J. Electrical Systems 20-4s (2024): 404-415

[1]Hong Jiang

Machine Learning in Teaching Piano Touch Technique and Tonal Expression

Journal of Electrical Systems

Abstract: - Music is the art of both sound and hearing. Among many musical instruments, the piano has the title of "king of instruments". Unlike the violin, which is delicate and soft, the piano's tone is noble and graceful. The pianist Nergoz once said that the purpose of all piano playing is to make sound and to create sound. So, how to use the piano to create a beautiful sound? This is an essential concern for all piano players. Therefore, it is important to practice tone, both in practice and in playing music, and only when the fingers touch the keys can a beautiful tone be produced. Tone is important in the whole process of piano playing. Both the expressiveness of the piano and its appeal to the listener are closely related to the quality of the tone. The part of the key that is touched, the intensity and the technique all have an important influence on the tone. The scientific approach to finger touch is a particularly important skill and technical criterion, both in the process of learning the piano and when performing it.

Keywords: performance; piano; tonal expression; key touch; machine learning

I. INTRODUCTION

People are more eager to meet higher spiritual needs after satisfying the basic needs of life, such as clothing, food, housing and transportation[1-2]. Learning to play the piano has become more and more popular among people[3]. In recent years, the number of people who want to learn to play piano has increased dramatically, and there are people of all ages who want to learn to play piano, with the largest number of young children learning to play piano[4]. In urban families of medium and above, one in five families has parents who want their children to learn piano, and the number of piano grade examination applicants is growing at a rate of 20% year by year[5]. On the other hand, the reform of the national education mechanism and the expansion of the scale of education have led to a significant increase in the demand for piano teachers. In the piano education industry, professional piano teachers have become a very scarce resource, and this state of affairs will not be greatly improved in a short time[6].

Piano education is very expensive, costing at least $10,000 for an average piano. Piano lessons are taught on a one-on-one basis, and the tuition for each piano lesson is between $100 and $400[7]. In this case, most of the time students learn piano is only one piano lesson a week, and the rest of the time students can only practice blindly by themselves, so the process of learning to play piano takes many detours, and the learning progress is very slow[8]. The learning of piano performance is also limited by the geographical and time constraints, which makes many enthusiasts who have the intention to learn to play piano to give up the learning.

It is a common problem when learning to play with excess motion. It is important to reduce the external forces as much as possible, to make the playing movements simple enough, and to improve the stability of the fingers in order

[1] Department of Preschool Education, Chaoyang teachers college, Chaoyang 122000, China

Email：jhmalaysia@163.com

to produce a focused, powerful and bright tone. In order to play a consistent tone on a horizontal line, it is important to avoid up and down wrist tremors as much as possible, to pay more attention to the octave, and to minimize the range of motion when practicing fast, and to make the motion as smooth as possible[9]. Whenever a note is played, it should be quickly released, and the power should be quickly transferred to the next note, especially when it comes to faster groups of notes, the power should be transferred between the fingers in a very relaxed way. It is also important to make sure that there is no deliberate pressure on the wrist. This will ensure that the weight of the hand always follows the movement of the fingertips. To ensure a smooth, even tone, the player's arm, elbow, and wrist must follow the fingertip movement well.

In addition, the tone is influenced by the support of the fingers. The key to good piano playing is the support of the fingers, which directly affects the technique and thus the quality of the tone. Mr. Ling Zhuping has said that the support of the fingers is limited and once the pressure of the arm is higher than the support of the fingers, it is easy to cause damage to the muscles of the hand, but the fingers should not be placed loosely on the piano, if the fingers do not feel it and are loose, the tone effect will be poor[10]. When playing the climax of the fourth variation of Glinka's "The Nightingale", the left hand has to play a skipping eighth-note decomposed chord and the right hand has to play an octave sixteenth-note decomposed chord, which requires a high degree of finger support[11][12]. In addition, it is important to feel the shoulder and the fingertips as two points of support when playing, although the palm is empty in the middle, but the palm of the hand can also play the role of support when playing[13]. No matter what kind of tone or speed you encounter, your fingers will always have a good support. This shows that the stability and support of the fingers play a particularly important role when playing piano works.

