Research Article

E-Sports Training System Based on Intelligent Gesture Recognition

Hui Li, Yao Lu, and Hongqiao Yan

School of Sport Communication and Information Technology, Shandong Sport University, Jinan 250000, China

Correspondence should be addressed to Hongqiao Yan; yanhongqiao@sdpei.edu.cn

Received 18 May 2022; Accepted 14 June 2022; Published 8 July 2022

Academic Editor: Xin Ning

Copyright © 2022 Hui Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to improve the effect of e-sport training, this paper combines the intelligent gesture recognition technology to construct an e-sport training system and judges the training effect of players through the recognition of players’ gestures. Moreover, this paper studies the commonly used feature extraction algorithms and proposes an improved SLC-Harris feature extraction algorithm, and the feasibility of this algorithm is verified by the experimental results on the EUROC dataset. In addition, this paper uses the KLT optical flow algorithm to track the extracted feature points and calculates the pure visual pose through epipolar geometry, triangulation, and PnP algorithms. The experimental research results show that the electronic economic training system based on intelligent gesture recognition proposed in this paper has certain effects.

1. Introduction

The reason why e-sports can become a sports competition is that it is closely related to the progress of society, the development of science and technology, and the spiritual and cultural needs of the people. Although there are countless people who enjoy this high-tech intellectual sports event, in fact, public opinion instills the harmful opinion of e-sports in people intentionally or unintentionally. Some media reported extensively that some students were addicted to games and could not extricate themselves, wasting their youth and studies, which made e-sports become “electronic heroin” that everyone shouted. The huge pressure of public opinion makes e-sports face severe survival pressure, and it is difficult for enterprises to enter this market justifiably. Moreover, athletes can only be called “players,” and their treatment cannot be compared with that of ordinary athletes. At the same time, the majority of fans can only engage in e-sports secretly. In addition, in the face of huge pressure from public opinion, it is difficult for the government to guide and supervise confidently, and sometimes, it has to prohibit escrow. The ban on television broadcasting of e-sports competitions can be described as a huge obstacle to the normal development of the current social discrimination against the e-sports sports industry.

Generally speaking, the development of e-sports is not yet mature, and the development of e-sports is still in its infancy [1], which is manifested in many aspects: the public recognition is not enough, there are few related large-scale events, there is no professional-scale operation, there is less research in this area, and so on [2]. Especially on college campuses, although students have more time for self-discipline than before, the school does not pay enough attention to e-sports, and there is no relatively formal organization and management of participants, which has led to many human resources problems waste [3].

In order to cater to the trend of e-sports development, vigorously develop e-sports business, improve the overall level of e-sports, and enable e-sports activities to develop well in colleges and universities, the current primary task is to deepen the characteristics of students participating in e-sports activities [4]. Among them, the analysis and research on the current situation, development trend, and participation significance of e-sports in colleges and universities are particularly important in order to discover the problems existing in the development of campus e-sports
and put forward reasonable suggestions for the development [5].

As emerging sports, e-sports are mainly participated by the younger generation, which has the characteristics of younger and younger age. E-sports can exercise people’s thinking ability, psychological pressure resistance, unity and cooperation, hand-eye coordination, and so on. It can also make the younger generation have the awareness of abiding by the rules in the process of participating in e-sports [6]; trained participants have a fair and open, never admit defeat, pursue a stronger competitive spirit, and have a positive impact on the lives of participants. Many colleges and universities have successively opened related majors in e-sports. Although e-sports is popular in the world, the related research and guiding theories on how to cultivate e-sports talents are rare [7].

Different scholars have different views on the attributes and characteristics of e-sports. Literature [8] proposed that “e-sports include three basic characteristics: one is electronics, the other is competitive sports, and the third is a confrontation between people. At the same time, e-sports are divided into virtualized e-sports sports and fictionalized sports.” Literature [9] pointed out that “the most fundamental characteristics of video games that distinguish them from other artificial games are: virtual environment, absence of the body and artificial intelligence,” emphasizing the main position of electronic communication technology in e-sports. Scholar Yang Fang believes that “e-sports should return to the essence of games, and games to competitive sports are based on the evolution trend of play-game-competitive sports” and based on the development process of traditional competitive sports, puts forward a plan for the development of e-sports. Jia Peng and Yao Jiaxin believe that e-sports has great characteristics: the diversity of functional structure requirements, the full expansion of self-awareness, the complexity of sports information pattern recognition, the agility of information processing efficiency, and the accuracy of intuitive thinking and decision-making. Sex analyzes and clarifies the various attributes of e-sports from many aspects [10].

The discussion on the attributes of e-sports is still going on. Based on the current research, it can be determined that the two essential attributes of e-sports are electronic interaction and confrontational competition. Without electronic interaction, it becomes traditional competitive sports; it becomes a video game, so the two are interdependent and indispensable. With the development of electronic interaction technology, various forms of e-sports have emerged [11].

Event services are mainly engaged in e-sports referees, coaches, club operation and management, game commentary, data and tactical analysis, and so on. Practitioners need to have data analysis capabilities, management capabilities, and commentary capabilities. The production and broadcast of the event include content production and external dissemination, mainly involving the design of live content and promotion plans, venue layout, equipment debugging, video data collection, postprocessing, background data analysis, and so on. The practitioners should have journalism, communication, broadcasting, TV technology, and other related professional abilities [12].

Since the e-sports industry is an emerging industry, most employees are not from e-sports majors and have not received a complete and systematic e-sports theoretical education, but nearly 90% of the employees believe that the e-sports industry needs prejob training [13]. Judging from the current situation of the development of the entire industry, it is undoubtedly the most attractive option to work for game manufacturers, but it is difficult for game manufacturers to absorb more human resources without major business adjustments. Therefore, the need to train practitioners in support organizations around e-sports events becomes more obvious [14]. For example, training content production capabilities (reporters, screenwriters, copywriters, and anchors) requires a professional background in journalism and communication; training event support capabilities (coaches, data analysts, nutritionists, and brokers) requires sports and information technology. Professional background: training public relations, marketing capabilities (products, business, brand marketing, and media), requires a professional background in marketing and management [15].

E-sports self-media is still a media, and you must have the ability to report news, or you can dig deep into a vertical field, such as specializing in video commentary of games, specializing in game clearance strategies, specializing in sharing game skills, and so on. After all, hot spots can bring traffic. WeMedia is a personalized media with social attributes; it communicates with users; and it has its own distinct character orientation [16]. To be a self-media, you should also have strong analytical skills and be able to interpret a topic or special event from a unique or professional perspective. Current e-sports professional ability training pathways.

1.1. Current E-Sports Professional Ability Training Pathways. Most training institutions in society position themselves as training professional players but basically lack training resources. Training institutions do not have coaches, data analysts, or club managers, and it is difficult for the trained people to find a suitable position in the e-sports circle. Rather than cultivating professional skills, it is better to make money from e-sports hot spots. Money has no intention or inability to contribute to the development of the e-sports industry [17].

