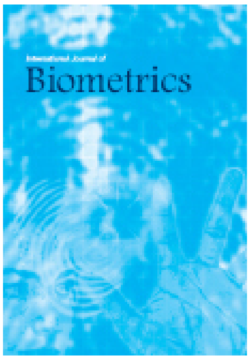


International Journal of Biometrics

ISSN online: 1755-831X - ISSN print: 1755-8301
https://www.inderscience.com/ijbm

An action recognition of track and field athletes based on Gaussian mixture model

Qin Yang, Zhenhua Zhou

DOI: 10.1504/IJBM.2025.10064813

Article History:

Received:	21 September 2023
Last revised:	25 December 2023
Accepted:	29 March 2024
Published online:	30 April 2025

An action recognition of track and field athletes based on Gaussian mixture model

Qin Yang

School of Physical Education,
Hunan City University,
Yi'yang, 413000, China
Email: 283489415@qq.com

Zhenhua Zhou*

School of Physical Education,
Hunan Normal University,
Chang'sha, 410012, China
Email: zzh5126@163.com
*Corresponding author

Abstract: To solve the problem of low recognition accuracy caused by the complexity of individual actions in track and field in the past, a method of action recognition for track and field athletes based on Gaussian mixture model was proposed. First, the data is analysed by the interaction of spatiotemporal features. Secondly, a low-pass filter is used to eliminate the impact of noise on the data to reduce the calculation loss. On the basis of pre-processing data, Hilbert Huang transform (HHT) was used for feature extraction to capture and understand athletes' motion features more accurately, thus significantly improving the accuracy of movement recognition. Then, the Gaussian mixture model is used to model the characteristic parameters, determine the number of mixed components and initialise the model parameters, and complete the movement recognition of track and field athletes. The experimental results show that the traditional method has high computational loss and low recognition accuracy, while the proposed method has very low computational loss and the highest recognition accuracy can reach 98%. The comparison shows that this method has the advantages of low computational complexity, high accuracy and good recognition performance.

Keywords: Gaussian mixture model; GMM; interaction of spatiotemporal features; action data; low pass filter; athlete movement recognition.

Reference to this paper should be made as follows: Yang, Q. and Zhou, Z. (2025) 'An action recognition of track and field athletes based on Gaussian mixture model', *Int. J. Biometrics*, Vol. 17, No. 3, pp.292–310.

Biographical notes: Qin Yang has Master's degree, a Lecturer, his research direction is a School Physical Education Teaching and Sports Training. He led a project by the Hunan Provincial Federation of Social Sciences: his research project on the development strategy of road running events based on urban marketing theory, he participated in multiple other projects, he won multiple provincial and municipal awards, and mainly he studied basic knowledge and skills in sports technology, sports ergonomics, psychology, and other aspects during teaching. He received training in physical function, skill technology,

strategy and tactics, psychology, and intelligence, and understood the competition rules and policies of sports competitions policies and regulations, training and management of athletes in sports competitions, as well as judgments and judgments of sports events, etc.

Zhenhua Zhou is a Professor, and a Master's Supervisor, majoring in Physical Education and Training (Athletics), and graduated as a Doctoral student in Physical Education and Training at Beijing Sport University in 2013. National level track and field judges, national level social sports instructors, and fitness coaches. He mainly engaged in research in areas such as competitive sports training, integrated talent cultivation between industry and education, integrated sports and medical services, and event management.

1 Introduction

The rapid development of contemporary technology has provided vast space for the research of motion recognition for track and field athletes. The advancement of computer vision, image processing, and machine learning technologies has made action recognition more feasible and accurate. At the same time, the advancement of modern technological equipment has also made data collection and analysis more convenient and efficient. For a highly technical sport like athletics, the requirements for athletes' movements are very high (Zhao et al., 2022; Li et al., 2022). The significance of motion recognition for track and field athletes lies in helping coaches or athletes analyse key motion details during the exercise process, guiding training, strengthening correct technical movements, helping athletes better grasp technical essentials, and improving their technical level. Through action recognition, we can find out the possible bad action habits or unreasonable sports patterns of athletes in the process of sports, correct these problems in time, help reduce the probability of Sports injury, and improve the health and safety of athletes (Alsawadi et al., 2023). In addition, action recognition research can also provide more objective and scientific evaluation criteria for track and field competition, and promote the development and progress of High-performance sport. Therefore, many relevant researchers have conducted research on the motion recognition of track and field athletes.