In piano playing, the keystrokes have a direct impact on the quality of the tone and the amount of tonal variation. If the touch is not good, the tone will not be good, and the variety of touches will be reduced. The piano is a very complex instrument, which requires not only a high level of technical skill, but also a complete understanding of the work and a full expression of the ideas and emotions contained in it[14]. The piano has a wide range of tones, and it is very important that the piano touches the keys in the right way to play the different levels of tone. This is because the tone of the piano is directly influenced by the explosive power of the fingertips, the stability, flexibility and independence of the fingers. Thus, the correct touch is a very important aspect of piano playing.

II. METHODOLOGY

To explore the link between piano touch technique and tonal expressiveness, we need to first analyze it with respect to the playing function, the block diagram of which is shown below.

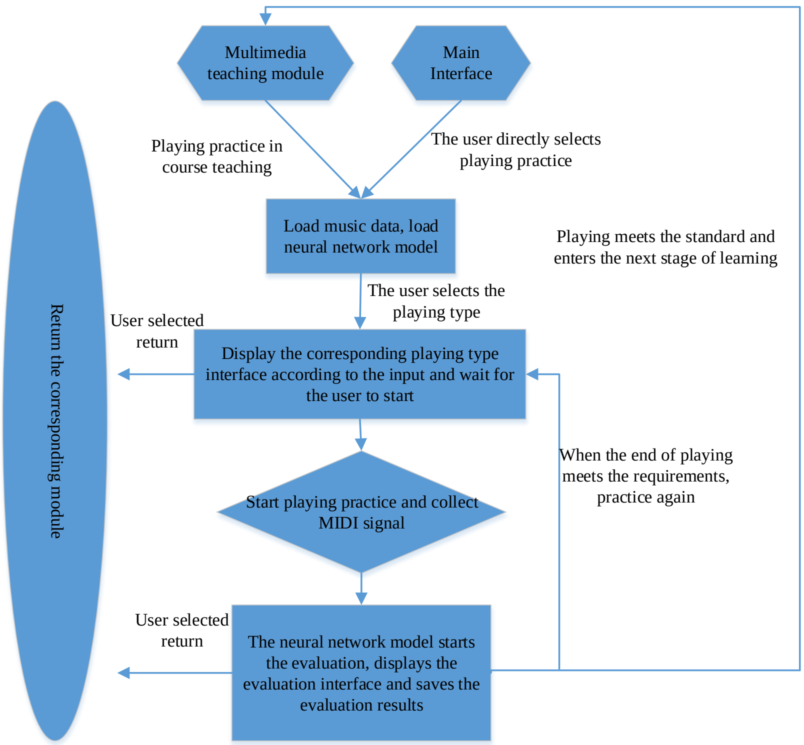


Figure 1 Block diagram of pop-up function analysis

When the object of Hand Controller is tracked through on Frame() callback function to get the data frame of each part of the object, the data frame structure is shown in Figure 2.

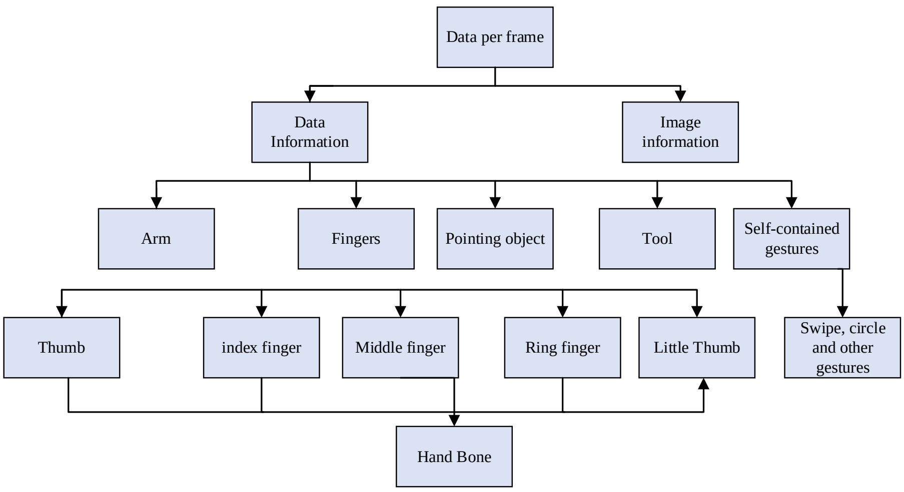


Figure 2 Data frame structure

With the function of tracking finger position with high accuracy, some high frequency noise will be generated when detecting motion gestures due to the tiny trembling of the finger or the sensor itself, etc. Therefore, data filtering and smoothing are required.