At present, the main e-sports talents are cultivated by e-sports companies and e-sports clubs. The club mainly trains professional players, coaches, and data analysts in order to achieve better results in the league. Game companies train referees, game developers, commentators, and other related talents to ensure the healthy development of the e-sports industry [18]. An analysis of the revenue structure of the e-sports industry can help us see the e-sports industry more transparently. The truly profitable institutions are still game manufacturers, which continuously create market value through development and operation. In the context of the continuous development and popularization of the video game industry, competition has become a starting point for expanding influence and creating new
commercial value. The comprehensive development of competitive value is inseparable from the promotion of surrounding formats, and new jobs such as video, live broadcast, and commentary emerge in an endless stream [19].

This paper combines the intelligent gesture recognition technology to construct an e-sports training system and uses the player’s gesture recognition to judge the player’s training effect to improve the e-sports training effect.

2. Intelligent Gesture Recognition

2.1. Gesture Intelligent Positioning. The structural framework of the gesture autonomous localization algorithm is shown in Figure 1.

Monocular visual-inertial odometry uses a pure camera in the front end for motion estimation. The algorithm firstly extracts the features of the image information collected by the camera, then uses the optical flow method to track the feature points, and finally uses PnP (Perspective-n-Point) to perform motion estimation on the tracked feature points. Then, the algorithm eliminates the mismatched point pairs through random sampling consistency (RANSAC) and uses nonlinear optimization to optimize the pose. The front-end process is shown in Figure 2.

2.2. SLC-Harris Feature Extraction. The feature is the digital expression of the object in the image, and the image can be quantitatively analyzed by extracting the feature. Commonly used feature extraction methods mainly include SIFT algorithm, SURF algorithm, and ORB algorithm.

The traditional Harris algorithm calculates the angular responsivity as shown below. It is mainly based on the weighted summation of the squared and multiplied gradients of all pixels in the window.

\[ R = |C| - k \times (\text{trace}(C))^2. \]  

(1)

Among them, there are

\[ |C| = \lambda_1 \times \lambda_2, \]  

(2)

\[ \text{trace}(C) = \lambda_1 + \lambda_2. \]  

(3)

In formula (1), \( k \) is a constant ranging from 0.04 to 0.06, and both \( \lambda_1 \) and \( \lambda_2 \) in formula (2) represent eigenvalues.

For a grayscale image, the value of any point \((x, y)\) in the integral image \( i(x, y) \) refers to the sum of all grayscale values from the upper left corner of the image to the area where this point is located, as shown in Figure 3.

The calculation formula of pixels in the rectangular window is as follows:

\[ \tilde{i}(x, y) = \sum_{x \leq x', y \leq y'} I(x', y'). \]  

(4)

The most complex calculation in the Harris algorithm is the calculation of diagonal responsivity. The original calculation method causes the calculation overlap between each pixel in the integration window, resulting in high computational complexity. For this, the gradient values in \( g_x, g_y \) and \( g_{xx}, g_{yy}, g_{xy} \) are used to integrate the image to speed up the calculation of the angular responsivity. The calculation formula is as follows:

\[ \tilde{i}_{xx}(x, y) = \sum_{x \leq x', y \leq y'} g_x^2(x', y'), \]  

\[ \tilde{i}_{yy}(x, y) = \sum_{x \leq x', y \leq y'} g_y^2(x', y'), \]  

\[ \tilde{i}_{xy}(x, y) = \sum_{x \leq x', y \leq y'} g_x(x', y') g_y(x', y'). \]  

(5)

Efficient nonmaximum suppression (E-NMS) is used to efficiently extract unique feature locations for each corner region, and the region thresholds are compared using image patches instead of pixels. The principle is shown in Figure 4.

2.3. KLT Optical Flow Tracking. After the key points are extracted, the optical flow method is used to calculate the minimum photometric error by establishing an error model. This method does not need to calculate descriptors or feature point matching, which will greatly save the amount of calculation.

The basic idea of LK optical flow is to assume that the optical flow in the local neighborhood of a pixel is invariant, and based on this assumption, construct a least-squares problem about the optical flow of the neighborhood pixels.

First, it is assumed that the light intensity of the pixel in each frame of the image is constant. According to this, for the pixel located at \((x, y)\) at time \(t\), moving to \((x + dx, y + dy)\) at time \(t + dt\), there are
Given another basic assumption of LK optical flow, the displacement of pixels in adjacent images is small; the Taylor expansion of formula (6) is

\[ I(x + dx, y + dy, t + dt) \approx I(x, y, t) + \frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt. \]  

(7)

Combining the above formulas and dividing by \( dt \) into both sides of the formula, we get:

\[ \frac{\partial I}{\partial x} \frac{dx}{dt} + \frac{\partial I}{\partial y} \frac{dy}{dt} = \frac{\partial I}{\partial t}, \]  

(8)

where \( dx/dt \) represents the motion speed of the pixel on the x-axis, \( dy/dt \) represents the motion speed of the pixel on the y-axis, and the two speeds are recorded as \( u \) and \( v \), respectively. At the same time, \( \partial I/\partial x \) represents the gradient value of the image in the x-axis direction at the pixel point; \( \partial I/\partial y \) represents the gradient value in the y-axis direction at the pixel point; and \( \partial I/\partial t \) represents the derivative value of the image in the \( t \) direction, which are denoted as \( I_x, I_y, \) and \( I_t \), respectively. Therefore, formula (8) can be written in matrix form as follows:

\[ \begin{bmatrix} I_x & I_y \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = -I_t. \]  

(9)

Finally, according to the third basic assumption of the LK optical flow method, adjacent pixels in the same image plane have similar motion; a \( w \times w \) size window is defined. According to the same motion of all pixels in the window, \( w^2 \) formulas can be listed; the overdetermined formulas can be constructed; and the motion parameters of the center point can be obtained by the least square method. Accordingly, its formula can be expressed as follows:

\[ \left( \sum I_x I_x, \sum I_x I_y \right) \begin{bmatrix} u \\ v \end{bmatrix} = -\left( \sum I_y I_t \right). \]  

(10)

Each image frame is downsampled by pyramid layering, and multilevel pyramids are established.

\[
i^l(x, y) = \frac{1}{4} i^{l-1}(2x, 2y) + \frac{1}{8} \left( i^{l-1}(2x-1, 2y) + i^{l-1}(2x+1, 2y) \right) + \frac{1}{16} \left( i^{l-1}(2x, 2y-1) + i^{l-1}(2x, 2y+1) \right) + \frac{1}{16} \left( i^{l-1}(2x+1, 2y-1) + i^{l-1}(2x+1, 2y+1) \right),
\]  

(11)
where \( L \) represents the \( L \)th layer image.