Han et al. (2021) proposed to identify athletes' movements from the perspective of deep learning. This method, on the basis of analysing the advantages and disadvantages of the classical Convolutional neural network and Recurrent neural network algorithms of deep learning, combined with the recognition and analysis of human movements under complex scenes, complex movements, and fast movements, an improved convolutional network and optimised neural network hybrid neural network algorithm were proposed. Han and Fan (2019) proposed a method for identifying incorrect actions in sports training based on 3D modelling detection, which achieves recognition of athlete actions through 3D modelling detection. From the simulation results, it can be seen that using this method can achieve a recognition rate of 86%. Liang (2021) proposes an automatic detection system for athlete training movements based on posture estimation. In terms of system hardware, RL2048P image sensor is used as the image capture device, and the system power circuit is optimised. In terms of software, gamma algorithm is used to obtain motion direction information, and OpenPose software is used to perform secondary processing on the obtained images to obtain athlete joint motion information and complete training motion detection. Combine the optimised system hardware with software to achieve automatic detection of athlete training movements. Lin (2021) proposed a research on intelligent recognition of motion based on laser sensors. The motion signal data of the moving target is collected using laser sensors, and wavelet analysis is used to denoise the collected data. After pre-processing, segmentation algorithms are used to find segmentation points and perform segmentation processing on the denoised motion. Time domain signal analysis is used to extract the feature values of the initial signals of each motion after segmentation, input into the BP neural network and output motion recognition results through the BP neural network classifier. The above methods have difficulties in accurately modelling these complex action patterns, resulting in poor recognition performance and low accuracy.

Gaussian mixture model (GMM) can model complex actions. Athletes' movements often involve coordinated movements of multiple joints and body parts. The GMM can capture this complexity and provide accurate action classification and recognition results. GMM can adapt well to multimodal data and fuse different types of information, improving the accuracy and robustness of action recognition. At the same time, a probability value can be assigned to each action category, indicating the likelihood that the action belongs to that category, which can provide more comprehensive and interpretable action recognition results. In this regard, this paper studies the action recognition of track and field athletes based on the GMM.

2 Extraction of movement features of athletes in track and field

2.1 Interaction of spatiotemporal features of action data

After collecting the movement data of track and field athletes, the original data obtained may contain a lot of information, including the athletes' posture, movement track, acceleration, speed, etc. However, the original data itself may be very large and complex, which is difficult to analyse and understand directly. In this case, spatiotemporal feature interaction can be used as a part of the data processing and analysis stage to help extract discriminative and expressive features from the collected data. By associating and interacting with video spatial and temporal features, the details and contextual information of actions can be captured, thereby improving the accuracy and effectiveness of subsequent tasks (Akbar et al., 2022).

Spatiotemporal feature interaction is the association of video spatial and temporal features to obtain effective comprehensive feature representations. A single spatial or temporal contact information cannot fully represent the action content in the video. Inspired by the attention mechanism, an attention based interaction method is proposed, which can effectively fuse two types of information. The temporal feature information exchange branch of the STI module has transformed the original spatial features into temporal features. Another branch of the STI module is responsible for preserving the original spatial feature information. Firstly, transform the temporal feature x and spatial feature X of two feature maps with the same dimension size into one-dimensional lengths, so that each pixel in the feature map is arranged in the same dimension, as shown in the following equation:

$$R^l = [r_1', r_2', r_3', ..., r_n'], n = T \times H \times W \tag{1}$$

$$X^l = [x_1, x_2, x_3, ..., x_n], n = T \times H \times W \tag{2}$$

In the formula, R and X represent the temporal and spatial features after deformation, l is the dimension, T is the number of frames, and H and W represent the height and width, respectively.