The trajectory of the finger in the process of movement can be approximated as a curve, and the relationship between the position of the coordinates of the finger and time can be expressed by the following function.

$$\begin{cases} x_t = m(t) \\ y_t = n(t) \end{cases} \quad (1)$$

Taking the first order derivative of equation (1) yields

$$\begin{cases} x_t^{'} = \dfrac{dx}{dt} m^{'}(t) \\ y_t^{'} = \dfrac{dy}{dt} n^{'}(t) \end{cases} \quad (2)$$

Equation (2) reflects the finger motion rate in the Z, J axes, from which the finger motion angle can be derived.

Since one discrete data is collected at fixed equal time intervals and the time interval is small, the differential in equation (2) can be replaced by the difference. From this, the direction of finger motion can be found as follows.

$$\varphi(t) = \begin{cases} \tan^{-1}\left(\dfrac{y_t - y_{t-1}}{x_t - x_{t-1}}\right) + \pi, x_t - x_{t-1} < 0 \\ \tan^{-1}\left(\dfrac{y_t - y_{t-1}}{x_t - x_{t-1}}\right) + 2\pi, x_t - x_{t-1} < 0, y_t - y_{t-1} < 0 \\ \tan^{-1}\left(\dfrac{y_t - y_{t-1}}{x_t - x_{t-1}}\right) (other) \end{cases} \quad (3)$$

The amount of angular change in the direction of the motion gesture between two adjacent two sampling points can thus be expressed by (4).

$$\Delta\varphi_t = \varphi_t - \varphi_{t-1} \quad (4)$$

Set a threshold τ , when τ is less than the amount of angular change in that direction then the data is retained, if greater than the amount of change is suspected of jitter noise, filter out the data, as shown in the following equation.

$$\begin{cases} \Delta\varphi_i < \tau, \text{ Data Retention} \\ \Delta\varphi_i > \tau, \text{ Data Filtering} \end{cases} \quad (5)$$

According to repeated experiments, the threshold value of equation (5) is taken as 30 degrees in this paper, i.e., when the amount of angular change in the direction of motion of the gesture is used as the initial data filter, the rate of change is retained when it is less than 30 degrees and filtered out when it is greater than 30 degrees.

In order to further suppress high frequency noise, the SMA (Simple Moving Average) filter method can also be used for data noise reduction processing. $S = \dfrac{\left(x_t + x_{t-1} + \ldots + x_{t-n+1}\right)}{n}, (n = 1, 2, 3, \ldots\ldots)$ (6)

(6) The S in Eq. is the length of the data sequence, and the value of n will affect the smoothing effect. If the value

is too large, the gesture detail information will be lost, and if it is too small, the smoothing effect is too poor.

Hidden Markov Models (HMM, Hidden Markov Model) can be used in labeling problems, in neighborhoods such as speech recognition, NLP, bioinformatics, pattern recognition, and nowadays popular artificial intelligence fields, where Hidden Markov Models play an increasingly important role.

The HMM is determined,

$$\lambda = (A, B, \pi), 2TN^2T$$
$$P(Q \mid \lambda) = \pi_{q_1} \alpha_{q_1 q_2} \alpha_{q_2 q_3}, \ldots, \alpha_{q_{r-1} q_r} Q \quad (7)$$

HMM not only contains the above three important elements in the learning process, but also needs to solve the learning problem and the prediction problem.