The algorithm calculates the value of the bottom layer from top to bottom according to the Gaussian pyramid and calculates the pixel value near the edge of the image based on the following formulas:

\[
\begin{align*}
I^{L-1}(-1, y) & \approx I^{L-1}(-1, y), \\
I^{L-1}(x, -1) & \approx I^{L-1}(x, 0), \\
I^{L-1}(n_x^{L-1}, y) & \approx I^{L-1}(n_x^{L-1} - 1, y), \\
I^{L-1}(x, n_y^{L-1}) & \approx I^{L-1}(x, n_y^{L-1} - 1), \\
I^{L-1}(n_x^{L-1}, n_y^{L-1}) & \approx I^{L-1}(n_x^{L-1} - 1, n_y^{L-1} - 1).
\end{align*}
\] (12)

The camera motion pose is estimated using SFM in the vision front end. For a monocular camera, the camera pose can be estimated by the geometric relationship between two points in different locations in real space and their projected points on their respective imaging planes. As shown in Figure 5, \( P \) is any point in the three-dimensional space, and its coordinates are \([X, Y, Z]^T\); \( O_1 \) and \( O_2 \) are the optical centers of the two camera positions. \( p_1 \) and \( p_2 \) are the projection points of point \( P \) on the imaging plane \( I_1 \) and the imaging plane \( I_2 \), respectively. According to the pixel positions of the two matched point pairs \( p_1 \) and \( p_2 \), the essential matrix \( E \) and the fundamental matrix \( F \) can be obtained.

According to the camera imaging model, we assume that \( K \) is the camera internal parameter matrix, and \( R \) and \( t \) represent the rotation matrix and translation vector from plane \( I_1 \) to plane \( I_2 \), and the following formula can be obtained:

\[
s_1p_1 = KP, s_2p_2 = K(RP + t). \tag{13}
\]

Homogeneous coordinate transformation and normalization between 2D and 3D, we can get

\[
x_1 = K^{-1}p_1, x_2 = K^{-1}p_2, \tag{14}
\]

where \( x_1 \) and \( x_2 \) represent the coordinates of pixels \( p_1 \) and \( p_2 \) in the normalized plane, respectively. The algorithm combines formulas (13) and (14) and multiplies by \( x_1^Tt^\wedge \) to obtain the essential matrix \( E \) and the fundamental matrix \( F \), which can be sorted out as follows:

\[
p_2^TK^{-T}t^\wedge RK^{-1}p_1 = 0, \\
E = t^\wedge R, F = K^{-T}EK^{-1}, \tag{15}
\]

where \( t^\wedge \) represents the antisymmetric matrix.

When there are more than eight sets of point pairs such as \( p_1 \) and \( p_2 \), the eight-point method can be used to construct a linear formula system for the simplified formula, and then the unique solution of \( R \) and \( t \) can be obtained.

\[
x_2^TEx_1 = p_1^TFp_1 = 0. \tag{16}
\]

When the monocular camera recovers the pose through the epipolar geometric relationship, the obtained translation is the normalized value, which has no practical significance.

In order to obtain the depth information on feature points, triangulation needs to be introduced. We assume that \( s_1 \) and \( s_2 \) represent the depth of the two feature points; we can get

\[
s_2x_2 = s_1Rx_1 + t. \tag{17}
\]

The feature point depth values \( x_2^T \) and \( x_1^T \) can be obtained by left-multiplying formula (17) by \( s_1 \) or \( s_2 \), respectively, as follows:

\[
s_1x_1^TRx_1 + x_1^Tt = 0, \\
s_2x_1^TRx_2 - x_1^Tt = 0. \tag{18}
\]

When the positions of multiple points in space are known, the camera pose can be estimated by the PnP algorithm. Common PnP algorithms include P3P, DLT, and BA optimization. Among them, the P3P algorithm is the most common method. The algorithm needs to know at least three points and their projection points on the camera imaging plane. Then, the camera pose can be estimated by solving the relationship between point pairs according to the similar triangle principle and the cosine theorem. A schematic diagram of the P3P relationship is shown in Figure 6.

The coordinate system convention is as follows: the world coordinate system is represented by \((·)^W\), and \((·)^b\) and \((·)^c\) represent the IMU coordinate system and the camera coordinate system, respectively. The relationship between the coordinate systems is shown in Figure 7. \((·)^r\) represents the visual reference frame in the sliding window, which is independent of the IMU measurement and can represent any frame in the visual structure. \((·)^s_{\bf q}\) represents the transformation from the IMU coordinate system to the world coordinate system; \(b_k\) represents the IMU frame of the \( k \)th image; \((·)^c_{\bf v}\) represents the transformation from the camera coordinate system to the visual reference system; and \(c_k\) represents the camera frame of the \( k \)th image. \((·)^r\) represents the measured value and parameter estimation value of the sensor; \((·)^{\hat{r}}\) represents the latest scale parameter of the sliding window; and the rotation can be represented by the rotation matrix \( R \) and the quaternion \( q \).

\[
\begin{align*}
&b_k \rightarrow (·)^b_{\bf v}, \\
&c_k \rightarrow (·)^c_{\bf v}, \\
&\hat{R} \rightarrow (·)^{\hat{r}}_{\bf v}, \\
&\hat{q} \rightarrow (·)^{\hat{r}}_{\bf q}.
\end{align*}
\]
2.4. IMU Preintegration. The sampling frequency of the camera used in this paper is 20 Hz, and the sampling frequency of the IMU is 200 Hz. It can be seen that the frequency of the IMU is much higher than that of the image. In order to avoid the repeated integration phenomenon caused by the frequency change of the visual frame optimization state caused by the high sampling rate of the IMU, a preintegration technique is used for all IMU sampling data aggregated into a relative motion constraint through a preintegration technique. The principle of preintegration is shown in Figure 8.

In Figure 8, from top to bottom are the time scale line, the number of image frames generated, the number of image key frames generated, the number of IMU samples, and the IMU preintegration value.

The measurement error of the system is mainly affected by bias random walk $b$ and white noise $\eta$, and other errors such as the Markov process are ignored. Then, the measurement model of the accelerometer and gyroscope in the IMU can be expressed by the following formula:

$$\ddot{\omega}^b = \omega^b(t) + b_\omega + \eta_\omega, \quad \ddot{a}^b = q_b^w (a^w - g^w) + b_a + \eta_a,$$

where $\ddot{\omega}^b, \ddot{a}^b$, and $a^w(t)$ represent the measured value and real value of angular velocity and acceleration, respectively; $b_\omega, b_a, \eta_\omega,$ and $\eta_a$ represent the random noise and measurement white noise of angular velocity and acceleration, respectively; and $q_b^w$ is the rotation matrix transformed from the world coordinate system to the IMU coordinate system.

White noise obeys a Gaussian distribution, that is, $\eta_\omega \sim N(0, \sigma_\omega^2), \eta_a \sim N(0, \sigma_a^2)$. The derivative of random walk noise also obeys the Gaussian distribution, that is, $\eta_b \sim N(0, \sigma_b^2), \eta_\omega \sim N(0, \sigma_\omega^2)$. The differential kinematic formulas for $P, V, Q$ (representing the position, velocity, and rotation expressed in quaternions, respectively) versus time can be written as follows:

$$P^w_k = P^w_1 + \int_{t \in [k,k+1]} \left[ R^w_r (\bar{a}_t - b_a) - g^w \right] dt,$$

$$V^w_k = V^w_1 + \int_{t \in [k,k+1]} \left[ R^w_r (\bar{a}_t - b_a) - g^w \right] dt,$$

$$Q^w_k = Q^w_1 + \int_{t \in [k,k+1]} \Omega (\bar{\omega}_t - b_\omega, q^b_\omega) dt,$$

where $\Omega$ represents quaternion multiplication.