Use the activation function sigmoid to activate R, and the activated R can be regarded as the self-attention information graph. The interaction of spatiotemporal features is realised by point multiplication of the self-attention information graph and spatial features. Finally, the obtained spatiotemporal features are restored to their original shape and input into the next interaction unit.

Spatiotemporal feature interaction can enable neural networks to focus on spatial features related to moving objects, injecting temporal information into spatial feature learning. The entire process of spatiotemporal interaction is as follows:

$$A^l = sigmoid(R^l) \tag{3}$$

$$Y_{n \times c}^l = A^l \otimes X^l \tag{4}$$

$$Y_{H \times W \times c}^l = reshape(Y_{n \times c}^l), n = T \times H \times W \tag{5}$$

Among them, A^l represents the temporal feature of activation, $Y_{n \times c}^l$ represents the spatial time of activation, and $\otimes$ represents the multiplication of corresponding elements. In this structure, by overlaying multiple interaction units, spatial and temporal features are interacted multiple times. Therefore, these two types of information can learn from each other, learn from each other, and integrate with each other many times. Integrating and interacting temporal and spatial features to obtain a more comprehensive and accurate data representation, the fused data representation can comprehensively consider the spatiotemporal action information of athletes, and has better discrimination and expression ability (Zhao et al., 2022).

2.2 Action data pre-processing based on low-pass filters

When obtaining movement data of track and field athletes, electromagnetic interference from the transmitting and receiving equipment itself, multipath components in the experimental environment, and interference from testers can cause anomalies in the collected movement data of track and field athletes, thereby affecting the classification accuracy of the movement recognition model. Figure 1 is the collected hand action amplitude data. It can be seen from the observation of Figure 1 that there are multiple outlier in the original data of channel state information (CSI). According to relevant literature analysis, it can be concluded that the frequency distribution of CSI amplitude fluctuations caused by track and field athletes' exercise is between 0–50 Hz, while features above this frequency or outliers deviating from the waveform belong to environmental interference. The existence of these Outlier will weaken the action characteristics and reduce the accuracy of the model's action recognition. In order to eliminate these outlier, Butterworth low-pass filter is selected and reasonable parameters are set to eliminate the outlier in CSI data. The processed CSI data can still maintain the integrity of the data and the original feature information (Zhan et al., 2022).

Input the collected CSI action data into Butterworth low-pass filter to eliminate high-frequency outlier. The amplitude frequency response function of Butterworth low-pass filter is as follows:

$$|H(j\Omega)|^2 = \frac{1}{1+\left(\frac{\Omega}{\Omega_c}\right)^{2N}} \tag{6}$$

In the formula, N represents the order of the low-pass filter, Ω_c represents the cut-off frequency, and Ω represents the overall frequency.

Figure 1 Original CSI data image (see online version for colours)

