the piano is played, when the fingers touch the keys, the corresponding effect will be feedback to the user. An additional particle system from Unity3D is adopted. The particle system is an orderly mechanism used to create and manage many moving objects, which is useful in creating visual effects such as flames, smoke, and water spray. The leap motion Software Development Kit (SDK) is input into Unity3D. When the button is pressed, the color on the button is set to follow the color of the flame. When the fingers approach the keys, the Fa in the high pitch area is activated, creating a visual effect of sparks [3]. For the construction process of the virtual piano scene needs to recognize the piano action gestures through the sensor, which comes with two high-definition high-resolution cameras to collect and analyze the piano gesture movements from different angles. When the sensors are collecting data on hand movements, the model detects the current hand movements being photographed through the hand-operated controller, and at the same time acquires the coordinate data of the palm position through data collection.

For the virtual piano scene design is to set the traditional piano keys, the bass, alto and treble are set to 7 tones, a total of 21 keys, the total number of black chromatic keys have 15, and the virtual scene of the piano through the three-dimensional slow rendering of the interface, will be rendered into two modes, screen space-overlay and screen space-camera two models, mode scene all the canvas is covered with the canvas space, and the entire piano element is placed in the top layer of the space and will not be affected by other factors and caused by the obscuring situation, at the same time, the scene can be based on the distance between the piano and the camera to the size of the piano for the scaling process, the entire training of the piano is made through the rectangular body padding [14, 22]. When practicing gestures, it is necessary to disassemble the gestures and extract the eigenvalue size of each frame, and then encode and divide the actions into speeds, so that the gestures that are disassembled into multiple actions can be divided into vectors of multiple dimensions, and then analyze the data through the model to output the parameter data of the current gestures, and then build a virtual simulation for teaching through these data [1].

3.3. The Application of HMM Model Combined with Gesture Recognition Technology in VR Piano Design

In the piano teaching design of VR, gesture recognition technology plays a key role. The technology can enhance the interactive experience of students by simulating natural gestures to achieve piano playing, thereby stimulating interest in learning and enhancing engagement. At the same time, gesture recognition technology can also provide teachers with information about students’ learning progress and skill level. Second, the oculus quest 2 uses hand tracking technology to easily mimic the keystrokes of a piano. However, when it comes to more complex playing techniques such as chords, arpeggios, etc., more accurate hand tracking and gesture prediction are required. Therefore, this research method provides a probability-based way to recognize and predict gestures. For simple note playing, it may only need to press a key or several keys, or even include a series of action patterns to enable the system to recognize. Therefore, the integration of piano teaching and gesture recognition technology in VR provides new possibilities for improving teaching effects, enhancing students' learning experience, and providing effective teaching feedback for teachers. The first-order derivative relationship between the position of finger coordinates and time is shown in Equation (12).

$$\begin{cases} x_t' = \dfrac{dx}{dt} = m'(t) \\ y_t' = \dfrac{dy}{dt} = n'(t) \end{cases} \tag{12}$$

In Equation (12), x_t' is the movement rate of the finger on the *x*-axis. y_t' is the movement rate of the finger on the *y*-axis. A discrete data is collected at fixed equal time intervals. The time interval is very small, which can be replaced by difference. The direction of finger movement is shown in Equation (13).

$$\varphi(t) = \begin{cases} \tan^{-1}\left(\dfrac{y_t - y_{t-1}}{x_t - x_{t-1}}\right) + \pi, x_t - x_{t-1} < 0 \\ \tan^{-1}\left(\dfrac{y_t - y_{t-1}}{x_t - x_{t-1}}\right) + 2\pi, x_t - x_{t-1} < 0, y_t - y_{t-1} < 0 \end{cases} \tag{13}$$

In Equation (13), (t) is the direction of finger movement. The change in gesture direction angle and threshold change are shown in Equation (14).

$$\begin{cases} \Delta\varphi_t = \varphi_t - \varphi_{t-1} \\ \Delta\varphi_t < \tau, \text{Data retention} \\ \Delta\varphi_t > \tau, \text{Data filter} \end{cases} \tag{14}$$

In Equation (14), τ is the threshold, and the threshold is taken as 30°. The amount of change in the direction and angle of motion of the gesture is used as a preliminary data filter. When the change rate is less than 30°, it is retained, and when it is greater than 30°, it is filtered out. To further suppress high-frequency noise, the Simple Moving Average (SMA) filter method can also be used for data denoising, as shown in Equation (15).

$$S = \frac{\left(x_i + x_{i-1} + \ldots + x_{i-n+1}\right)}{n}, (n = 1, 2, 3, \ldots) \tag{15}$$

In Equation (15), n is the length of the data sequence. The size of the value directly affects the detailed information of the gesture. The effect schematic diagram of SMA is shown in Figure 3.

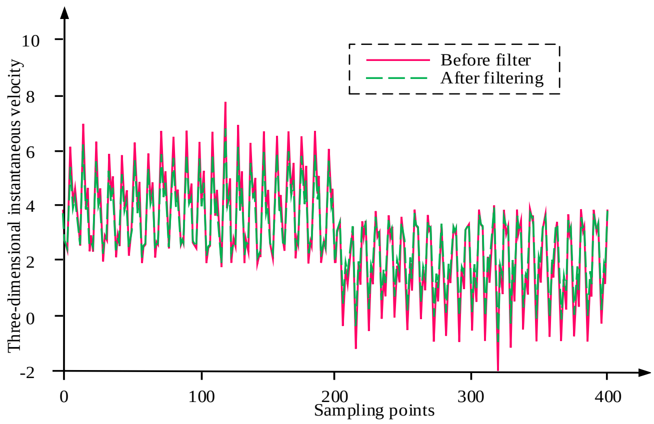


Figure 3. Schematic diagram of SMA smooth filtering effect.