Through the above derivative relationship, the position, velocity, and rotation at time $k+1$ can be obtained from the position, velocity, and rotation at time $k$ and by integrating the measured values of the IMU over time $\Delta t_k$. The continuous integration formula for $P, V, Q$ is as follows:

$$P^w_k = P^w_1 + \int_{t \in [k,k+1]} \left[ R^w_r (\bar{a}_t - b_a) - g^w \right] dt,$$

$$V^w_k = V^w_1 + \int_{t \in [k,k+1]} \left[ R^w_r (\bar{a}_t - b_a) - g^w \right] dt,$$

$$Q^w_k = Q^w_1 + \int_{t \in [k,k+1]} \Omega (\bar{\omega}_t - b_\omega, q^b_\omega) dt,$$

where $\bar{a}_t$ and $\bar{\omega}_t$ represent the acceleration and angular velocity measured in the IMU coordinate system, respectively. $\Delta t_k$ represents the time difference from the $k$th frame to the $k+1$ frame. $R^w_r$ represents the rotation matrix from
the world coordinate system to the IMU coordinate system. Because the measured \( \ddot{a}_t \) and \( \dot{\omega}_t \) belong to the IMU coordinate system, in order to transform the IMU measured value to the world coordinate system, the rotation matrix needs to be left-multiplied. \( \Omega(\omega) \) means quaternion right multiplication; \( \omega_x \) means antisymmetric matrix in quaternion multiplication (\( \omega \) means the imaginary part value of quaternion). We assume that the quaternion is \( q = [x \ y \ z \ s] = [\omega \ s] \); then we have

\[
\Omega(\omega) = \begin{bmatrix}
-\omega_x & -\omega_y & -\omega_z & 0 \\
\omega_y & -\omega_z & \omega_x & 0 \\
\omega_z & \omega_x & -\omega_y & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

\( \omega_x = \begin{bmatrix} \omega_z & 0 & -\omega_x \\ 0 & \omega_x & \omega_y \\ -\omega_y & \omega_x & 0 \end{bmatrix} \) (22)

By observing the continuous integral formula of PVQ, it can be seen that the current state is recursively obtained from the state of the previous time, and the estimated value is constantly changing. This will cause the IMU measurements to be repropagated, causing the velocity and rotation to be reinitialized after each nonlinear optimization iteration, resulting in a higher computational cost. Therefore, the optimization variables are separated from the IMU preintegration terms of the two key frames, and the rotation matrix \( R_{t_0}^w \) of the world coordinate system to the IMU coordinate system can be obtained by simultaneously left-multiplying the left and right sides of the continuous integration formula of PVQ:

\[
R_{t_0}^w \phi_{t_0}^w = R_{t_0}^w (p_{t_0}^w + v_{t_0}^w \Delta t_k - \frac{1}{2}g^w \Delta t_k^2) + \alpha_{t_0}^{b_{k+1}},
\]

\[
R_{t_0}^w \nu_{t_0}^w = R_{t_0}^w (v_{t_0}^w - g^w) + \beta_{t_0}^{b_{k+1}},
\]

\[
q_{t_0}^w \otimes q_{t_0}^w = Y_{t_0}^{b_{k+1}}.
\]

The image frames \( b_k \) and \( b_{k+1} \) of two consecutive moments are given, and the linear acceleration and angular velocity are preintegrated in the local coordinate system \( b_k \) to obtain

\[
\alpha_{t_0}^{b_{k+1}} = \int_{t_{k+1} \mid t_k} R_{t_0}^{R_{t_0}^w(b_k)} (\tilde{a}_t - b_a) dt^2,
\]

\[
\beta_{t_0}^{b_{k+1}} = \int_{t_{k+1} \mid t_k} R_{t_0}^{R_{t_0}^w(b_k)} (\tilde{\omega}_t - b_\omega) dt,
\]

\[
y_{t_0}^{b_{k+1}} = \int_{t_{k+1} \mid t_k} (1/2)\Omega(\tilde{\omega}_t - b_\omega) y_{t_0}^{b_{k+1}} dt,
\]

where \( \alpha_{t_0}^{b_{k+1}}, \beta_{t_0}^{b_{k+1}}, y_{t_0}^{b_{k+1}} \) represent the relative pose, velocity, and rotation constraints, respectively, and are also the relative motion of \( b_{k+1} \) to \( b_k \). It can be seen that they are only related to \( \tilde{a}_t \) and \( \tilde{\omega}_t \) in \( b_k \) and \( b_{k+1} \), and they have nothing to do with the initial position and velocity of coordinate system \( b_k \).

Therefore, the preintegration formula is redescribed, in terms of \( \alpha_{t_0}^{b_{k+1}}; \) it is related to \( \bar{a}_t \) and \( \bar{\omega}_t \) of the IMU; and \( \bar{a}_t \) and \( \bar{\omega}_t \) are also variables that need to be optimized. When the bias change is small, \( \alpha_{t_0}^{b_{k+1}}, \beta_{t_0}^{b_{k+1}}, y_{t_0}^{b_{k+1}} \) are adjusted according to their first-order approximations to the bias.

\[
\alpha_{t_0}^{b_{k+1}} = \bar{\alpha}_{t_0}^{b_{k+1}} + f_{t_0}^a \Delta b_a + f_{t_0}^a \Delta b_\omega,
\]

\[
\beta_{t_0}^{b_{k+1}} = \bar{\beta}_{t_0}^{b_{k+1}} + f_{t_0}^\beta \Delta b_a + f_{t_0}^\beta \Delta b_\omega,
\]

\[\begin{aligned}
Y_{t_0}^{b_{k+1}} &= Y_{t_0}^{b_{k+1}} \otimes \begin{bmatrix} 0 \\
\frac{1}{2}f_{t_0}^\gamma \Delta b_\omega \end{bmatrix},
\end{aligned}\]

where \( f_{t_0}^a \) and \( f_{t_0}^\beta \) are the block matrices in \( f_{t_0}^a \) and \( f_{t_0}^\beta \) are the block matrices in \( f_{t_0}^{\gamma, b_k} \).