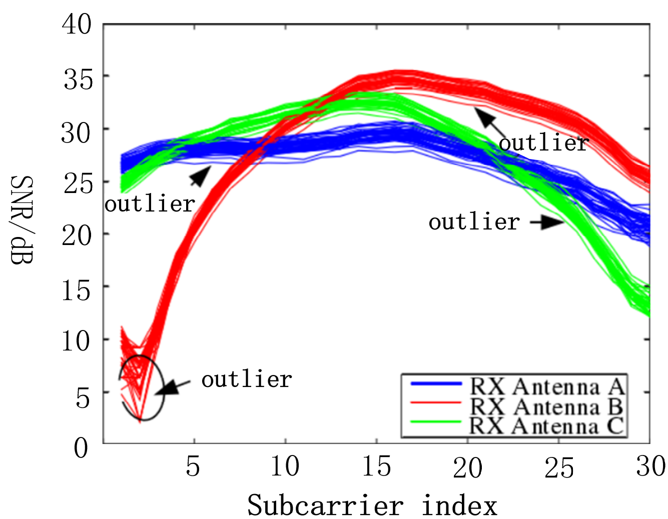
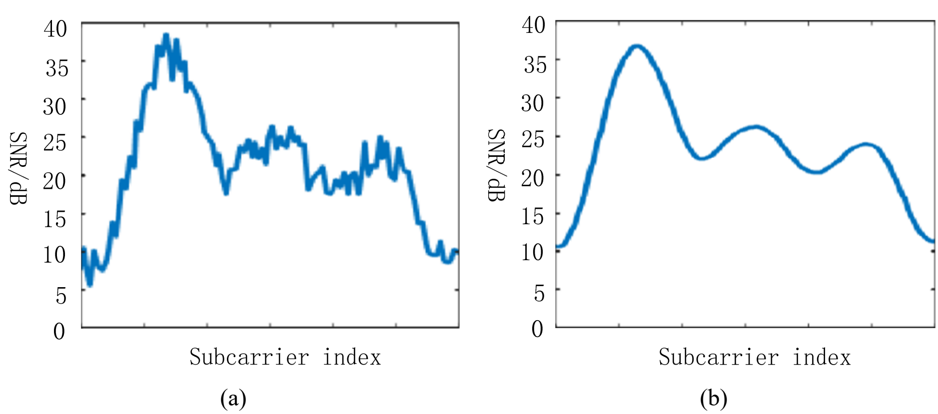


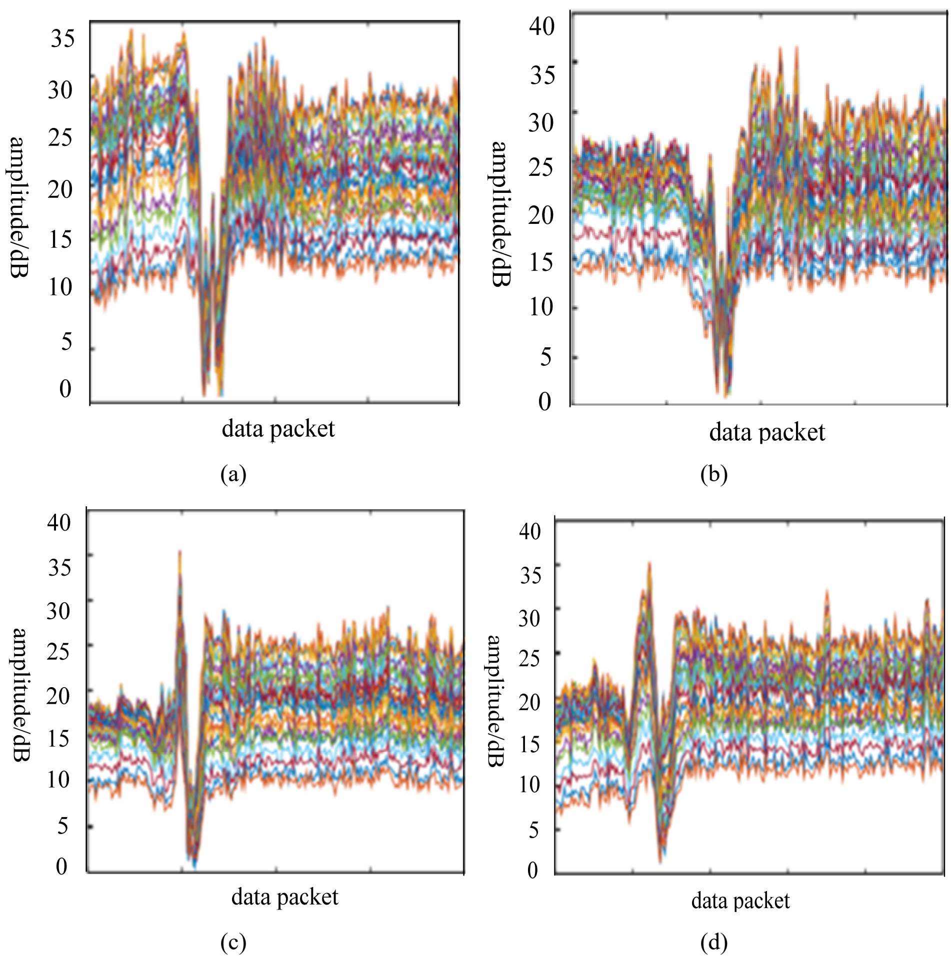
Figure 2 Comparison results of CSI data processed by Butterworth low-pass filtering, (a) outlier data (b) Butterworth low-pass filter processing (see online version for colours)



