

Received: 03 August 2026; **Revised:** 14 September 2026; **Accepted:** 29 September 2026; **Published:** 18 October 2026

Archs Sci. (2026) Volume 76(4), Pages 197-204, Paper ID 2026421.
<https://doi.org/10.68304/as/76421>

Machine Learning in Teaching Piano Touch Technique and Tonal Expression

Da Li^{1,*}, Jing Zhou² and Qiyao Tan¹

¹School of Arts, Nanchang University, Nanchang, Jiangxi 330027, China

²School of Music, Jiangxi Science and Technology Normal University, Nanchang, Jiangxi 330038, China

Correspondence: 18970945811@163.com

Abstract Music is an art involving both sound and hearing, and it is experienced through a combination of musical performance and attentive listening. Among the many musical instruments, the piano is often called the “king of instruments.” Unlike the violin, which is frequently described as delicate and soft, the piano possesses a tone that may be perceived as noble and graceful. A statement attributed to the pianist Nergoz emphasizes that the purpose of piano playing is to make sound and to create sound. This observation draws attention to the central importance of sound production in piano education, although the attribution needs confirmation from an identifiable original source. How, then, can a pianist use the instrument to produce a beautiful sound? This remains an essential concern for piano players at different levels of study and performance. Therefore, tonal quality deserves consistent attention during both technical practice and musical performance, because the fingers must engage with the keys to produce each intended sound. Tone plays an important role throughout the entire process of piano performance. Both the expressiveness of piano playing and its appeal to listeners are closely associated with the quality of the resulting tone. The area of key contact, the intensity of the touch, and the technique employed all have an important influence on the resulting sound. A systematic approach to finger touch is consequently an important skill and a valuable technical consideration, both when learning to play the piano and when performing musical compositions.

Index Terms performance, piano, tonal expression, key touch, machine learning

I. Introduction

People increasingly seek to satisfy higher-level cultural and artistic needs after meeting basic requirements such as clothing, food, housing, and transportation. Learning to play the piano has consequently become an attractive activity for many people. In recent years, interest in piano learning has reportedly increased substantially among learners of different ages, with young children representing a particularly important group of students. In some urban middle-income and higher-income families, it has been claimed that one in five households has parents who want their children to learn the piano and that applications for piano grade examinations have increased by approximately 20% each year. Such figures are highly dependent on the particular population and educational system being considered, and their interpretation requires reliable information about the surveyed households, geographical location, and reporting period. Moreover, reforms in the national education system and the expansion of educational provision have reportedly increased the demand for piano instructors. In the field of piano education, professionally qualified teachers are sometimes described as a scarce resource, and such shortages may not be easily resolved within a short period of time.

Piano education can also be expensive, with some accounts suggesting that an average piano costs at least \$10,000. Piano instruction is often delivered through individual, one-to-one lessons, and tuition fees have been reported as ranging between \$100 and \$400 per lesson. These estimated costs can differ considerably by instrument type, country, teaching institution, and period, and published comparisons of piano education therefore require a carefully defined market and a consistent pricing basis. In this situation, some students attend only one formal piano lesson per week and must undertake much of their remaining practice independently. Without sufficient guidance, learners may develop inefficient practice habits, encounter avoidable difficulties, and progress at a slower rate than expected. Independent practice and the development of supportive learning technologies have therefore attracted attention in research on piano education [1], [2]. Access to piano instruction may also be restricted by geographical distance, scheduling difficulties, and the limited availability of qualified teachers, which can discourage interested learners from continuing their studies. Studies of asynchronous lessons and comparisons of on-site and distance instruction offer relevant evidence about the possibilities and limitations of remote teaching [3], [4].

Excessive or unnecessary movement is a common technical difficulty encountered during piano training. Pianists may need to reduce avoidable external force, simplify their playing movements, and improve finger stability in order to produce a focused, powerful, and bright tone. To maintain a consistent tone throughout a passage, it is useful to avoid unnecessary vertical wrist movement, pay particular attention to octave passages, and minimize excessive movement during rapid playing. These adjustments can help performers achieve smoother and more coordinated motion. Biomechanical studies comparing experienced and novice pianists, investigating tactile feedback, and examining the influence of trunk motion and touch provide a research basis for considering movement organization and control, although they do not establish a universal prescription for every performer [5]–[7]. After each note is sounded, the hand should be prepared to release the key and transfer the required movement to the next note, particularly during rapid groups of notes. The transfer of movement and effort between the fingers should remain as relaxed and efficient as possible. It is also important to avoid consciously applying excessive pressure through the wrist. This helps the weight and movement of the hand remain coordinated with the movements of the fingertips. For a smooth and even tone, the pianist's arm, elbow, and wrist should move in a manner that supports the intended fingertip action.