There are errors in the integral values of the IMU at different times, and the errors at time \( t \) are mainly related to the measured values of \( \alpha_{t_0}^{b_{k+1}}, \beta_{t_0}^{b_{k+1}}, \gamma_{t_0}^{b_{k+1}}, b_a \), and \( b_\omega \) at time \( t \). The following definition is given to represent the error vector:

\[
\delta \omega_{t_0}^{b_{k+1}} = \begin{bmatrix} \delta \alpha_{t_0}^{b_{k+1}} \\
\delta \beta_{t_0}^{b_{k+1}} \\
\delta b_a \\
\delta b_\omega \end{bmatrix}.
\]

The derivation is based on the derivative of the error term kinetic formula. First, two concepts are introduced: true and nominal, where true represents the real measurement value containing noise and nominal represents the theoretical value without noise, and \( \delta \) represents the measurement error; there are

\[
\delta \alpha = \bar{\alpha}_{true} - \bar{\alpha}_{nominal},
\]

\[
\delta \beta = \bar{\beta}_{true} - \bar{\beta}_{nominal}.
\]

Among them, there are:

\[
\hat{\beta}_{true} = R_t^{b_k} [\bar{a}_t - \eta_a - b_a - \delta b_a - (\bar{a}_t - b_a) \delta \theta],
\]

\[
\hat{\beta}_{nominal} = R_t^{b_k} (\bar{a}_t - b_a).
\]

Combining the above formulas, we can obtain

\[
\delta \hat{\beta} = -R_k^{b_k} \bar{a}_t - b_a \delta \theta - R_k^{b_k} \delta b_a - R_k^{b_k} \eta_a.
\]

The derivation of \( \delta \theta \) is as follows, and according to the formula in the literature, it can be known that

\[
q_{true} = q_{nominal} \otimes \delta q.
\]

In this paper, according to the noise model and bias, we can get

\[
\delta \theta = -\bar{a}_t - b_a \delta \theta - \eta_a - \delta b_a.
\]

In summary, the derivative of the IMU measurement error term at time \( t \) can be as follows:
The pure visual initialization method lacks absolute scale information. Therefore, the value estimated by the visual SFM is aligned with the IMU after preintegration to estimate the true scale. Visual-inertial alignment initialization is mainly to solve the following problems, including the initialization of gyroscope bias, the initialization of velocity, gravitational acceleration, and scale.

The first is to initialize the gyroscope. The gyroscope bias can be obtained from two consecutive key frames with known orientations, considering two consecutive frames \( b_k \) and \( b_{k+1} \) in the sliding window, and \( \Delta q_{k}^{v} \) and \( \Delta q_{k}^{w} \) represent the rotations obtained from the pure visual sliding window optimization, respectively. Linearize the IMU preintegration term for the gyroscope bias and minimize the following function:

\[
\min_{\delta b_w} \sum_{k \in \mathcal{B}} \left\| \Delta q_{k+1}^{v} \otimes \Delta q_{k}^{w} \otimes \delta b_w \right\|^2. \tag{38}
\]

Among them, there are:

\[
\Delta q_{k+1}^{v} \otimes \Delta q_{k}^{w} = 1_{0}. \tag{39}
\]

In formula (42), \( \mathcal{B} \) represents all the frames in the window, and the product of the two quaternions indicates that the camera rotates from the \( k \)th frame to the \( k+1 \)th frame, and the gyroscope rotates from the \( k \)th frame to the \( k \)th frame, and the optimized objective function is

\[
\Delta q_{k+1}^{v} \otimes \Delta q_{k}^{w} \otimes \delta b_w = \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \tag{40}
\]

The algorithm takes \( \tilde{y} \) into the formula and multiplies the inverse moment ordering of the relative constraints obtained from the preintegration to the left on both sides of formula (40) and obtains by Cholesky decomposition (multiplying the transpose of \( J_{b_w}^{T} \) on both sides of the formula):

\[
J_{b_w}^{T} \delta b_w = 2J_{b_w}^{T} \left( \Delta q_{k+1}^{v} \otimes \Delta q_{k}^{w} \otimes \delta b_w \right). \tag{41}
\]

In this way, the initial calibration value of the gyroscope bias \( \delta b_w \) can be estimated, and then the IMU preintegration terms \( \Delta q_{k+1}^{v} \otimes \Delta q_{k}^{w} \otimes \delta b_w \) are corrected with the new gyroscope bias.

The second is the initialization of velocity, gravitational acceleration, and scale. The initialized state vector is as follows:

\[
\chi_{i} = \left[ v_{i}^{v}, \hat{v}_{i}^{v}, \ldots, v_{n}^{v}, g_{i}^{v}, s \right]. \tag{42}
\]

where the state vector \( v_{i}^{v} \) represents the speed of the visual coordinate system of the \( k \)th frame image, \( g_{i}^{v} \) represents the gravity vector in the visual coordinate system, and \( s \) is the estimated scale parameter. To sum up, the dimension of \( \chi_{i} \) is \( 3(n+1)+3+1 \). The constraint relationship between the scale parameter and the speed of the visual SFM is as follows:

\[
\eta_{v} \otimes \eta_{b}\]

We set \( F = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \cdot \frac{\delta z_{t}^{h}}{\delta t} \left( 1 + F_{1} \right) \delta z_{t}^{h} + \left( G_{t} \delta t \right) n_{t}. \tag{33}
\]

According to the definition of the derivative, the prediction formula of the mean is as follows:

\[
\delta z_{t+\delta t}^{h} = \frac{\delta z_{t+\delta t}^{h} - \delta z_{t}^{h}}{\delta t}. \tag{34}
\]

According to the error value at the current moment, the mean and covariance at the next moment can be predicted. The prediction formula for covariance is as follows:

\[
P_{t+\delta t}^{h} = \left( 1 + F_{1} \delta t \right) P_{t}^{h} \left( 1 + F_{1} \delta t \right)^{T} + \left( G_{t} \delta t \right) Q \left( G_{t} \delta t \right)^{T}, \tag{35}
\]

where \( P_{t}^{h} \) represents the initial value of the iteration and its value is zero and \( Q \) represents the diagonal covariance matrix of the noise term as follows:

According to formula (35), the iterative formula of the Jacobian matrix Jacobian can be obtained as follows:

\[
J_{t+\delta t} = \left( 1 + F_{1} \right) J_{t}, \tag{36}
\]

where the iterative initial value of the Jacobian matrix \( J_{t} \) is \( J \).