In Figure 3, the Chebyshev low-pass filter successfully filters out excess noise, leaving behind low-frequency information. This can effectively reduce the frequency of ineffective work. Chebyshev low-pass filter successfully filters out excess noise and preserves low-frequency information through the excellent roll off characteristics. It can effectively reduce the frequency of ineffective work, improve system performance and efficiency. In the field of signal processing and electronic engineering, Chebyshev low-pass filter is an important tool with wide application. The blocking method of sliding windows is to divide a series of continuous action signals into several blocks. Then the action signals of each block are averaged. The average movement speed of the finger in the window is the obtained average value. If the average movement speed is less than a certain critical value, there are no piano performance actions during the window period. When the average speed is the maximum value, it means that there is an action. This time window exactly covers an action. The instantaneous convergence speed of the fingertips and the average convergence speed within the entire window are shown in Equation (16).

$$\begin{cases} v_i = \dfrac{\sqrt{\left(x_{i+1}-x_i\right)^2+\left(y_{i+1}-y_i\right)^2+\left(z_{i+1}-z_i\right)^2}}{\Delta t} \\ V_i = \arg\max_{i} \left[\dfrac{1}{k}\sum_{i}^{i+k-1} v_i\right] \end{cases} \tag{16}$$

In Equation (16), v_i is the 3D instantaneous velocity of the i-th sample point. k is the window width. The direction of finger movement best reflects the characteristics of the action. Firstly, based on the transformation relationship between spherical and Cartesian coordinate systems, the normalization and velocity vector are determined, as shown in Figure 4.

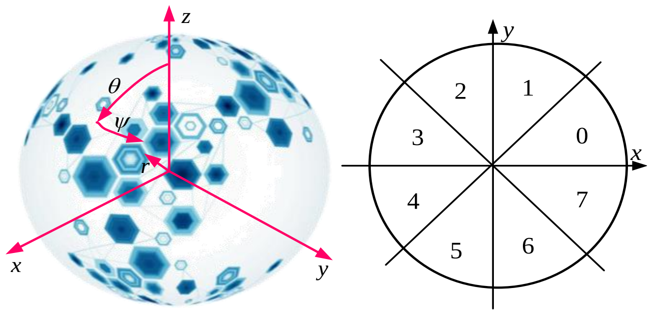


Figure 4. Schematic diagram of vector sphere.

In Figure 4, there is a transformation relationship between vector spherical coordinates and Cartesian coordinate systems. Through this transformation relationship, users can convert the speed direction of finger movement from vector spherical coordinates to Cartesian coordinate systems. Thus, users can better understand the characteristics of gesture movements and conduct relevant analysis. In VR, the speed and direction of finger movement can be seen as a feature of gesture. It can be used as input information to control objects in the virtual environment, Gesture recognition and gesture control. By collecting and analyzing the speed and direction of finger movement, the system can recognize different gesture movements. Based on these gesture actions, corresponding feedback and control are completed. The encoding of finger movement speed direction can also be applied in VR pianos. The speed direction of finger movement is recognized and encoded in real-time. The game can achieve a more realistic and accurate interactive experience based on the player's gestures and actions. It can realize more natural and intuitive Gesture recognition and interactive control, and improve the user experience and interactivity in the virtual environment. With the continuous progress of VR technology and the continuous enrichment of application scenarios, the application potential of finger motion velocity direction coding has also been further explored and expanded. The established HMM training and recognition flowchart is shown in Figure 5.

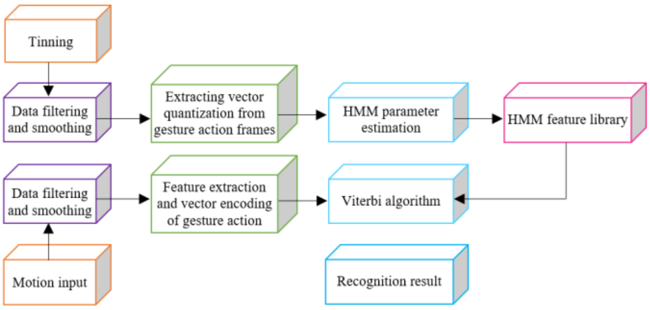


Figure 5. HMM training and recognition flowchart.

In Figure 5, the HMM method based on time series probability is proposed to be used. Multiple samples are taken for each pre-set posture behavior. By using methods such as filtering and smoothing, inter frame extraction and vector quantization encoding of pose behavior are achieved. Furthermore, an HMM model based on time series probability is established to achieve precise modeling of posture behavior. The initial probability, transition probability, observation probability, and other parameters are processed into observation sequences through vector code processing. The formulas in the Viterbi algorithm are used to solve various parameter terms until they finally converge.