In addition, tonal production is affected by the support and coordination of the fingers. Effective finger support is an important element of piano technique because it influences both mechanical control and the resulting quality of sound. Mr. Ling Zhuping has stated that the supporting capacity of the fingers is limited and that excessive pressure from the arm may contribute to discomfort or injury in the hand muscles. Nevertheless, the fingers should not be placed loosely on the keys without appropriate contact or control, because insufficient finger support may also weaken the intended tonal effect. In the climax of the fourth variation of Glinka's "The Nightingale," for example, the left hand plays a leaping eighth-note broken-chord figure while the right hand performs a sixteenth-note broken-chord figure in octaves, requiring considerable coordination and finger stability. The fingers should maintain secure but moderate contact with the keys rather than exerting excessive pressure, which may restrict finger flexibility; meanwhile, the palms should remain relaxed instead of tense. Because the five fingers differ in their anatomical characteristics and mechanical capacities, coordinated and even finger control may require additional practice, particularly for the little finger. When each finger can provide independent support, the hands maintain an appropriate level of relaxation, and movement is transferred smoothly from the arms toward the keys, the pianist is better positioned to coordinate tonal quality and loudness. It is also important to maintain a sense of support between the shoulder and the fingertips during performance. Although the centre of the palm may remain slightly arched, the hand as a whole continues to contribute to controlled support. Across different tonal demands and playing speeds, stable finger positioning and controlled movement can help

performers retain the necessary mechanical support. These considerations demonstrate the importance of finger stability, coordination, and support in the performance of piano works. Research on piano biomechanics and historical approaches to touch supports their pedagogical relevance, while suggesting that instructions about hand position should be adapted to the individual performer and the musical task [7], [8].

In piano performance, the manner in which the keys are touched can influence the perceived quality of the tone and the range of expressive variations available to the pianist. If touch is poorly controlled, the tonal result may be unsatisfactory, and the variety of expressive effects can be restricted. The piano is a complex instrument whose performance requires not only technical competence but also a thorough understanding of the musical work and an ability to convey its intended ideas and emotions. Studies of piano pedagogy and timbral perception have highlighted the importance of coordinating touch, auditory perception, and musicians' embodied understanding of sound [8]–[10]. The relationship between a pianist's description of timbre and acoustically measurable changes remains an important distinction, and educational discussion should not treat every perceived tonal difference as proof of an independently altered acoustic spectrum. The piano is capable of producing a wide range of dynamic effects, and appropriate key-touch techniques are important for performing different levels of tonal intensity. Tonal production is influenced by fingertip movement, the stability of the hand, and the flexibility and relative independence of the fingers. Accordingly, carefully developed touch technique is a central aspect of effective piano performance. Digital interfaces and interactive learning technologies have also been investigated as possible supports for piano education, providing a wider context for the sensor-based methods discussed in the following sections [1], [2], [11]. These published applications establish the feasibility of studying instructional technologies; they do not, by themselves, establish that the particular system described in the present manuscript improves students' performance.

II. Methodology

To explore the relationship between piano touch technique and tonal expressiveness, the playing function must first be analysed. The overall analytical procedure is presented in the block diagram shown in Figure 1.

When an object tracked by the Hand Controller is processed through the `onFrame()` callback function, the system obtains data frames representing individual parts of the tracked object. The structure of the resulting data frame is illustrated in Figure 2.

High-accuracy tracking of finger positions can produce high-frequency noise during gesture detection, because both small involuntary finger tremors and fluctuations in the sensor measurements may affect the recorded coordinates. Consequently, filtering and smoothing procedures are required to reduce unwanted fluctuations in the recorded motion data. Published evaluations of the Leap Motion Controller document measurement accuracy and reliability limitations, including

differences between static and dynamic tracking conditions [12], [13].

The trajectory followed by a moving finger can be approximated by a continuous curve. Accordingly, the relationship between the finger's coordinate position and time can be expressed using the following functions.

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

Taking the first-order derivatives of the coordinate functions in Eq. (1) gives the following expressions:

$$\begin{cases} x'_t = \frac{dx}{dt} m'(t), \\ y'_t = \frac{dy}{dt} n'(t). \end{cases} \quad (2)$$

Eq. (2) describes the rate of finger movement along the indicated Z and J axes, from which the angle associated with the direction of finger movement can be determined.