The formula for the output probability in the state sequence $Q = q_1, q_2, \ldots, q_r$ in the Hidden Markov given model λ is as follows.

$$P(O \mid \lambda) = \pi_{q_1} \alpha_{q_1 q_2} \alpha_{q_2 q_3}, \ldots, \alpha_{q_{r-1} q_r} \quad (12)$$

The probability formula for the observation sequence generated is

$$P(O \mid Q, \lambda) = \prod_{t=1}^{T} P\left(o_t \mid q_t, \lambda\right) = b_{q_1}\left(o_1\right) b_{q_2}\left(o_2\right) \ldots b_{q_r}\left(o_r\right) \quad (13)$$

Q in a given model λ can be derived from the above two equations.

$$P(O \mid \lambda) = \sum_{q_{11} q_{12} \cdots q_{i\pi}} \pi_{q_{11}} b_{q_{11}}\left(o_1\right) a_{q_{1,1} q_{12}} b_{q_{12}}\left(o_2\right) a_{q_{12} q_{13}} \ldots a_{q_{1(\tau-1)} q_{i\pi}} b_{q_\pi}\left(o_T\right) \quad (14)$$

The result of Equation (14) illustrates $P(Q \mid \lambda)$ contains all possible scenarios that can arise from the sequence of implied states of the observation sequence 0. The complexity of the calculation is about $2TN^T$, the HMM and belong to the dynamic planning algorithm and reduces the weight of complexity in the calculation to N^2T.

(1) Forward algorithm

Given model λ, the probability that the sequence of observations up to moment t is partially $o_1, o_2, \ldots, o_r$ and the state is q_i is defined as the forward probability and is denoted as

$$MQ(\lambda, \overline{\lambda}) = \sum_{I} \log P(O \mid I, \lambda) P(O \mid I, \overline{\lambda}) \quad (12)$$

(2) Backward algorithm

Given model λ, the observations from time series t +1 to T is $o_{t+1}, o_{t+2}, \cdots, o_T$ is defined as the backward probability given that the state at moment t is q_i.

$$P(O|\ \lambda) = \sum_{i=1}^{N} \pi_i b_i\left(o_1\right) \beta_1(i) \quad (22)$$

The relationship between the forward and backward probabilities is shown in Figure 3.

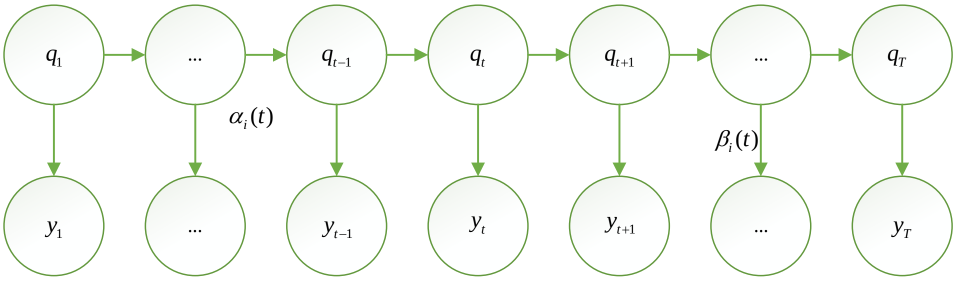


Figure 3 Relationship between forward and backward probabilities

If the training data is only the sequence of observations, the learning of HMM is required to use EM algorithm belongs to unsupervised learning.

(1) To find the parameter π_i, the first term of equation (25) can be written as

$$\sum_{I} \log \pi_{t_0} P(O, I|\ \bar{\lambda}) = \sum_{i=1}^{N} \log \pi_r P\left(O, i_1 = i|\ \bar{\lambda}\right) (26)$$

Since there is a constraint π_i that needs to be satisfied by constraint $\sum_{i=1}^{N} \pi_i = 1$, write the Lagrange function using the Lagrange multiplier method.