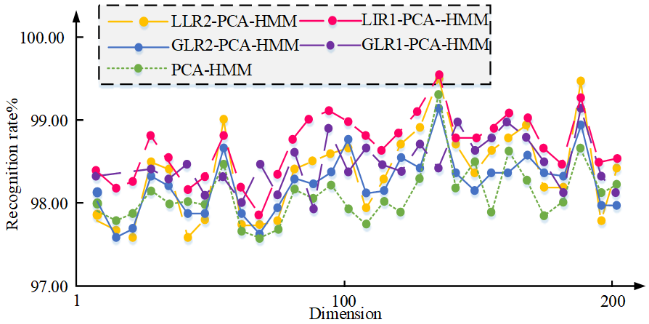
4. Application Analysis of Gesture Recognition Technology in VR Piano Playing Teaching Game based on HMM Model

4.1. Analysis of HMM Model Integrating VR Piano Performance

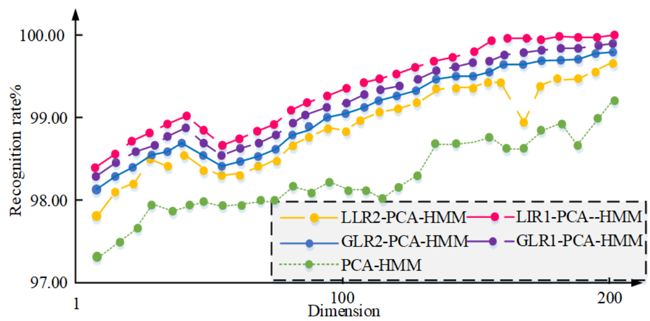
In the experiment, it involves cutting facial images. The window slides from top to bottom, with a certain overlap between adjacent windows. For the UCF101 database, the sampling window height L=11 and the overlapping window height P=7. For VGGFace/VGGFace2 databases, the sampling window height L=12 and the overlapping window height P=5. In the collection of gesture data, firstly, the gestures are preset, and then each gesture movement is subjected to repeated data collection, and the collected data are processed by smoothing and data filtering. Secondly, the different frames of the gesture are encoded and quantized for analysis, and then the currently collected gesture data are encoded and modeled by studying the algorithmic model used to solve the main parameters of the current model, the original probability, the movement probability, and the observation probability, and finally, the optimal arrangement probability of the current gesture action is calculated by studying the algorithmic model, and the maximum probability model output from the algorithmic model is the current gesture action. Experimental parameter Settings are shown in Table 2.

Table 2. Time synchronization error and program run time of different data sets.

Parameter	UCF101 database	VGGFace/VGGFace 2 database
Sampling window height (L)	11	12
Overlapping Window Height (P)	7	5
CPU Configuration	Intel i5	AMD Ryzen 5
RAM size	8 GB	16 GB
Graphics card	Nvidia GTX 1050	AMD Radeon RX 570
Operating system	Windows 10	Ubuntu 18.04
Version of VR system used	Oculus Rift	HTC Vive
Number of states in HMM	5	7
Type of emission probabilities	Gaussian	Multinomial



a) The accuracy of Piano key position at different dimensions on UCF101 database.



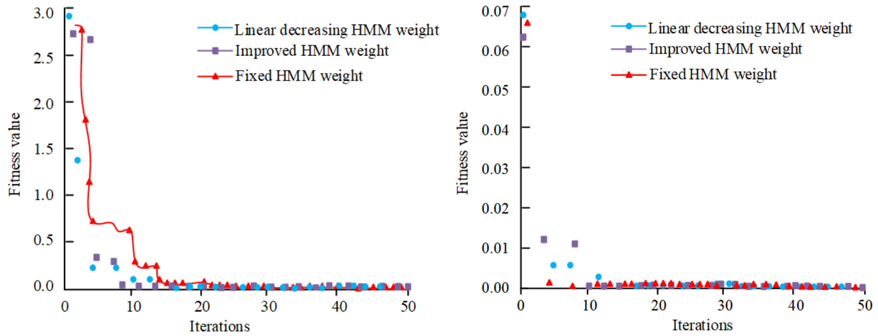
b) Piano key position accuracy at different dimensions in VGGFace/VGGFace2 database.

Figure 6. HMM models for various VR piano performances in different dimensions.

In Table 2, Gesture recognition plays a crucial role in VR piano teaching games, which can analyze user gestures in real-time and convert them into keyboard operations in a virtual environment. To address the temporal nature of gesture sequences, Hidden Markov Models (HMMs) are applied to gesture recognition. The performance analysis experiment of the HMM based gesture recognition model used UCF101 and VGGFace/VGGFace2 databases. UCF101 provides diverse human action videos for model training and testing, while VGGFace and VGGFace2 provide a large number of facial images, helping to achieve user authentication and personalized settings in VR environments. The purpose of the experiment is to verify the application effect of the HMM gesture recognition model in VR piano teaching games, and to explore optimizing the model and algorithm to improve its performance, in order to promote the development of VR piano teaching games and meet users' needs for efficient, accurate, and natural interactive experiences. HMM models for various VR piano performances in different dimensions are shown in Figure 6.