2.5. Sliding Window Initialization. When the camera extrinsic parameter matrix \( \begin{bmatrix} p_{c}^{0}, p_{c}^{1} \end{bmatrix} \) is known, the pose obtained by the initialization of the monocular camera is transformed from the visual coordinate system to the IMU coordinate system to obtain the following formula:

\[
q_{i}^{v} = q_{i}^{v} \otimes q_{i}^{w}, \]

\[
s_{i}^{p_{i}} = s_{i}^{p_{i}} \otimes q_{i}^{w}, \tag{37}
\]

where \( s \) is the translation factor obtained by visual initialization, which has no real information.
\[
\begin{bmatrix}
\alpha^b_{k+1} \\
\beta^b_{k+1} \\
\gamma^b_{k+1}
\end{bmatrix} = \begin{bmatrix}
H^b_{k+1}
\end{bmatrix} \begin{bmatrix}
\delta^b_{k+1} + \delta^b_{k+1} \approx \delta^b_{k+1}
\end{bmatrix},
\]

where \(q^V_{k+1}, q^V_{k+1}, q^V_{k+1}\) are all obtained from visual SFM. \(q^V_{k+1}\) and \(q^V_{k+1}\) are mutually square matrices. The following linear least squares problem is constructed to complete the initialization of velocity, gravitational acceleration, and scale:

\[
\min_{\xi} \sum_{k=\infty} \|z^b_{k+1} - H^b_{k+1} z^b_{k+1}\|^2.
\]

### 2.6. Monocular Visual Inertial Coupling Nonlinear Optimization

When coupling the visual constraint value and the IMU constraint value, the data of the inertial sensor should be introduced first, and the constraint value of the IMU on the state should be added to the optimized state vector. Then, nonlinear optimization is performed within a sliding window, and all state vectors of the sliding window are as follows:

\[
\begin{align*}
\chi &= [x_0, x_1, \ldots, x_n, \lambda_0, \lambda_1, \ldots, \lambda_m], \\
x_k &= [p^w_{k}, v^w_{k}, q^w_{k}, b_k, b_g], k \in [0, n], \\
x^b_k &= [p^b, q^b, \chi^b_k],
\end{align*}
\]

where \(x^b_k\) represents the state of the IMU when the \(k\)th image is captured. There are \(n + 1\) states in the sliding window, and each state contains the position, velocity, and rotation in the world coordinate system, and the IMU offsets in the IMU coordinate system. \(\lambda_m\) represents the inverse depth information of the \(m\)th 3D point, and \(\chi^b\) represents the external parameter from the camera to the IMU. The objective function is

\[
\min_{\chi} \left\{ \left\| r_p - H_p \chi \right\|^2 + \sum\left\| r_p \left( \tilde{z}^b_{k+1, \chi} \right) \right\|^2_p + \sum_{(l, j) \in C} \rho \left( \left\| r_c(\tilde{z}^f_{j}) \right\|^2_p c^j \right) \right\},
\]

where \(H_p\) is the Huber norm, which is defined as follows:

\[
\rho(s) = \begin{cases} 
1, & s \geq 1, \\
2\sqrt{s} - 1, & s < 1.
\end{cases}
\]

In formula (52), \(\| \cdot \|_p\) represents the Mahalanobis distance weighted by the covariance matrix \(P\), and \(j, r_p(z^b_{k+1}), r_p(z^f_{j}), r_c(\tilde{z}^f_{j}), X\) represent the marginalized prior information, the IMU measurement residual, and the visual reproject error, respectively. \(\mathcal{B}\) is the set of all IMU measurement frames, and \(\mathcal{C}\) is the set of visual features in the sliding window.

According to the Gauss–Newton method, the incremental method can be used to calculate the minimum value of the Gaussian objective function, as follows:

\[
\min_{\Delta X} \left\| r_{\mathcal{B}} \left( \tilde{z}^b_{k+1, \chi} + r_{\mathcal{B}} \right) \right\|_{\mathcal{B}}^2
\]

\[
= \min_{\Delta X} \left\| r_{\mathcal{B}} \left( \tilde{z}^b_{k+1, \chi} + f^b_{k+1} \right) \right\|_{\mathcal{B}}^2
\]

\[
\text{where } f^b_{k+1} \text{ is the Jacobian matrix of the error term } r_{\mathcal{B}} \text{ with respect to all state vectors } \chi.
\]

The algorithm differentiates the above formula from \(\Delta X\) and then sets its derivative to 0, and the formula for the increment \(\Delta X\) can be calculated as follows:

\[
f^b_{k+1} T P^b_{k+1} f^b_{k+1} + f^b_{k+1} T P^b_{k+1} f^b_{k+1} \Delta X = -f^b_{k+1} T P^b_{k+1} f^b_{k+1} r_{\mathcal{B}}
\]

\[
\text{The overall incremental equation of the objective function is as follows:}
\]

\[
\min_{\Delta X} \left( \begin{bmatrix} H_p + \sum_{j \in \mathcal{B}} f^b_{k+1} P^b_{k+1} r_{\mathcal{B}} \end{bmatrix} + \sum_{j \in \mathcal{C}} f^b_{k+1} P^b_{k+1} c^j \right) \Delta X = b_p + \sum_{j \in \mathcal{B}} f^b_{k+1} P^b_{k+1} r_{\mathcal{B}} + \sum_{j \in \mathcal{C}} f^b_{k+1} P^b_{k+1} c^j
\]

where \(P^b_{k+1}\) represents the covariance of the IMU preintegrated noise term and \(P^b_{k+1}\) represents the visually observed noise covariance. When the noise covariance \(P^b_{k+1}\) of the IMU is larger, the inverse of \(P^b_{k+1}\), that is, its information matrix is smaller, which means that the IMU observations are not as reliable as the visual observations. Formula (50) can be simplified to

\[
(\Lambda_p + \Lambda_B + \Lambda_C) \Delta X = b_p + b_B + b_C,
\]

where \(\Lambda_p, \Lambda_B, \text{ and } \Lambda_C\) represent the Hessian matrix. Using the perturbation method to calculate, we can get

\[
J = \frac{\partial r}{\partial \Delta X} = \lim_{\Delta X \to 0} \frac{r(X + \delta X) - r(X)}{\delta X},
\]

where \(\delta X\) represents the small disturbance of the state vector instead of the incremental \(\Delta X\), \(\oplus\) represents the disturbance of the state vector.

The continuous preintegration formula is derived in the IMU preintegration, and the IMU residuals are as follows:
According to the above formula, it can be seen that the optimization variables of the IMU residual mainly include the position, rotation, speed, and inertia bias at the \( i \) and \( j \) times:

\[
\begin{bmatrix}
\dot{p}_{b_{k_{1}}}, q_{b_{k_{1}}}, v_{b_{k_{1}}}, b_{\omega_{k_{1}}}, b_{\alpha_{k_{1}}}, b_{\beta_{k_{1}}}, \ldots
\end{bmatrix},
\begin{bmatrix}
\dot{p}_{b_{k_{2}}}, q_{b_{k_{2}}}, v_{b_{k_{2}}}, b_{\omega_{k_{2}}}, b_{\alpha_{k_{2}}}, b_{\beta_{k_{2}}}, \ldots
\end{bmatrix}.
\]

(54)

To calculate the Jacobian matrix, perturbation is added to each optimization variable to obtain

\[
\begin{bmatrix}
\delta \dot{p}_{b_{k_{1}}}, \delta \dot{q}_{b_{k_{1}}}, \delta \dot{v}_{b_{k_{1}}}, \delta b_{\omega_{k_{1}}}, \delta b_{\alpha_{k_{1}}}, \delta b_{\beta_{k_{1}}}, \ldots
\end{bmatrix},
\begin{bmatrix}
\delta \dot{p}_{b_{k_{2}}}, \delta \dot{q}_{b_{k_{2}}}, \delta \dot{v}_{b_{k_{2}}}, \delta b_{\omega_{k_{2}}}, \delta b_{\alpha_{k_{2}}}, \delta b_{\beta_{k_{2}}}, \ldots
\end{bmatrix}.
\]