Taking the partial derivative and making the result zero, it follows that

$$\frac{\partial}{\partial \pi_t}\left[\sum_{i=1}^{N} \log \pi_t P\left(O, i_1 = i|\ \bar{\lambda}\right) + \gamma\left(\sum_{i=1}^{N} \pi_t - 1\right)\right] = 0 \quad (28)$$

The result is found π_i.

$$\pi_i = \frac{P\left(O, i_1 = i|\ \bar{\lambda}\right)}{P(O|\ \bar{\lambda})} \quad (29)$$

To find the parameter B, the third term of equation (25) can be written as

$$\sum_{i}\left(\sum_{i=1}^{T}\log b_h\left(o_t\right)\right)P(O,I|\ \bar{\lambda})=\sum_{j=1}^{N}\sum_{l=1}^{T}\log b_j\left(o_t\right)P\left(O,i_t=j|\ \bar{\lambda}\right)\quad(32)$$

The Lagrangian multiplier method is still used, with the constraint $\sum_{k=1}^{M}b_j(k)=1$. Note that the partial derivative of $b_j(o_t)$ with respect to $b_j(k)$ is not zero only at $o_t=v_k$.

III. EXPERIMENTS

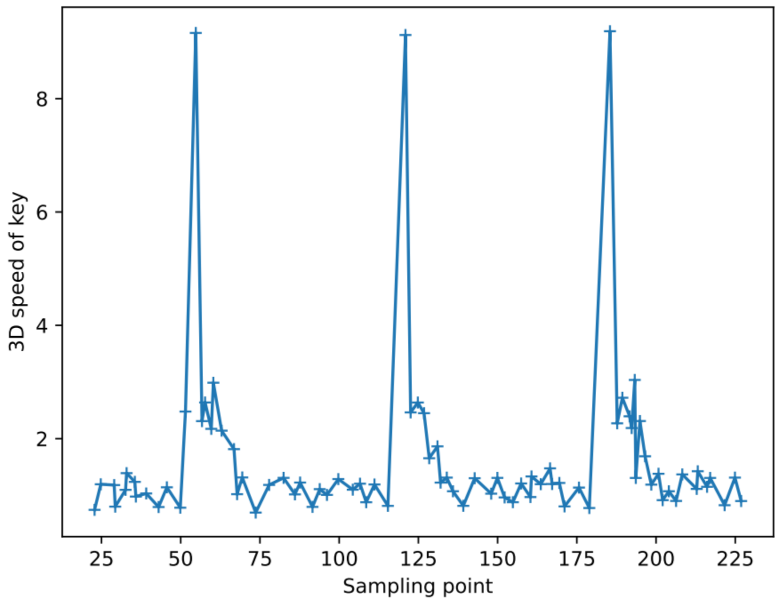
First of all, we look at the aspects of manual evaluation of music performance, the elements involved are the playing environment, the player, the instrument, and the listener, and the listener's perception of the music is an evaluation process. It is not feasible to use the characteristics of each tone as input parameters, because all these individual tones have a lot of redundant information in terms of the factors affecting the performance. The characteristics of the tone include height, strength, length, and timbre. Ignoring the timbre the messages obtained from the MIDI instrument can be resolved to other characteristics. The length of the synthesized features correspond to the beat and rhythm, respectively, and here we only need to synthesize the characteristics of the pitch.

The pitch properties obtained from piano messages are discretized into 127 levels. The difference between each level is one semitone, corresponding to the two adjacent keys of the piano keyboard (no distinction is made between black and white keys), and it is a keystroke error. The key error has the most serious effect on the music, so it is judged as an error regardless of the difference in pitch. The combined factor for each measure is simply a quantification of the number of key errors.