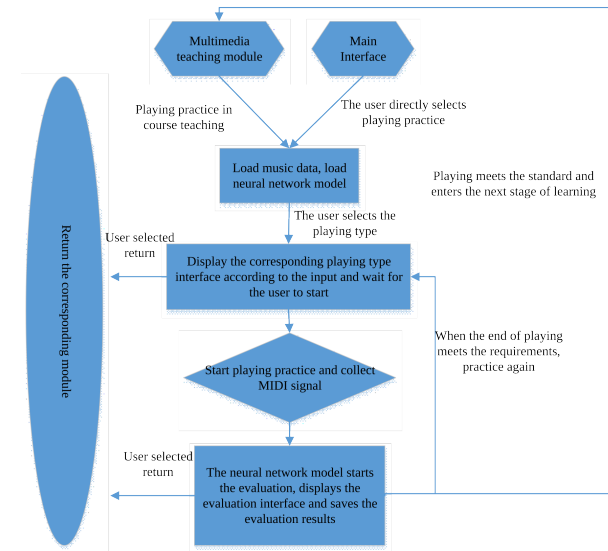


Figure 1: Block diagram for the analysis of the playing function

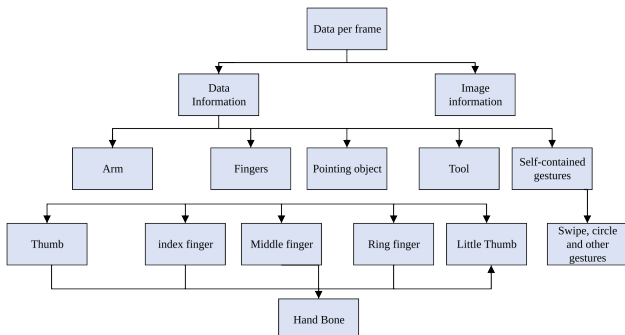


Figure 2: Data frame structure

Because the recorded discrete observations are collected at fixed, equal time intervals, and because these sampling intervals are relatively short, the derivatives in Eq. (2) can be approximated using differences between adjacent observations. The direction of finger motion can therefore be calculated using the following expression.

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

The angular change in the direction of the recorded movement between two consecutive sampling points can consequently be expressed using Eq. (4).

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

A threshold, τ , is then applied to identify possible jitter in the movement data. When the angular change remains below the selected threshold, the observation is retained; when the angular change exceeds the threshold, the observation is treated as potential jitter noise and filtered out, according to the following decision rule.

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

According to the repeated experiments described in this study, the threshold in Eq. (5) was set at 30 degrees. In particular, changes in the direction of a movement gesture were retained when their angular magnitude was less than 30 degrees, whereas observations with an angular change greater than 30 degrees were removed during the initial filtering stage.

To suppress remaining high-frequency noise more effectively, the simple moving average (SMA) filter can additionally be applied as a data-smoothing technique. The operation of this filter is represented by the following equation.

$$S = \frac{(x_t + x_{t-1} + \dots + x_{t-n+1})}{n}, \quad (n = 1, 2, 3, \dots). \quad (6)$$

In Eq. (6), S represents the calculated moving-average value of the data, while n denotes the number of observations included in the averaging window. The choice of n affects the overall smoothing result: an excessively large window may remove meaningful details of the movement gesture, whereas an excessively small window may fail to suppress unwanted fluctuations adequately.

Hidden Markov models (HMMs) are probabilistic models that can be applied to sequential classification and labelling problems. They have been employed in several fields, including speech recognition, natural language processing (NLP), bioinformatics, pattern recognition, and artificial intelligence. Their ability to describe sequences of hidden states and observable events also makes them relevant to the analysis of

time-dependent motion data. Established treatments explain both HMM inference and its relationship to statistical pattern-recognition methods [14], [15].

The HMM is represented by the following model parameters and expressions:

$$\lambda = (A, B, \pi), 2TN^2T, \quad (7)$$

$$P(Q | \lambda) = \pi_{q_1} \alpha_{q_1 q_2} \alpha_{q_2 q_3} \dots \alpha_{q_{r-1} q_r} Q.$$

In addition to the three principal elements specified above, applying an HMM to observed data involves addressing both the parameter-learning problem and the sequence-prediction problem. These tasks must be distinguished because they describe different stages of model estimation and interpretation.