In order to filter out high-frequency outlier in CSI action data, Butterworth low-pass filter can be used and parameters can be set as required. In this case, set the passband cut-off frequency to 30 Hz and the stopband cut-off frequency to 80 Hz. Such setting can effectively filter out high-frequency noise or Outlier higher than 30 Hz, thus improving the quality and accuracy of data.

As shown in Figure 2(a) and 2(b) respectively represent the original CSI data with outlier and the data image filtered by Butterworth low-pass filtering, it can be observed that the waveform after filtering is more smooth and stable, which is conducive to subsequent feature extraction and classification recognition, and improves the recognition accuracy of falls (Zhou et al., 2020).

Figure 3 Amplitude waveform diagram, (a) fall (b) sit down (c) squat down (d) stand up (see online version for colours)



2.3 *Analysis of CSI Characteristics of Athletes' movements*

The behavioural actions during different competition events will generate CSI amplitude signal waveforms with different intensities and contours. In order to verify the mapping relationship between hand movements and CSI amplitude, in Figures 3(a), 3(b), 3(c) and

3(d) respectively display the CSI amplitude waveform of left and right hand movements and foot movements, indicating the one-to-one mapping relationship between several movements and CSI signal amplitude. It can be seen from these figures that the amplitude waveform change trend of several actions is roughly the same, which will affect the recognition accuracy of track and field athletes' actions, and it is easy to misjudge other actions as suspected violations. Therefore, this paper conducts HHT processing on the pre-treated CSI fusion signal to obtain the instantaneous amplitude and Instantaneous phase and frequency, then create a feature matrix, and finally input the classifier, which is conducive to identifying the actions of track and field athletes (Meng et al., 2022).

2.4 Feature extraction algorithm based on HHT

Based on the above content, the HHT algorithm is used for feature extraction (Ge et al., 2022). The specific implementation steps of the HHT algorithm are as follows:

1 Input the pre-treated CSI fusion signal $X_{CSI}(t)$, obtain the local maximum and minimum points of the fusion signal, fit the upper and lower envelopes, and obtain the local mean value $m(t) = \frac{X_h(t) + X_l(t)}{2}$. Where, $X_h(t)$ represents the upper envelope, which is composed of local maximum, and $X_l(t)$ is the lower envelope composed of local minimum, which are all obtained by cubic Spline interpolation.

2 Based on the size characteristics of CSI itself, it is decomposed into a group of IMF_s:

$$X_{CSI}(t) = \sum_{j=1}^{n} IMF_j(t) | + r_n(t) \tag{7}$$

In the equation, $r_n(t)$ represents the residual, j represents the number of IMF signals, n represents the highest number of IMF signals, and $IMF_j(t)$ represents the j the IMF signal at time t.

3 Find the remaining signal $H(t) = X_{CSI}(t) - m(t)$, and if the number of extreme points and zero crossing points of H(t) in the fused signal is 0 or ± 1, and the average of the extreme levels of the upper and lower envelope lines is 0, then it can be considered as the first IMF component (Sun et al., 2020). If this condition is not met, the previous steps need to be repeated until the screening stop criteria proposed by Huang are met:

$$sd = \sum_{k=1}^{T} \frac{H_{k+1}(t) - H_k(t)}{H_k^2(t)} \tag{8}$$

In the equation, $0.2 < sd < 0.3$ and $H_{k+1}(t)$ represent the $k + 1$ remaining signal, and $H_k(t)$ represents the k remaining signal.