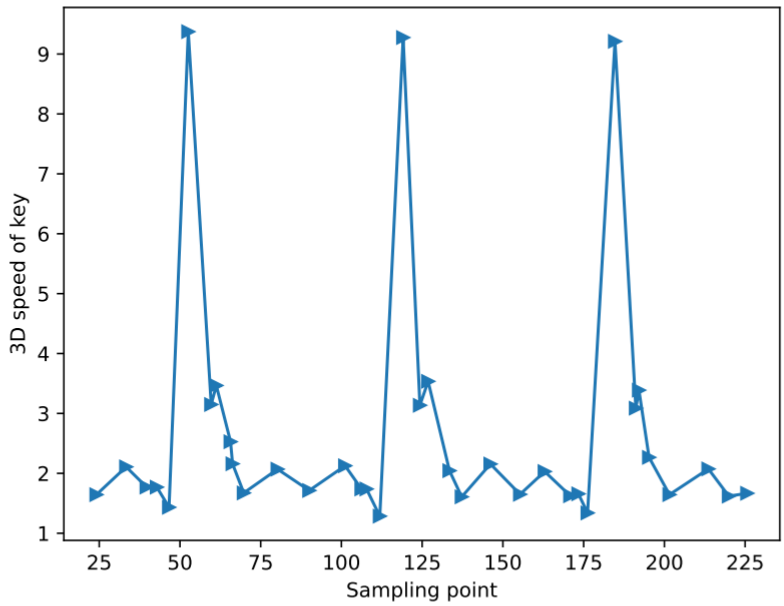
The MIMD messages are 71, 7l, 72, and 74 for pitches B, B, C, and D. When playing four notes correctly, the input parameter is 4/4=l; if only three notes are played correctly, the parameter is 3/4=0.75, and so on. The MIDI signal quantifies the strength and weakness values in 127 discrete quantities, and to find out the absolute value of strength and weakness, we consider here that the input sample with good grasp of the music beat is averaged as the standard value, i.e., a player with a very high level of performance (e.g., a piano teacher) is used to play and obtain the absolute value of strength and weakness. absolute value. For example, in the first measure of Ode to Joy, the standard values of each tone strength and weakness are 100, 70, 90, and 70, which correspond to strong, weak, weak, and second strong. The values of 98, 75, 80 and 70 are obtained when the performer plays the first bar of the ode, so the value of this bar for the beat is 2+5+10+0=17. Also corresponding to 16 bars, 16 input layers of neural network elements are needed to correspond to them.

The pitch, intensity and duration of each chord are calculated according to the method mentioned earlier, with the difference of the standard values. The difference in pitch is determined by first determining the right or wrong fundamental note.

An example of SMA smoothing filtering effect is given in Figure 4, where Figure 4 (a) shows the original data of the 3D ensemble velocity when the finger does the clicking action, and Figure 4 (b) shows the filtered data.



(a) 3D instantaneous combined velocity raw data



(b) Data after SMA filtering

Figure 4 Example of SMA smoothing filtering effect

When Leap Motion captures gesture acquisition data, it will last for a period of time and contain many gestures, forming a series of discrete points to form a lengthy data, so it is necessary to segment and extract the gestures for subsequent gesture classification and recognition.

In this section, a sliding window segmentation technique is used to extract individual gesture movements. The sliding window-based data segmentation technique is to segment a fixed-length window from a continuous segment of motion signals that slides along the time axis, and to average all the data within the window. If the average velocity is small and below a certain threshold, then there is no gesture motion in the window; if the average velocity reaches the maximum, then there is a gesture motion in the window and the window time period just covers a single gesture motion.

The instantaneous combined velocity of fingertips and the average combined velocity within the whole window can be calculated by the following equation.

$$V_i = \arg\max_l \left[\frac{1}{k} \sum_{i}^{t+k-1} v_1 \right] (36)$$

V_i denotes the 3D instantaneous ensemble velocity of the i th sample point. Since the time interval between two adjacent two sample points is fixed, it is advisable to Δt =1, k is the window width, and since the Leap Motion data sampling rate is 100fps/s, and the general gesture movement does not exceed ls, it is advisable to k =100. Set the minimum gaps value to 1, and if the average ensemble velocity in the window is lower than this threshold then it is considered that there is no gesture movement.