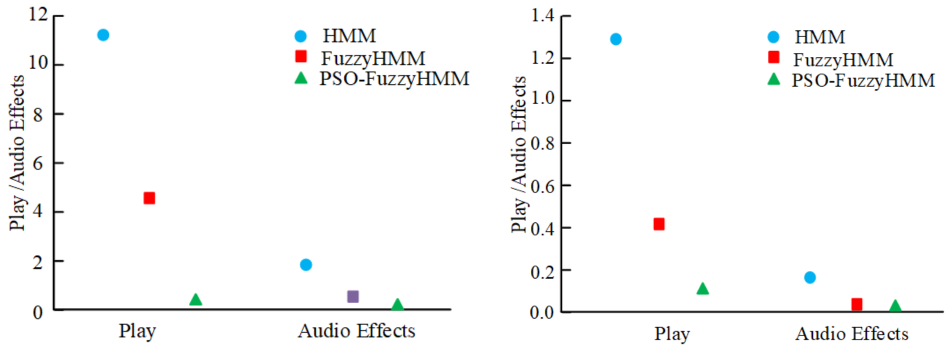
Figure 6. HMM models for various VR piano performances in different dimensions. Firstly, in Figure 8-a), as the dimension increases, the performance of various algorithms gradually stabilizes. Among them, the Log-Likelihood Ratio (LLR2) algorithm performs the best among all algorithms, while the Principal Component Analysis-HMM (PCA-HMM) algorithm performs the worst. The experimental results show that compared with the overall Gas Lantern Routine (GLR) method, blocky LLR can better improve the classification performance of images. Compared with other correction methods for performance, LLR2 is more suitable for VR pianos because it uses more virtual training samples. Therefore, higher recognition accuracy can be achieved. Secondly, the identification accuracy of the PCA-HMM model without performance correction is relatively low. Compared with the LLR2 model, the identification accuracy is lower, only 97.02%, indicating the necessity of performance correction. In Figure 8-b), the accuracy of the performance correction algorithm can be further improved by using performance correction. The recognition accuracy of keyboard and finger touch accuracy algorithm in piano teaching and professional performance reaches 99.685%. Meanwhile, due to the addition of virtual training samples, the performance of this method is superior to that of general single training sample performance correction methods. Therefore, it can be more suitable for playing change recognition problems. AVR piano performance parameter is designed based on the HMM model. The HMM online control model achieves fast, accurate, and stable online control of process parameters such as playing and music accuracy. Finally, the performance effect of the model is compared and evaluated through simulation results. Iterative curves based on three HMM algorithms are shown in Figure 7.

In Figure 7, compared to linearly decreasing HMM weights and fixed HMM weights, the improved HMM weight Ackley function converges at the 26th and 13th iterations. The convergence rate is 4 times and 3 times that of the other two algorithms. The improved HMM weight Griewank function falls into local optimal fitness values in the 3rd and 8th times. The improved dynamic HMM weight algorithm is significantly superior to the other two algorithms in terms of iterative convergence efficiency and avoiding falling into local optima. The comparison of performance and audio effects before and after decoupling is shown in Figure 8.

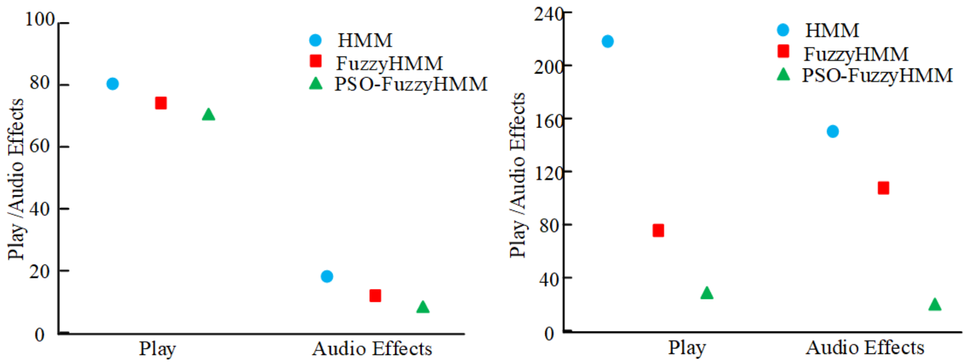


a) Ackley function. b) Griewank function.

Figure 7. Iterative curve based on three HMM algorithms.



a) Comparison chart of overshoot performance. b) Steady state error performance comparison chart.



c) Response time performance comparison chart. d) Comparison chart of adjust time performance.

Figure 8. Performance comparison under three different algorithms.

The experimental results obtained based on Python software are shown in Figure 9. The results show that the overshoot of performance and audio effect curves is reduced by 96.08% and 89.48% compared to ordinary HMMs, respectively. The steady-state error is reduced by 1.24 and 0.14. The response time is shortened by 12.4% and 59.98%, respectively. The adjustment time is shortened by 86.37% and 80.01%, respectively. Compared to the fuzzy HMM control model, the overshoot of the performance and audio effect curves is reduced by 4.14 and 0.54, respectively. The steady-state error is decreased by 88.09% and 58.32% respectively. The response time is reduced by 5.40% and 33.2%, respectively. The adjustment time is shortened by 60.53% and 81.47%, respectively. The one-time performance deviation and audio effect deviation are 0.44 and 0.21, respectively. For the three algorithms of conventional HMM control, fuzzy HMM control, and improved Particle Swarm Optimization (PSO) optimized fuzzy HMM control, the improved PSO optimized fuzzy HMM algorithm performs the best in both performance and audio effects. The one-time performance deviation and audio effect deviation are 0.46 and 0.18, respectively, which are 96.05% and 89.48% lower than traditional HMM algorithms. Compared to the fuzzy HMM algorithm, it reduces by 90.20% and 73.35%, respectively. The conventional HMM algorithm and fuzzy HMM algorithm have rapidity, stability, and accuracy in online control of piano performance and audio effects.