The expression for the output probability associated with the state sequence $Q = q_1, q_2, \dots, q_r$ under the specified hidden Markov model λ is given as follows.

$$P(O | \lambda) = \pi_{q_1} \alpha_{q_1 q_2} \alpha_{q_2 q_3} \dots \alpha_{q_{r-1} q_r}. \quad (8)$$

The probability expression for generating the observed sequence, conditional on the underlying state sequence, is presented below.

$$P(O | Q, \lambda) = \prod_{t=1}^T P(o_t | q_t, \lambda) = b_{q_1}(o_1) b_{q_2}(o_2) \dots b_{q_r}(o_r), \quad (9)$$

Using the expressions above, the observation probability for a given model λ can be written in the following form.

$$P(O | \lambda) = \sum_{q_1 q_2 \dots q_{iT}} \pi_{q_1} b_{q_1}(o_1) a_{q_1, q_{12}} b_{q_{12}}(o_2) a_{q_{12} q_{13}} \dots a_{q_{1(\tau-1)} q_{i\tau}} b_{q_{i\tau}}(o_T). \quad (10)$$

The expression in Eq. (10) is intended to account for the possible hidden-state sequences associated with the observed sequence O . The direct calculation is described as having approximate computational complexity $2TN^T$, whereas the corresponding dynamic-programming approach is described as having complexity N^2T . These complexity expressions depend on the definitions of the numbers of observations, states, and transitions and should be checked against the implemented model. This comparison is presented as part of the computational motivation for using sequential models in the proposed gesture-recognition procedure.

(1) Forward algorithm

Given a model λ , the forward probability is defined as the probability of observing the partial sequence $o_1, o_2, \dots, o_r$ up to time t and reaching the specified state q_i at that time. This probability is expressed in the following form.

$$MQ(\lambda, \bar{\lambda}) = \sum_I \log P(O | I, \lambda) P(O | I, \bar{\lambda}). \quad (11)$$

(2) Backward algorithm

Given a model λ , the backward probability describes the probability of the future observations $o_{t+1}, o_{t+2}, \dots, o_T$, conditional on the state at time t being q_i . The corresponding probability expression is presented below.

$$P(O | \lambda) = \sum_{i=1}^N \pi_i b_i(o_1) \beta_1(i). \quad (12)$$

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

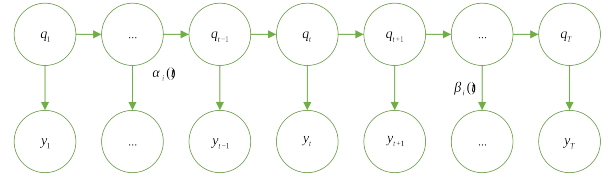


Figure 3: Relationship between forward and backward probabilities

When the available training data consist only of observation sequences, HMM parameters can be estimated using the expectation-maximization (EM) algorithm, which is a standard approach to unsupervised parameter learning in probabilistic sequential models. The general estimation principle is established in foundational work on maximum-likelihood estimation from incomplete data [16]. For the hidden Markov model, let $O = (o_1, \dots, o_T)$ denote the observed sequence, $I = (i_1, \dots, i_T)$ denote a possible hidden-state sequence, $\lambda = (A, B, \pi)$ denote the model parameters to be estimated, and $\bar{\lambda}$ denote the current parameter estimates. Following the HMM auxiliary-function formulation [14], [16], the unnormalized EM objective can be written as

$$\begin{aligned} \mathcal{Q}(\lambda, \bar{\lambda}) &= \sum_I P(O, I | \bar{\lambda}) \log P(O, I | \lambda) \\ &= \sum_I P(O, I | \bar{\lambda}) \left[\log \pi_{i_1} + \sum_{t=1}^{T-1} \log a_{i_t i_{t+1}} \right. \\ &\quad \left. + \sum_{t=1}^T \log b_{i_t}(o_t) \right]. \end{aligned} \quad (13)$$

The terms a_{ij} and $b_j(o_t)$ represent the state-transition and observation-emission probabilities, respectively. The objective in Eq. (13) uses the joint probability $P(O, I | \bar{\lambda})$ as a weight; dividing by the fixed quantity $P(O | \bar{\lambda})$ gives the conventional normalized conditional-expectation objective and does not change its maximizing parameter values.

(1) To estimate the initial-state probability π_i , the first term in Eq. (13) can be collected by initial hidden state as follows:

$$\sum_I P(O, I | \bar{\lambda}) \log \pi_{i_1} = \sum_{i=1}^N P(O, i_1 = i | \bar{\lambda}) \log \pi_i. \quad (14)$$

The parameter π_i must satisfy the normalization constraint $\sum_{i=1}^N \pi_i = 1$. Therefore, the Lagrange multiplier method can

be applied to construct an appropriate constrained objective function and derive the required parameter estimate.