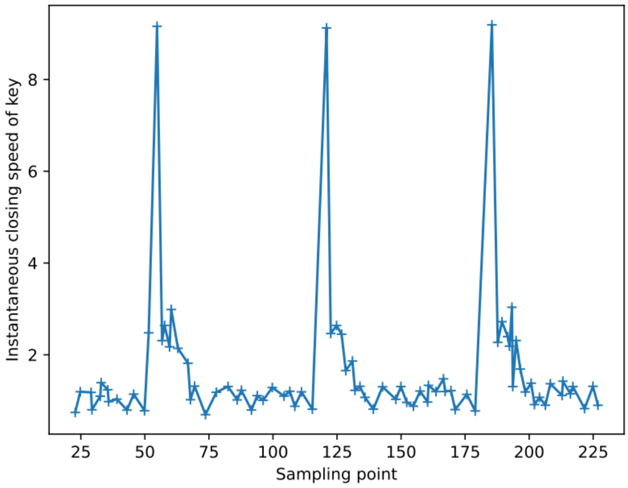


Figure 5 Schematic diagram of sliding window for click action

Figure 5 shows an example of segmentation and extraction of finger clicks. As can be seen from the figure, the sliding window segmentation technique allows the window to be accurately positioned at each click action moment, thus segmenting and extracting individual click actions and laying the foundation for subsequent gesture recognition.

The single gesture extracted in the previous section has 100 sample points, each of which contains the coordinate data of x, y, z axes. To facilitate the HMM modeling operation, each gesture action is divided into frames, and the feature values of each frame are extracted and then encoded in the velocity direction, so that each gesture action can be decomposed into multiple meta-actions, and the feature vectors of each meta-action form an observation sequence after encoding in the velocity direction, so that a gesture action is divided into an n -dimensional vector associated with a time series. In other words, action framing is performed to reduce the dimensionality while fully guaranteeing the characteristics of the gesture action data.

The velocity direction of the finger movement best reflects the characteristics of the gesture action. To this end, the normalized combined velocity vector of the finger is first derived. The normalized combined velocity vector can be found according to the following transformation relationship between the spherical coordinates and the right-angle coordinate system with the angle of z -axis and x -axis respectively θ, φ, as shown in Figure 6.

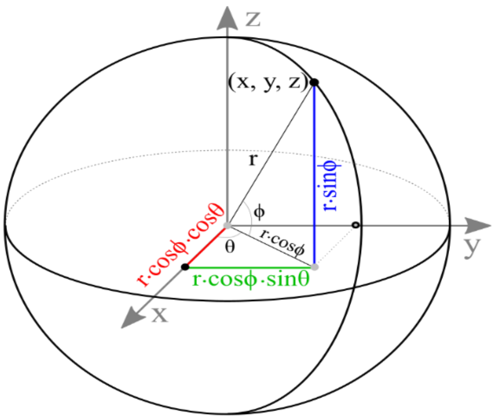


Figure 6 Schematic diagram of vector direction in the sphere

As shown in Table 1, the velocity vector quantization is coded into 26 levels according to θ and φ in this paper. It is equivalent to say that firstly, θ is divided into one code segment every 45, and when θ is far from z -axis, the xy plane is divided into 8 equal parts, i.e., each code segment contains 8 codes, as shown in Figure 7. When θ =0 and 180 degrees each occupy one code, so that a total of 3*8+2=26 codes are generated.

Table 1 Table of speed vector codes

θ value range	Coding segment
0°±22.5°	0
45°±22.5°	1-8
90°±22.5°	9-16
135°±22.5°	17-24
180°±22.5°	25

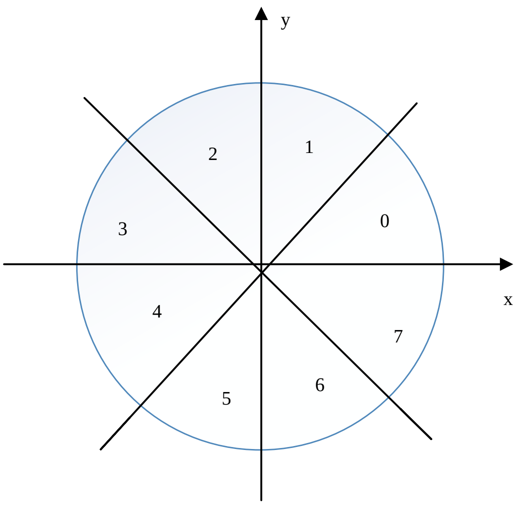


Figure 7 Quantization encoding of velocity direction vectors in the $x-y$ plane

For each sub-frame of the gesture action, the average of the θ, φ values of the normalized combined velocity vector within a sub-frame is obtained, and then encoded according to Table 1.