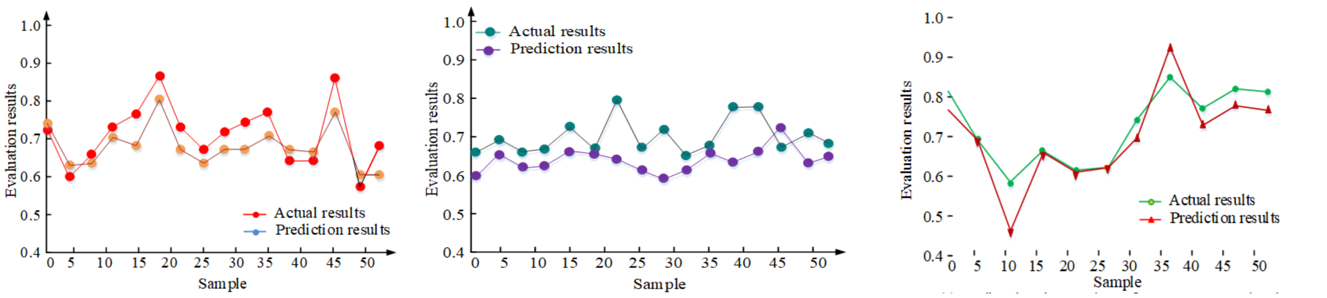
4.2. VR Piano Performance Gesture Recognition Test

In the VR piano playing gesture recognition test, basic piano playing gestures are investigated and the performance of different gesture recognition algorithms is evaluated. 440 participants were tested in 22 VR scenes. Of these, 220 participants wore sensors equipped with different algorithms, and the data collected show only some of the main data, as shown in Table 3.

In Table 3, the mean absolute error of Truncated Singular Value Decomposition (TSVD) on each training set is 0.26s and the standard deviation is 0.204s. The mean error and standard deviation of Simultaneous Localization and Mapping (SLAM) are 13.17s and 5.676s, respectively. This shows that TSVD is superior to SLAM in both accuracy and stability. In terms of processing time, the average processing time of TSVD was 0.059s, while the average processing time of SLAM was 0.833s. This further confirms that TSVD is superior to SLAM in processing efficiency. In addition, the study compares the fitness of different algorithms with the HMM algorithm, and the results are shown in Figure 10.

Table 3. Time synchronization error and program run time of different data sets.

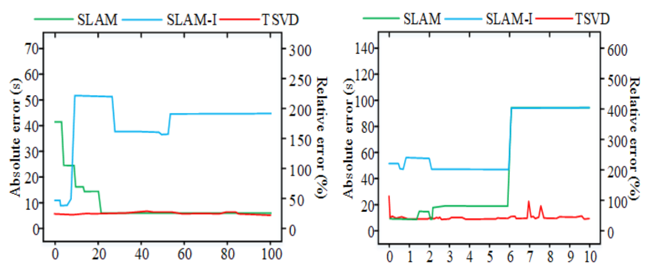
Test object	Absolute error(s)			Program run time(s)		
	SLAM	SLAM-I	TSVD	SLAM	SLAM-I	TSVD
Dataset 1	11.4	10.5	0.0	1.234	0.331	0.042
Dataset 2	4.3	6.7	0.3	0.234	0.139	0.052
Dataset 3	9.3	7.8	0.4	0.573	0.243	0.021
Dataset 4	13.4	11.4	0.3	1.331	0.471	0.072
Dataset 5	16.0	14.9	0.0	0.962	0.142	0.061
Dataset 6	5.8	7.2	0.8	0.784	0.098	0.045
Dataset 7	21.7	12.4	0.6	1.365	0.377	0.087
Dataset 8	19.7	14.5	0.6	1.197	0.278	0.072
Dataset 9	17.3	11.3	0.1	0.832	0.181	0.061
Dataset 10	14.2	8.5	0.2	0.913	0.059	0.043
Dataset 11	12.2	10.5	0.0	1.234	0.361	0.052
Dataset 12	9.7	8.7	0.6	0.234	0.113	0.059
Dataset 13	10.3	7.7	0.7	0.561	0.235	0.064
Dataset 14	13.7	17.7	0.4	1.311	0.453	0.043
Dataset 15	16.1	14.2	0.8	0.982	0.107	0.043
Dataset 16	5.3	7.1	0.1	0.737	0.093	0.056
Dataset 17	21.7	16.7	0.4	1.347	0.382	0.064
Dataset 18	19.7	13.4	0.7	1.182	0.294	0.055
Dataset 19	17.4	18.7	0.4	0.821	0.198	0.057
Dataset 20	11.1	8.7	0.3	0.911	0.087	0.026
Average	13.17	10.57	0.28	0.937	0.231	0.049
Standard deviation	5.678	2.927	0.209	0.367	0.152	0.016



a) Predicated and true values of HMM algorithm. b) Predicated and true values of FuzzyHMM algorithm. c) Predicated and true values of POS-FuzzyHMM algorithm.

Figure 10. Comparison of adaptability of HMM algorithms using different algorithms.