4 The residual signal $r_n(t)$ is a low-frequency component that can be omitted. Analyse the high-frequency component IMF. Therefore, the signal analysed by Hilbert spectrum is:

$$s(t) = \mathrm{Re} \sum_{k=1}^{n} a_i(t) e^{-j\theta_i(t)} = \mathrm{Re} \sum_{i=1}^{n} a_i(t) e^{-j\int \omega_i(t)dt} \tag{9}$$

In the equation, Re represents a constant independent of the signal, $a_i(t)$ represents the cumulative probability distribution of the signal, $\theta_i(t)$ represents the frame length, ω_i represents the average energy, and E is the weight.

The right side of the above equation represents the Hildert spectrum, denoted as $H(\omega, t)$. Therefore, $H(\omega, t)$ can be reliably expressed as the integration of Hilert's Spectrogram in time, that is, the Spectrogram accumulates with the change of time (Chen et al., 2020). This integration operation can provide more comprehensive and comprehensive information, thereby more accurately describing the marginal characteristics of the signal. See the following equation:

$$h(\omega) = \int_{-\infty}^{\infty} H(\omega, t)dt \tag{10}$$

In the formula, $h(\omega)$ represents the distribution of cumulative amplitude corresponding to each frequency, which can accurately reflect the actual frequency value of the original signal; $H(\omega, t)$ represents the judgment threshold.

5 The instantaneous phase and frequency $f_j(t)$ and instantaneous amplitude $a_j(t)$ at any time can be calculated from HT of IMF component:

$$\left[f_j(t), a_j(t)\right] = HT\left(IMF_j(t)\right) \tag{11}$$

In the equation, HT represents the time dimension.

In this way, for each leg action, a feature matrix about it can be obtained through HHT to identify the leg action. The characteristic matrix consists of a group of Instantaneous phase and frequency and a group of instantaneous amplitudes:

$$H_i = \left[f_{1i}, f_{2i}, \dots, f_{ni}, a_{1i}, a_{2i}, \dots, a_{ni}\right] \tag{12}$$

3 Action recognition based on Gaussian mixture model

3.1 Construction of Gaussian mixture model

After the extraction of action feature parameters is completed, the feature parameters need to be modelled. This paper uses the GMM to complete the modelling and recognition of feature parameters (Zhang et al., 2020). The GMM is established, and the training of the GMM is realised by using the action characteristic parameters. Based on the above action feature parameter samples, these samples are divided into M categories. Assuming that the samples of each category follow a Gaussian probability distribution model. So, we can use a mixed Gaussian model to describe the probability distribution of these samples, where each category corresponds to a Gaussian distribution:

$$P\left(H_i \middle| \lambda\right) = \sum_{i=1}^{M} \omega_i P\left(H_i \middle| \lambda_i\right) \tag{13}$$

$$P\left(H_i \middle| \lambda_i\right) = \frac{1}{(2\pi)^{D/2} \left|\sum_i 1/2\right|} \exp\left\{-\frac{1}{2}\left(H_i - \mu_i\right)^T\right\} \sum_i^{-1} \left(H_i - \mu_i\right) \tag{14}$$

where D is the mixing number of the model, H_i is the covariance matrix, and μ_i is the mean.

To facilitate calculation, the covariance matrix uses the diagonal matrix, as shown in equation (15):