IV.CONCLUSION

The artistic style, melodic changes and structure of piano performance are unique, so when processing the tone of piano performance, it is necessary to analyze in depth the key touching techniques of different piano keys, the key touching order and strength under the melodic changes of the repertoire, which is an effective way to improve the effect of piano performance and increase the musical impact. The effect of piano tone processing can be significantly different with different key touch methods chosen. The requirement of music image and music expression mode, the clear relationship of progressive promotion between the two, and the proficiency of different key touch methods can promote the performance effect, artistic effect, emotional expression and music expression of piano repertoire.

REFERENCES

[1] Palanisamy, S.; Thangaraju, B.; Khalaf, O.I.; Alotaibi, Y.; Alghamdi, S.; Alassery, F. A Novel Approach of Design and Analysis of a Hexagonal Fractal Antenna Array (HFAA) for Next-Generation Wireless Communication. Energies 2021, 14, 6204. https://doi.org/10.3390/en14196204

[2] Alsubari, S. N., Deshmukh, S. N., Alqarni, A. A., Alsharif, N., H., T. et al. (2022). Data Analytics for the Identification of Fake Reviews Using Supervised Learning. CMC-Computers, Materials & Continua, 70(2), 3189–3204.

[3] Verdugo, F., Pelletier, J., Michaud, B., Traube, C., & Begon, M. (2020). Effects of trunk motion, touch, and articulation on upper-limb velocities and on joint contribution to endpoint velocities during the production of loud piano tones. *Frontiers in Psychology*, *11*, 1159.

[4] Weinstein-Reiman, M. (2021). Printing Piano Pedagogy: Experimental Psychology and Marie Jaëll's Theory of Touch. *Nineteenth-Century Music Review*, *18*(3), 427-453.

[5] Li, S., & Timmers, R. (2020). Exploring pianists' embodied concepts of piano timbre: An interview study. *Journal of New Music Research*, *49*(5), 477-492.

[6] Zhang, L. X. (2019). The "Four Main Factors Theory" of Piano Teaching and its Systematic Thinking. *Advances in Social Science Education and Humanities Research*, *322*, 359-362.

[7] Pala, F. K., & Türker, P. M. (2019). Developing a haptic glove for basic piano education. *World Journal on Educational Technology: Current Issues*, *11*(1), 38-47.

[8] Yücetoker, İ., Angi, Ç. E., & Kaynak, T. (2021). Evaluation of asynchronous piano education and training in the Covid-19 era. *Educational Research and Reviews*, *16*(4), 109-117.

[9] Bobbe, T., Oppici, L., Lüneburg, L. M., Münzberg, O., Li, S. C., Narciss, S., ... & Muschter, E. (2021). What early user involvement could look like—developing technology applications for piano teaching and learning. *Multimodal Technologies and Interaction*, *5*(7), 38.

[10] Timbrell, C. (2019). Beethoven: Piano Sonata in E Major, Op. 109; Piano Sonata in A-flat Major, Opus 110; Piano Sonata

in C Minor, Opus 111. *The Beethoven Journal*, *34*(1), 30-30.

[11] Hassan Mohamed Mahmoud PASHA, S. (2018). Advantages of breathing exercises on piano performance for beginners. *International Journal of Education and Learning Research*, *1*(1), 10-17.

[12] Comeau, G., Lu, Y., & Swirp, M. (2019). On-site and distance piano teaching: An analysis of verbal and physical behaviours in a teacher, student and parent. *Journal of Music, Technology & Education*, *12*(1), 49-77.

[13] Micheloni, E., Tramarin, M., Rodà, A., & Chiaravalli, F. (2019). Playing to play: a piano-based user interface for music education video-games. *Multimedia Tools and Applications*, *78*(10), 13713-13730.

[14] Ioniță, R. D. (2021). The Second Sonata for Piano and Violin, Ópus 6 by George Enescu. *Review of Artistic Education*, (21), 149-156.

[15] Wang, X., Liu, J. & Zhang, C. Network intrusion detection based on multi-domain data and ensemble-bidirectional LSTM. EURASIP J. on Info. Security 2023, 5 (2023). https://doi.org/10.1186/s13635-023-00139-y