Taking the partial derivative of the constrained objective function and setting the resulting expression equal to zero gives the following condition:

$$\frac{\partial}{\partial \pi_i} \left[\sum_{j=1}^N P(O, i_1 = j | \bar{\lambda}) \log \pi_j + \gamma \left(\sum_{j=1}^N \pi_j - 1 \right) \right] = 0. \quad (15)$$

Solving the resulting condition yields the following estimate for the parameter π_i :

$$\pi_i = \frac{P(O, i_1 = i | \bar{\lambda})}{P(O | \bar{\lambda})}. \quad (16)$$

To estimate the observation-emission parameters B , the third term in Eq. (13) can be reorganized by hidden state and observation time as follows:

$$\begin{aligned} \sum_I P(O, I | \bar{\lambda}) \sum_{t=1}^T \log b_{i_t}(o_t) \\ = \sum_{j=1}^N \sum_{t=1}^T P(O, i_t = j | \bar{\lambda}) \log b_j(o_t). \end{aligned} \quad (17)$$

The Lagrange multiplier method is again applied, this time with the constraint $\sum_{k=1}^M b_j(k) = 1$. The partial derivative of $b_j(o_t)$ with respect to $b_j(k)$ is nonzero only when the observation o_t is equal to v_k , which is the corresponding discrete observation symbol.

III. Experiments

First, consider the manual evaluation of musical performance. Its principal elements include the performance environment, the performer, the musical instrument, and the listener. The listener's perception of the resulting music forms an important component of the evaluation process. Using every individual tonal characteristic as an input parameter may be impractical because several features contain redundant information concerning the factors that influence performance quality. The principal characteristics of a musical tone include pitch, intensity, duration, and timbre. If timbre is excluded from the immediate analysis, the remaining characteristics can be obtained from messages generated by a MIDI instrument. Representing pitch, onset, duration, and expressive timing as computational features is consistent with established approaches to music-information processing [17]. Duration-related features are associated with the beat and rhythm, while the present example focuses on the synthesis and evaluation of pitch characteristics.

The pitch properties obtained from the piano's MIDI messages are represented using 127 discrete levels. A difference of one level represents a semitone, corresponding to adjacent notes on the piano keyboard without distinguishing between black and white keys. Playing an incorrect key produces a

pitch error that may have a substantial effect on the musical result. Consequently, a pitch deviation is classified as an error regardless of the size of the difference. The combined evaluation factor for each measure is therefore determined by quantifying the number of incorrect keystrokes.

The MIDI note numbers in the source draft are given as 71, 71, 72, and 74 for the example pitches B, B, C, and D. The second code contains a possible transcription error (the letter "l" rather than the numeral "1") and should be confirmed using the original MIDI record. If all four notes are played correctly, the input parameter is $4/4 = 1$; if only three notes are played correctly, the corresponding parameter is $3/4 = 0.75$, and the same calculation is extended to other outcomes. MIDI messages also provide a discrete measure of note-on velocity associated with performance intensity. The coding of intensity depends on the precise MIDI message definition, the interpretation of note-on velocity, and the device-specific treatment of zero velocity, all of which should be specified for a reproducible analysis. To establish a reference level for intensity, the performance of a musician with a strong command of the beat, such as an experienced piano teacher, can be used to obtain standard measurements of dynamic intensity. For example, in the first measure of "Ode to Joy," the reference intensity values are 100, 70, 90, and 70, representing the intended dynamic pattern. If a performer produces the corresponding values 98, 75, 80, and 70, the total absolute intensity deviation for that measure is $2 + 5 + 10 + 0 = 17$. For a musical passage containing 16 measures, the proposed representation consequently requires 16 input-layer elements in the neural network, with one element associated with each measure.

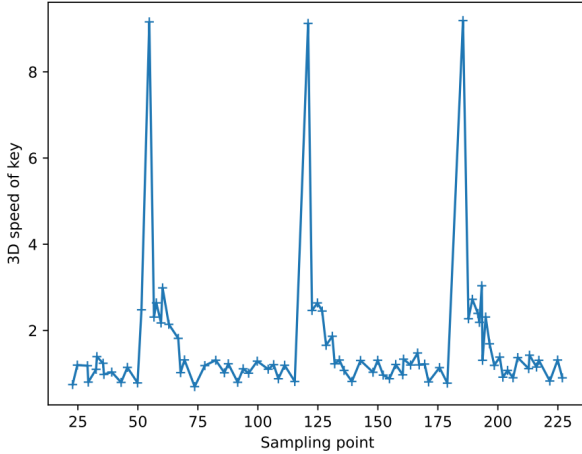
The pitch, intensity, and duration of each chord are evaluated using the method described above, with the observed values compared against the corresponding reference values. The difference in pitch is determined by first identifying whether the fundamental note of the chord has been played correctly and then examining any remaining deviations.