In Figure 10, the fit of the three methods is compared with the actual fit. The predicted and true values of the three algorithms are compared with their fitness. The PSO-FuzzyHMM algorithm model shows that as the number of samples increases, the curve changes between the true and predicted values are basically in the same curve, with a slight difference in numerical values. In the analysis of the changes in the true and predicted curve values, when the sample size reaches 67, the PSO-FuzzyHMM algorithm model has the maximum deviation between the true and expected values. The overall deviation is about 0.05. The fitness of the PSO-FuzzyHMM algorithm model was calculated to be 94.8% by comparing the curves. The HMM algorithm has a significant deviation between the true and predicted values, with a fitness of 88.6%. The RNN algorithm model has a small deviation in the trend of changes between the true and predicted values, with a fitness of 91.5%. The fitness of the PSO-FuzzyHMM algorithm model is 6.8% higher than that of HMM, and 2.6% higher than that of the FuzzyHMM algorithm model. By experiencing the time error of VR, the algorithm performance is verified. The synchronization error of VR experience time under different preset parameters is shown in Figure 11.



a) Window size. b) Thershold(m/s^2).

Figure 11. Synchronization error of VR experience time algorithm under different preset parameters.

Figure 11-a) shows the error situation for different window sizes. The experimental results show that as the window width increases, the time synchronization accuracy of TSVD reaches the best, and the error is always between 0-1s. The SLAM method has the lowest

time synchronization accuracy. As the window width increases, the accuracy tends to stabilize between 59 and 60s. The SLAM-I method has low accuracy when the window width is small. But as the window width increases, the accuracy is equivalent to that of the TSVD method. In Figure 11-b), error scenarios under different thresholds are shown. The results indicate that the time stability of this method is the best. It always remains relatively stable, with only occasional fluctuations. The accuracy of the Zero-Velocity Update (ZUPI) algorithm gradually improves with the increase of the threshold, and finally stabilizes at about 140s. The error of SLAM-I method also increases with the increase of threshold. However, in the early stages of the threshold, it falls between the two methods. In later stages, it is similar to both methods. The flash recognition experiment is shown in Figure 12.

In Figure 13, the horizontal coordinate is the user ID, and the vertical coordinate is the rating given by the user, from low to high, in order of 1-10. The highest rating for utility was given by user 3, who gave it an 8. In the rating of mastery, that is, friendliness to novices, user 1 thinks it is easy, while users 3 and 5 think it is slightly difficult. In the rating of interest level, users 1, 3 and 6 all think that they are very interested in this project, while user 4 thinks that he is not interested in this project. Different research algorithm models such as Convolutional Neural Network (CNN) algorithm, Support Vector Machine (SVM) algorithm and Scale Invariant Feature Transform (SIFT) algorithm are compared and analyzed with the current algorithm models as shown in Table 4.

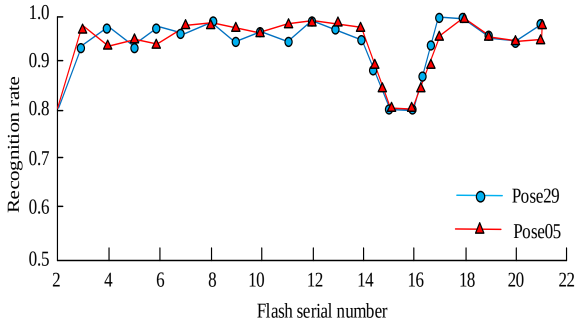


Figure 12. Recognition results for the subset of ‘illum’.

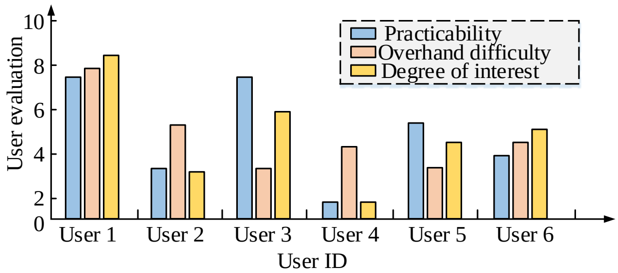


Figure 13. User evaluation model.

Table 4. Comparison of performance parameters of different algorithms.

Test frequency	CNN			SVM			SIFT			Research on using algorithms		
	F1 (%)	RMSE	Accuracy (%)	F1 (%)	RMSE	Accuracy (%)	F1 (%)	RMSE	Accuracy (%)	F1 (%)	RMSE	Accuracy (%)
1	81.1	8.1	82.6	83.7	7.1	85.6	86.8	6.7	92.5	92.9	3.7	96.4
2	82.4	9.4	84.6	82.7	8.2	88.2	88.1	5.8	93.6	93.4	3.6	97.0
3	81.9	7.9	83.4	83.5	8.8	87.5	83.6	6.5	91. 2	92.8	3.5	96.8
4	81.1	8.1	82.5	84.3	8.6	86.2	89.5	5.7	90.3	92.2	4.7	97.1

As can be seen in Table 4, in the comparison of the detection performance of several algorithmic models research using algorithmic model of several parameter values significantly higher than other algorithmic models, the error is significantly lower. Among them, the lowest F1 value of the research use model is 92.25, which will be compared to the worst CNN model, whose model F1 value is 9.8% higher than its highest value. Meanwhile comparing the Root Mean Square Error (RMSE) value in the research using algorithm model the parameter value of the lowest minimum value is 3.5, compared to the highest CNN model its error value is lower by 5.9,. Meanwhile the accuracy value of the model used in the study is 97.1% higher in the accuracy comparison. This shows that the accuracy of the model used in the study is significantly higher than the other traditional algorithmic models in the comparison of accuracy and other parameters. In order to verify the real experience effect of the current research model, 100 college students in the user experience survey, the survey experience through the form of the form to mention, will be analyzed by its experience survey is shown in Table 5.