(55)

Taking the partial derivatives for the above disturbance variables, respectively, we get

\[
J[0] = \begin{bmatrix}
\frac{\partial \mathbf{R}_{B}}{\partial \dot{p}_{b_{k_{1}}}}, & \frac{\partial \mathbf{R}_{B}}{\partial \dot{q}_{b_{k_{1}}}}, & \frac{\partial \mathbf{R}_{B}}{\partial \dot{v}_{b_{k_{1}}}}, & \frac{\partial \mathbf{R}_{B}}{\partial b_{\omega_{k_{1}}}}, & \frac{\partial \mathbf{R}_{B}}{\partial b_{\alpha_{k_{1}}}}, & \frac{\partial \mathbf{R}_{B}}{\partial b_{\beta_{k_{1}}}}, & \ldots
\end{bmatrix},
\]

\[
J[1] = \begin{bmatrix}
\frac{\partial \mathbf{R}_{B}}{\partial \dot{p}_{b_{k_{2}}}}, & \frac{\partial \mathbf{R}_{B}}{\partial \dot{q}_{b_{k_{2}}}}, & \frac{\partial \mathbf{R}_{B}}{\partial \dot{v}_{b_{k_{2}}}}, & \frac{\partial \mathbf{R}_{B}}{\partial b_{\omega_{k_{2}}}}, & \frac{\partial \mathbf{R}_{B}}{\partial b_{\alpha_{k_{2}}}}, & \frac{\partial \mathbf{R}_{B}}{\partial b_{\beta_{k_{2}}}}, & \ldots
\end{bmatrix},
\]

(56)

\[
J[2] = \begin{bmatrix}
\frac{\partial \mathbf{R}_{B}}{\partial \dot{p}_{b_{k_{1}}}}, & \frac{\partial \mathbf{R}_{B}}{\partial \dot{q}_{b_{k_{1}}}}, & \frac{\partial \mathbf{R}_{B}}{\partial \dot{v}_{b_{k_{1}}}}, & \frac{\partial \mathbf{R}_{B}}{\partial b_{\omega_{k_{1}}}}, & \frac{\partial \mathbf{R}_{B}}{\partial b_{\alpha_{k_{1}}}}, & \frac{\partial \mathbf{R}_{B}}{\partial b_{\beta_{k_{1}}}}, & \ldots
\end{bmatrix},
\]

\[
J[3] = \begin{bmatrix}
\frac{\partial \mathbf{R}_{B}}{\partial \dot{p}_{b_{k_{2}}}}, & \frac{\partial \mathbf{R}_{B}}{\partial \dot{q}_{b_{k_{2}}}}, & \frac{\partial \mathbf{R}_{B}}{\partial \dot{v}_{b_{k_{2}}}}, & \frac{\partial \mathbf{R}_{B}}{\partial b_{\omega_{k_{2}}}}, & \frac{\partial \mathbf{R}_{B}}{\partial b_{\alpha_{k_{2}}}}, & \frac{\partial \mathbf{R}_{B}}{\partial b_{\beta_{k_{2}}}}, & \ldots
\end{bmatrix}.
\]

The visual residual is a reprojection error, which represents the difference between the estimated value and the observed value of the feature point in the normalized camera coordinate system. The small receiver camera used in this paper belongs to the fisheye camera model and belongs to the fisheye with a large viewing angle, so its projection on the unit sphere needs to be considered, as shown in Figure 9.

Through the unit spherical projection model illustrated in the figure above, the value of the visual residual is decomposed into two directions. The final visual residual model looks like this:

\[
r_{c}(\hat{z}_{j}^{c}, \chi) = \left[ b_{1}, b_{2} \right]^{T} \cdot \left( \hat{p}_{l}^{c} - \frac{\hat{p}_{l}^{c}}{||\hat{p}_{l}^{c}||} \right),
\]

\[
\hat{p}_{l}^{c} = \pi_{c}^{-1} \left( \begin{bmatrix} \hat{u}_{l}^{c} \\ \hat{v}_{l}^{c} \end{bmatrix} \right),
\]

\[
P_{l}^{c} = R_{c} \left[ R_{c}^{v} \left( R_{c}^{b} \left( \frac{1}{\chi} \hat{F}_{l} + F_{l}^{b} \right) + \hat{p}_{b}^{w} \right) - \hat{p}_{l}^{w} \right] - F_{l}^{c},
\]

(57)

where \( \hat{p}_{l}^{c} \) and \( p_{l}^{c} \) represent the estimated and observed coordinates of the landmark point 1 in the \( j \)-th frame image under the normalized camera coordinate system.
The optimization variables of the visual residual are as follows:

\[ [p_{b_i}^w, q_{b_i}^w], [p_{b_i}^w, q_{b_i}^w], [p_{c_j}^b, q_{c_j}^b], \lambda_j, \]  

where \( \lambda_j \) represents the inverse depth value when the landmark point 1 is first observed by the \( j \)-th image. The inverse depth is used as the optimization variable because the inverse depth satisfies the Gaussian distribution, and it can reduce the parameter variables in the actual computations.

**Table 1:** Gesture recognition effect of the system.

| Number | Gesture recognition | Number | Gesture recognition | Number | Gesture recognition |
|--------|---------------------|--------|---------------------|--------|---------------------|
| 1      | 86.095              | 13     | 87.363              | 25     | 86.633              |
| 2      | 88.782              | 14     | 89.124              | 26     | 90.414              |
| 3      | 90.695              | 15     | 90.055              | 27     | 86.697              |
| 4      | 86.325              | 16     | 89.968              | 28     | 86.916              |
| 5      | 86.705              | 17     | 88.482              | 29     | 89.003              |
| 6      | 89.557              | 18     | 91.278              | 30     | 88.444              |
| 7      | 86.623              | 19     | 89.215              | 31     | 89.911              |
| 8      | 91.281              | 20     | 86.272              | 32     | 90.202              |
| 9      | 88.050              | 21     | 91.253              | 33     | 89.507              |
| 10     | 87.621              | 22     | 86.770              | 34     | 88.130              |
| 11     | 88.498              | 23     | 86.843              | 35     | 91.326              |
| 12     | 86.304              | 24     | 90.614              | 36     | 86.846              |

**Table 2:** The application effect of the method proposed in this paper in the e-sports training system.

| Number | Training effect | Number | Training effect | Number | Training effect |
|--------|-----------------|--------|-----------------|--------|-----------------|
| 1      | 82.685          | 13     | 82.234          | 25     | 79.365          |
| 2      | 80.315          | 14     | 78.537          | 26     | 81.356          |
| 3      | 79.700          | 15     | 81.913          | 27     | 78.790          |
| 4      | 78.179          | 16     | 82.938          | 28     | 79.700          |
| 5      | 82.942          | 17     | 78.167          | 29     | 78.194          |
| 6      | 80.821          | 18     | 78.250          | 30     | 78.253          |
| 7      | 81.176          | 19     | 81.512          | 31     | 80.133          |
| 8      | 78.183          | 20     | 80.076          | 32     | 82.668          |
| 9      | 82.636          | 21     | 81.925          | 33     | 83.747          |
| 10     | 81.613          | 22     | 82.857          | 34     | 82.354          |
| 11     | 80.514          | 23     | 82.951          | 35     | 78.416          |
| 12     | 80.913          | 24     | 79.066          | 36     | 83.128          |