An example illustrating the effect of SMA smoothing is provided in Figure 4. Figure 4(a) presents the original three-dimensional combined-velocity data recorded during a finger-click action, whereas Figure 4(b) shows the corresponding data after filtering.

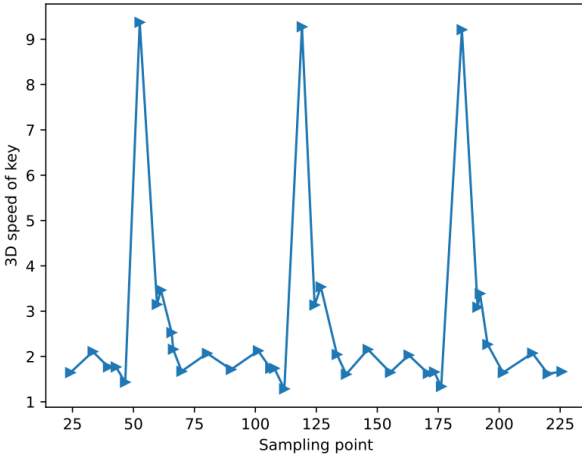
During gesture acquisition, the Leap Motion device records movements over a period of time and may capture multiple consecutive gestures. These recordings contain long sequences of discrete observations representing different stages of movement. As a result, it is necessary to segment the continuous data and extract individual gestures before performing subsequent gesture classification and recognition. Research on Leap Motion-based gesture recognition demonstrates how tracked fingertip positions and orientations can be transformed into features for classification [18].

A sliding-window segmentation technique is used in this study to isolate individual gesture movements. The technique extracts a fixed-length window from a continuous sequence of motion signals, moves that window along the time axis, and

calculates the average velocity of observations within each window. When the average velocity falls below a selected threshold, the corresponding window is interpreted as containing no significant gesture movement. Conversely, when the average velocity reaches a sufficiently high value, the window is treated as containing gesture activity. The chosen window can then be positioned to cover the interval associated with an individual gesture movement.



(a) Original three-dimensional instantaneous combined-velocity data



(b) Velocity data after SMA filtering

Figure 4: Example illustrating the effect of SMA smoothing

The instantaneous combined velocity of the fingertips and the average combined velocity over the selected window are calculated using the following expression:

$$V_i = \arg \max_i \left[\frac{1}{k} \sum_{i=t-k+1}^{t+k-1} v_i \right], \quad (18)$$

Here, V_i denotes the three-dimensional instantaneous combined velocity of the i th sampled point. Because the sampling interval between adjacent observations is fixed, the calculation adopts $\Delta t = 1$ for the discrete-time representation. The parameter k denotes the width of the sliding window. The

Leap Motion sampling rate is stated as 100 frames per second, and typical gesture movements are assumed to last no longer than one second; consequently, the window width is set to $k = 100$ observations. The minimum gap threshold is set to 1, and any window whose average combined velocity falls below this threshold is treated as containing no detectable gesture movement.

Figure 5 illustrates the segmentation and extraction of individual finger-click actions. The example indicates that the sliding-window technique can identify the time intervals associated with successive clicks, thereby allowing individual movements to be isolated for subsequent recognition. This segmentation stage provides the data structure needed for the next steps in gesture classification and motion analysis.

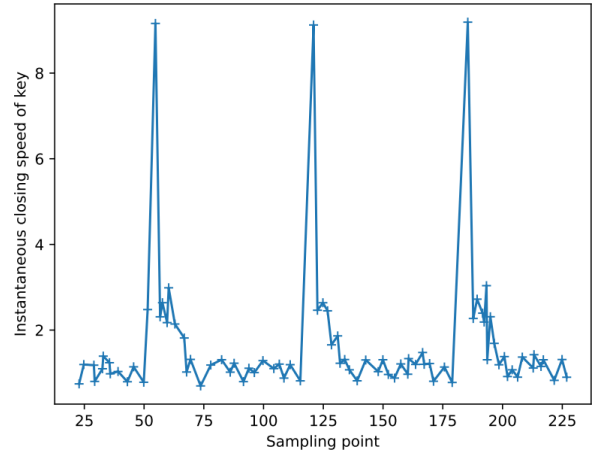


Figure 5: Schematic representation of the sliding window used for a finger-click action