As can be seen in Table 5, in the users' evaluation of the analysis of the real situation, 92% of the users think that the model has a better sense of realism, and 8% of the users think that the model does not have the effect of real experience. In the evaluation of teaching help, 89% of the students thought that they could do better teaching assignments on the piano through this model, and 11% of the students thought that it did not help piano teaching. Finally, 94% of the students thought that the current model could be more helpful to their learning, while 6% of the students thought that the model was not helpful to their learning. This shows that the model can effectively help piano teaching.

Table 5. Analysis of user experience.

Experience indicators	Number of evaluators	Evaluation proportion	Total proportion
Realistic model	100	92	92%
The model is not realistic	100	8	8%
Helpful for teaching	100	89	89%
Not helpful for teaching	100	11	11%
Helpful for learning	100	94	94%
Not helpful for learning	100	6	6%

5. Conclusions

Through the research of hand gesture recognition technology in VR piano teaching game based on HMM, the purpose is to improve the accuracy and efficiency of piano teaching and professional performance. The research methods include the application of keyboard and finger touch accuracy algorithms, as well as a modified Griewank function for HMM weights and a modified dynamic HMM weighting algorithm. The results show that the recognition accuracy of keyboard and finger touch accuracy algorithm in piano teaching and professional performance reaches 99.685%. The modified Griewank function of HMM weights is trapped in the fitness value of the local optimum at the third and eighth times. In terms of iterative convergence efficiency and avoiding falling into local optimum, the improved dynamic HMM weight algorithm is significantly better than the other two algorithms. The overshoot of performance curve and audio effect curve are reduced by 96.08% and 89.48% respectively compared with ordinary HMM. The mean absolute error and standard deviation of TSVD on each training set are 0.26s and 0.204s, respectively. By comparing the curves, the fitness of PSO-FuzzyHMM algorithm model is calculated to be 94.8%. There is a significant deviation between the true and predicted values of the HMM algorithm with a fitness of 88.6%. The research contribution is to improve the accuracy and efficiency of piano teaching, and provide new technical support for piano teaching games. However, there are some differences in the degree of interest of some users in the study in this project, which may affect the universality and generalization of the project. Future research and optimization will focus on how to further improve the recognition accuracy and stability of the algorithm, and how to better meet the needs and interests of different users, so that VR piano teaching games can be applied in a wider range of fields. In the subsequent research will not only piano gesture teaching, but also apply it to the teaching of other musical instruments, and secondly, the research cost of the software built by the research is still high, and the subsequent research will also control the cost of the research to find a cheaper virtual simulation software. Meanwhile, in the process of constructing the HMM model, in order to improve its probability accuracy of gesture recognition, it can improve the recognition accuracy of the model by classifying and solving different data in the process of analyzing the data.

References

[1] Albahri O. and AlAmoodi A., "Navigating the Metaverse of Big Data: A Bibliometric Journey," *Mesopotamian Journal of Big Data*, vol. 2023, pp. 92-106, 2023. https://doi.org/10.58496/MJBD/2023/013

[2] Amaya M., Linde N., and Laloy E., "Adaptive Sequential Monte Carlo for Posterior Inference and Model Selection among Complex Geological Priors," *Geophysical Journal International*, vol. 226, no. 2, pp. 1220-1238, 2021. https://doi.org/10.1093/gji/ggab170

[3] Briggs A., Robstad C., and Rosell F., "Using Radio-Frequency Identification Technology to Monitor Eurasian Beavers," *Wildlife Society Bulletin*, vol. 45, no. 1, pp. 154-161, 2021. https://doi.org/10.1002/wsb.1147

[4] Cui K., "Artificial Intelligence and Creativity: Piano Teaching with Augmented Reality Applications," *Interactive Learning Environments*, vol. 31, no. 10, pp. 7017-7028, 2023. https://doi.org/10.1080/10494820.2022.2059520

[5] De Wael A., De Backer A., and Van Aert S., "Hidden Markov Model for Atom-Counting from Sequential ADF STEM Images: Methodology, Possibilities and Limitations," *Ultramicroscopy*, vol. 219, no. 1, pp. 1-14, 2020. https://doi.org/10.1016/j.ultramic.2020.113131

[6] Deja J., Mayer S., Pucihar K., and Kljun M., "A Survey of Augmented Piano Prototypes: Has Augmentation Improved Learning Experiences," *Proceedings of the ACM on Human-Computer Interaction*, vol. 6, no. ISS, pp. 226-253, 2022. https://doi.org/10.1145/3567719

[7] Deng H., Shang W., Lu G., Guo P., Ai T., Fang C., and Tian J., "Targeted and Multifunctional Technology for Identification between Hepatocellular Carcinoma and Liver Cirrhosis," *ACS Applied Materials and Interfaces*, vol. 11, no. 16, pp. 144526-14537, 2019. DOI:10.1021/acsami.8b20600

[8] Don M. and Khan F., "Process Fault Prognosis Using Hidden Markov Model-Bayesian Networks (HMM-BN) Hybrid Model," *Industrial and Engineering Chemistry Research*, vol. 58, no. 27, pp. 12041-12053, 2019. DOI:10.1021/acs.iecr.9b00524