Figure 11: Overall system framework.
optimization process. According to the above formula, by adding disturbance to each optimization variable, the following Jacobian is obtained:

\[
    J[0]^{3 \times 7} = \left[ \frac{\partial r_c}{\partial p_b}, \frac{\partial r_c}{\partial q_c} \right],
\]

\[
    J[1]^{3 \times 7} = \left[ \frac{\partial r_c}{\partial p_b}, \frac{\partial r_c}{\partial q_c} \right],
\]

\[
    J[2]^{3 \times 7} = \left[ \frac{\partial r_c}{\partial p_b}, \frac{\partial r_c}{\partial q_c} \right],
\]

\[
    J[3]^{3 \times 1} = \frac{\partial r_c}{\partial l_1}.
\]

3. E-Sports Training System Based on Intelligent Gesture Recognition

This paper combines the finger joints and the sensor in the data glove to demarcate the movement of the finger joints. This paper mainly considers the distal phalanx of the thumb (TDP) and the proximal joint proximal phalanx of the thumb (TPP) of the thumb as shown in Figure 10 and the changes in the intermediate joints middle phalanges (MP) and proximal phalanges (PP) of the remaining four fingers.

This paper combines the algorithm part of the second part to construct the e-sports training system, and the overall framework of the system is shown in Figure 11.

The simulation of the system proposed in this paper is carried out through the MATLAB platform, and the gesture recognition effect of the system and the application effect in the e-sports training system are evaluated, and the obtained results are shown in Tables 1 and 2.

It can be seen from the above research that the electronic economic training system based on intelligent gesture recognition proposed in this paper has certain effects.

Data Availability

The labeled data set used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This study was sponsored by Shandong Sport University.

References

[1] D. J. D. Q. Martins, L. C. L. Moraes, and W. M. Júnior, “COVID-19 impacts on school sports events: an alternative through E-sports,” Managing Sport and Leisure, vol. 27, no. 1-2, pp. 45–49, 2022.
[2] M. Yasumoto and T. Teraoka, “Physical e-Sports in VAIR Field system,” in SIGGRAPH Asia 2019 XR, pp. 31–33, ACM, New York, NY, U.S.A, 2019.
[3] U. Hussain, B. Yu, G. B. Cunningham, and G. Bennett, “‘I can be who I Am when I play tekken 7’; E-sports women participants from the islamic republic of Pakistan,” Games and Culture, vol. 16, no. 8, pp. 978–1000, 2021.
[4] B. B. Romanenko, C. C. Pyatitsotska, B. C. Ashanin, and A. M. Yefremenko, "Analysis of sensorimotor abilities and nervous system properties of players of different e-sports disciplines. Scientific journal of the National Pedagogical University named after MP Drahomanov. Series 15," Scientific and pedagogical problems of physical culture (physical culture and sports), vol. 11, no. 143, pp. 125–131, 2021.
[5] W. Lu, G. Xu, X. Gong, and Q. Chen, “On the interdisciplinary education among undergraduate students majoring in E-sports,” in Proceedings of the 2021 International Conference on Culture-oriented Science & Technology (ICCST), pp. 63–66, IEEE, Beijing, China, November 2021.
[6] M. Jasny, “Doping in e-sports. An empirical exploration and search for sociological interpretations,” Acta Universitatis Lodzienis. Folia Sociologica, vol. 26, no. 75, pp. 85–99, 2020.
[7] S. Akkaya, B. S. Gezer, and M. Kapidere, “A multi-directional assessment related to E-sports as A new game experience field and socialising tool,” Atatürk üniversitesi sosyal bilimler Enstitüsü Dergisi, vol. 25, no. 3, pp. 968–988, 2021.
[8] D. Чалык, “E-sports broadcasts as a new format in journalism,” Al-Farahi kazakh national university, vol. 52, no. 2, pp. 202–209, 2019.
[9] R. Ardianto, T. Rivanie, Y. Alkhalifi, F. S. Nugraha, and W. Gata, “Sentiment analysis on E-sports for education curriculum using naive Bayes and support vector machine,” Jurnal Ilmu Komputer dan Informasi, vol. 13, no. 2, pp. 109–122, 2020.
[10] S. Kwag, W. J. Lee, and Y. D. Ko, “Optimal seat allocation strategy for e-sports gaming center,” International Transactions in Operational Research, vol. 29, no. 2, pp. 783–804, 2022.
[11] M. Marelić, D. Vukušić, and D. Vukušić, “E-sports: definition and social implications,” Exercise and Quality of Life, vol. 11, no. 2, pp. 47–54, 2019.
[12] A. Çetin and M. Coşkuner, “A conceptual overview of E-sports tourism as a new trend in the tourism industry,” Turk Turizm Arastirmalari Dergisi, vol. 3, no. 1, pp. 28–34, 2021.
[13] R. F. Marta, A. A. Prasetya, B. Laurensia, S. Stevani, and K. L. Syarnubi, "Imbalance identity in E-sports news intersectionality on covid-19 pandemic situation," *Jurnal ASPI-KOM*, vol. 5, no. 2, pp. 206, 2020.

[14] D. H. Shin, J. S. Kim, and C. W. Kim, "A Study on subtitle synchronization calibration to enhance hearing-impaired persons' viewing convenience of e-sports contents or game streamer contents," *Journal of Korea Game Society*, vol. 19, no. 1, pp. 73–83, 2019.

[15] J. Hwang, "Cheating in E-sports: a proposal to regulate the growing problem of E-doping," *Northwestern University Law Review*, vol. 116, no. 5, pp. 1283–1318, 2022.

[16] A. Ersin, H. C. Tezeren, N. O. Pekyavas et al., "The relationship between reaction time and gaming time in e-sports players," *Kinesiology*, vol. 54, no. 1, pp. 36–42, 2022.

[17] J. Hou, X. Yang, and E. Panek, "How about playing games as a career? The evolution of E-sports in the eyes of mainstream media and public relations," *International Journal of Sport Communication*, vol. 13, no. 1, pp. 1–21, 2020.

[18] C. Tjokrodinata, C. R. A. Bangun, F. Dinansyah, and A. R. Farmita, "Gamers with different ability and the role of EAI (E-Sports ability Indonesia)," *Jurnal Interaksi: Jurnal Ilmu Komunikasi*, vol. 6, no. 1, pp. 52–66, 2022.

[19] B. Göktaş, "Level of awareness of e-sports concept and its effect on image of sports clubs in Turkey," *Uluslararası Toplumsal Bilimler Dergisi*, vol. 3, no. 1, pp. 68–85, 2019.