Each individual gesture extracted during the preceding stage contains 100 sample points, with each point recording coordinate information along the x , y , and z axes. To facilitate HMM-based modelling, the complete gesture is divided into smaller frames, and characteristic features are extracted from each frame. These features are subsequently encoded according to the direction of movement velocity. In this way, each gesture can be represented as a sequence of simpler component actions, or meta-actions, whose feature vectors form an ordered sequence of observations. The original gesture is therefore represented as an n -dimensional, time-dependent feature vector. Frame-based decomposition is intended to reduce the dimensionality of the recorded data while preserving the movement characteristics that are most relevant for recognising the gesture.

The direction of finger velocity is treated as an important descriptor of each gesture's characteristic movement. Accordingly, the normalized combined-velocity vector is calculated before directional encoding is performed. The transformation between spherical coordinates and the Cartesian coordinate system is described using the angles θ and φ , measured with respect to the indicated z and x axes. This geometric representation is illustrated in Figure 6.

As shown in Table 1, the velocity-direction vector is quantized into 26 coding levels on the basis of the angles θ and φ . First, the polar angle θ is divided into segments separated by 45 degrees. For directions away from the z axis, the xy plane is then divided into eight equal angular sectors, with each sector receiving a distinct code, as illustrated in Figure 7. The special cases $\theta = 0^\circ$ and $\theta = 180^\circ$ each receive a single code. This procedure produces a total of $3 \times 8 + 2 = 26$ directional codes.

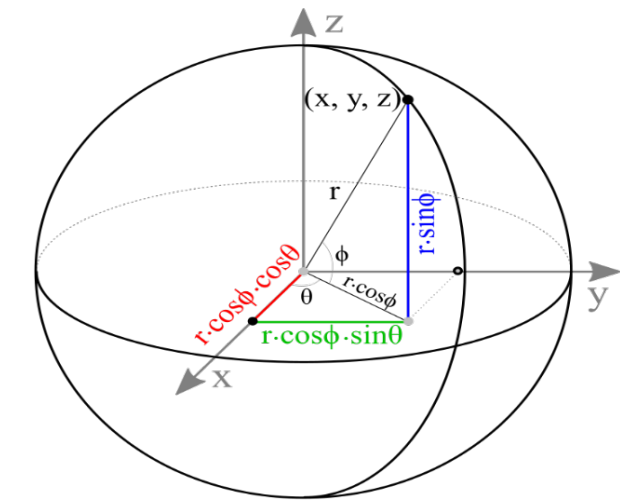


Figure 6: Schematic representation of the vector direction in spherical coordinates

Table 1: Coding levels for the velocity-direction vector

θ value range	Coding segment
$0^\circ \pm 22.5^\circ$	0
$45^\circ \pm 22.5^\circ$	1-8
$90^\circ \pm 22.5^\circ$	9-16
$135^\circ \pm 22.5^\circ$	17-24
$180^\circ \pm 22.5^\circ$	25

For each subframe of a recorded gesture, the mean values of the angles θ and φ are calculated from the normalized combined-velocity vectors belonging to that subframe. The resulting mean direction is then encoded using the quantization levels specified in Table 1.

IV. Conclusion

The artistic style, melodic development, and structural organization of a piano performance provide important contexts for understanding tonal expression. Therefore, a detailed analysis of piano tone should consider the techniques used to touch individual keys, the sequence in which those keys are played, and the intensity of the keystrokes as the melody and musical texture change. Such an approach may help performers develop more deliberate control over tonal production and improve the expressive quality and musical impact of their playing. Different key-touch methods can produce noticeably different performance outcomes, depending on the musical passage and the intended expressive effect. The relationship