[9] Frimane A., Munkhammar J., and Van der Meer D., "Infinite Hidden Markov Model for Short-Term Solar Irradiance Forecasting," *Solar Energy*, vol. 244, no. 9, pp. 331-342, 2022. DOI:10.1016/j.solener.2022.08.041

[10] Goodall V., Ferreira S., Funston P., and Maruping-Mzileni N., "Uncovering Hidden States in African Lion Movement Data Using Hidden Markov Models," *Wildlife Research*, vol. 46, no. 4, pp. 296-303, 2019. DOI:10.1071/WR18004

[11] Jiang J., Chen R., Chen M., Wang W., and Zhang C, "Dynamic Fault Prediction of Power Transformers Based on Hidden Markov Model of Dissolved Gases Analysis," *IEEE Transactions on Power Delivery*, vol. 34, no. 4, pp. 1393-1400, 2019. DOI:10.1109/TPWRD.2019.2900543

[12] Li Y., Li X., Chen Z., Ma Y., Yang G., Zhang Y., Sun M., Zhang L., Chu J., Lu J., Zhu J., and You J., "Application of Plant Print Identification Technology in Salt-Resistant Soybean Breeding," *Journal of Plant Growth Regulation*, vol. 40, no. 4, pp. 1557-1565, 2021. DOI:10.1007/s00344-020-10213-1

[13] Li Y., Wu M., Zhang Y., Xu L., and Yu L., "Piano: A Parametric Hand Bone Model from Magnetic Resonance Imaging," *in Proceedings of the 30th International Joint Conference on Artificial Intelligence*, Montreal, pp. 816-822, 2021. https://doi.org/10.24963/ijcai.2021/113

[14] Patel S. and Makwana R., "Connectionist Temporal Classification Model for Dynamic Hand Gesture Recognition Using RGB and Optical flow Data," *The International Arab Journal of Information Technology*, vol. 17, no. 4, pp. 497-506, 2020. https://doi.org/10.34028/iajit/17/4/8

[15] Raskar C. and Nema S., "Metaheuristic Enabled Modified Hidden Markov Model for Traffic Flow Prediction," *Computer Networks*, vol. 206, no. 7, pp. 1-11, 2022. DOI:10.1016/j.comnet.2022.108780

[16] Rosenblatt J., Ritov Y., and Goeman J., "Discussion of 'Gene Hunting with Hidden Markov Model Knockoffs'," *Biometrika*, vol. 106, no. 1, pp. 29-33, 2019. DOI:10.1093/biomet/asy062

[17] Shin Y., Liu D., Sang H., Ferguson T., and Song P., "A Binary Hidden Markov Model on Spatial Network for Amyotrophic Lateral Sclerosis Disease Spreading Pattern Analysis," *Statistics in Medicine*, vol. 40, no. 13, pp. 3035-3052, 2021. DOI:10.1002/sim.8956

[18] Soper B., Nygrd M., Abdulla G., Meng R., and Nygard J., "A Hidden Markov Model for population Level Cervical Cancer Screening Data," *Statistics in Medicine*, vol. 39, no. 25, pp. 3569-3590, 2020. DOI:10.1002/sim.8681

[19] Stolz B., "The Author and the Piano Student: The Transferability of the Creative Process in Practice," *International Journal of Music Education*, vol. 40, no. 10, pp. 53-65, 2022. DOI:10.1177/02557614211019144

[20] Tallon M., Greenlee M., Wagner E., Rakoczy K., and Frick U., "How do Art Skills Influence Visual Search?-Eye Movements Analyzed with Hidden Markov Models," *Frontiers in Psychology*, vol. 12, pp. 1-17, 2021. https://doi.org/10.3389/fpsyg.2021.594248

[21] Xia T. and Chen X., "A Weighted Feature Enhanced Hidden Markov Model for Spam SMS Filtering," *Neurocomputing*, vol. 444, no. 15, pp. 48-52, 2021. DOI:10.1016/j.neucom.2021.02.075

[22] Yildiz A., Adar N., and Mert A., "Convolutional Neural Network Based Hand Gesture Recognition in Sophisticated Background for Humanoid Robot Control," *The International Arab Journal of Information Technology*, vol. 20, no. 3, pp. 368-375, 2023. https://doi.org/10.34028/iajit/20/3/9

[23] Zhiyong G., Jiwu L., and Rongxi W., "Prognostics Uncertainty Reduction by Right-Time Prediction of Remaining Useful Life Based on Hidden Markov Model and Proportional Hazard Model," *Eksploatacja i Niezawodność-Maintenance and Reliability*, vol. 23, no. 1, pp. 154-164, 2020. DOI:10.17531/ein.2021.1.16

Min Zeng 1982.12.19, Han, male, Changsha, Hunan Province, 2001.9-2004.7 Hunan City College of vocal specialized specialty. 2004.9-2007.7 Bachelor's degree in vocal performance, China Conservatory of Music. 2009.9-2012.7 Master's degree in Vocal Performance, Wuhan Conservatory of Music. 2021.9-2024.7 Ph.D. in Music Education, College of Music, Hunan Normal University. Work experience: 2007.9-2021.7 Hunan Vocational College of Science and Technology, Faculty of Humanities and Arts, vocal teacher. Academic situation: academic paper "Russian vocal teacher Su Shilin research" published in the May 2024 issue of the total No. 358.