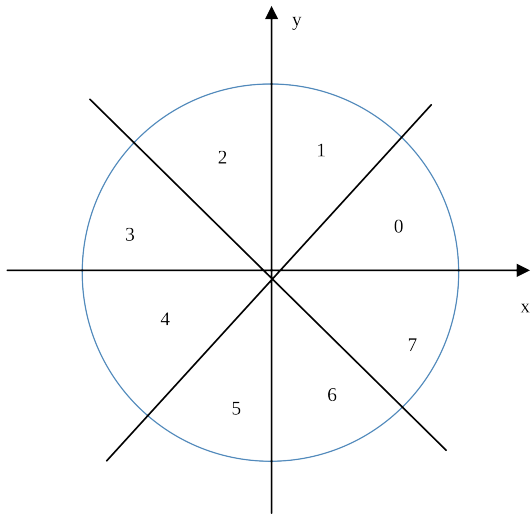


Figure 7: Quantization encoding of velocity direction vectors in the $x - y$ plane

between the musical image and its mode of expression, together with the performer’s ability to select and coordinate suitable key-touch techniques, can contribute to improvements in technical execution, artistic presentation, emotional communication, and the overall interpretation of a piano composition. The analytical framework discussed in this paper emphasizes the role of detailed movement characteristics in understanding finger gestures and their possible relationship to performance technique. The material presented does not provide a controlled comparison of student learning outcomes, classification accuracy, or measured improvements in tonal expression; those educational and technical outcomes would require a separately documented evaluation.

Author contributions

The authors contributed equally.

Conflict of interests

The authors declare no conflicts of interest.

Data availability

Data are available from the corresponding author upon reasonable request.

Funding information

No funding was received for this study.

References

[1] 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. doi: 10.18844/wjet.v11i1.4008.

[2] Bobbe, T., Oppici, L., Lüneburg, L.-M., Münzberg, O., Li, S.-C., Narciss, S., Simon, K.-H., Krzywinski, J., & Muschter, E. (2021). What early user involvement could look like—Developing technology applications for piano teaching and learning. *Multimodal Technologies and Interaction*, 5(7), Article 38. doi: 10.3390/mti5070038.

- [3] Yücetoker, İ., Angı, 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. doi: 10.5897/ERR2021.4136.
- [4] 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. doi: 10.1386/jmte.12.1.49_1.
- [5] Furuya, S., & Kinoshita, H. (2008). Organization of the upper limb movement for piano key-depression differs between expert pianists and novice players. *Experimental Brain Research*, 185(4), 581–593. doi: 10.1007/s00221-007-1184-9.
- [6] Goebel, W., & Palmer, C. (2008). Tactile feedback and timing accuracy in piano performance. *Experimental Brain Research*, 186(3), 471–479. doi: 10.1007/s00221-007-1252-1.
- [7] 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, Article 1159. doi: 10.3389/fpsyg.2020.01159.
- [8] Weinstein-Reiman, M. (2021). Printing piano pedagogy: Experimental psychology and Marie Jaëll's theory of touch. *Nineteenth-Century Music Review*, 18(3), 427–453. doi: 10.1017/S1479409819000715.
- [9] Li, S., & Timmers, R. (2020). Exploring pianists' embodied concepts of piano timbre: An interview study. *Journal of New Music Research*, 49(5), 477–492. doi: 10.1080/09298215.2020.1826532.
- [10] Zhang, L.-X. (2019). The “four main factors theory” of piano teaching and its systematic thinking. In *Proceedings of the 2nd International Seminar on Education Research and Social Science (ISERSS 2019)* (Advances in Social Science, Education and Humanities Research, Vol. 322, pp. 161–164). Atlantis Press. doi: 10.2991/iserss-19.2019.94.
- [11] 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. doi: 10.1007/s11042-018-6917-1.
- [12] Weichert, F., Bachmann, D., Rudak, B., & Fisseler, D. (2013). Analysis of the accuracy and robustness of the Leap Motion Controller. *Sensors*, 13(5), 6380–6393. doi: 10.3390/s130506380.
- [13] Guna, J., Jakus, G., Pogačnik, M., Tomažič, S., & Sodnik, J. (2014). An analysis of the precision and reliability of the Leap Motion sensor and its suitability for static and dynamic tracking. *Sensors*, 14(2), 3702–3720. doi: 10.3390/s140203702.
- [14] Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), 257–286. doi: 10.1109/5.18626.
- [15] Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
- [16] Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1), 1–22. doi: 10.1111/j.2517-6161.1977.tb01600.x.
- [17] Müller, M. (2015). *Fundamentals of music processing: Audio, analysis, algorithms, applications*. Springer. doi: 10.1007/978-3-319-21945-5.
- [18] Marin, G., Dominio, F., & Zanuttigh, P. (2014). Hand gesture recognition with Leap Motion and Kinect devices. In *2014 IEEE International Conference on Image Processing (ICIP)* (pp. 1565–1569). IEEE. doi: 10.1109/ICIP.2014.7025313.

...