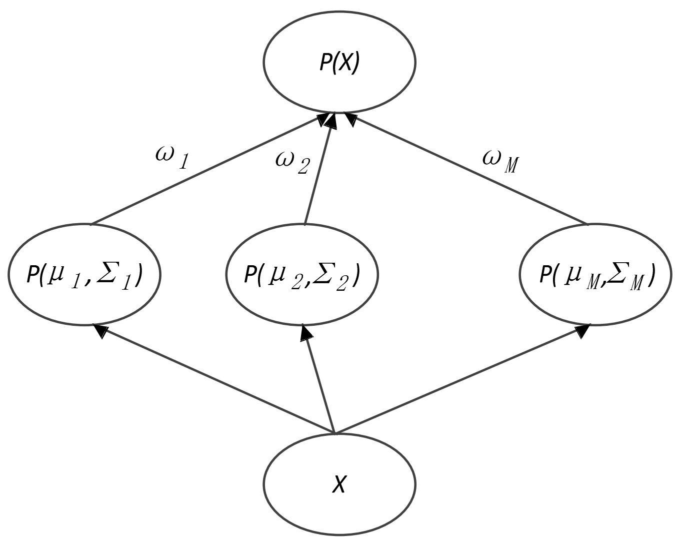
$$\sum_i = diag\left(\sigma_{i1}^2, \sigma_{i2}^2, ..., \sigma_{iD}^2\right) \tag{15}$$

In the formula, $\sigma_i d^2 (d = 1, 2, \cdots D)$ is the variance value of the d dimensional component of the feature parameter sample corresponding to the i Gaussian model.

The GMM can be expressed by equation (16), and its structure is shown in Figure 4.

$$\lambda = \left\{\omega_i, \mu_i, \sum_i (i = 1, 2, ..., M)\right\} \tag{16}$$

Figure 4 Composition structure of Gaussian mixture model



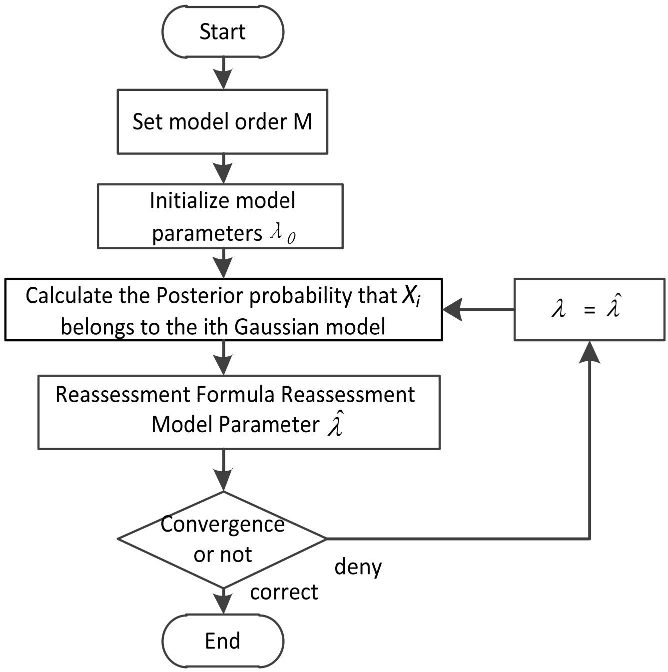
3.2 Gaussian mixture model training

The action signal of track and field athletes can be seen as a Stochastic process. The above feature parameters are composed of the feature vectors of each frame of track and field athletes' action signals extracted by HHT. According to the training data of the feature parameters of each tested athlete, a group of parameters λ of the GMM is obtained to represent the corresponding tested track and field athletes. The process of estimating GMM model parameters based on EM algorithm is as follows 5.

1 It is unknown which Gaussian distribution sample H_t comes from. The observation data $Y = \{Y_1, Y_2, \ldots, Y_T\}$ is introduced, where Y_t represents H_t coming from the i Gaussian model, and Y_t takes values from $1 \sim M$. $Y_t = \{Y_{t1}, Y_{t2}, \ldots, Y_{tM}\}$. $Y_{ti}(t = 1, 2, \ldots T, i = 1, 2, \ldots, M)$ is defined as follows:

$$Y_{ti} = \begin{cases} 1, & \textit{The t observation data comes from the i Gaussian model} \\ 0, & \textit{other} \end{cases} \tag{17}$$

Figure 5 Parameter estimation flowchart



Complete data $Z = (H_i, Y)$ is composed of known data H_i and observation data Y. The Likelihood function of complete data can be expressed as:

$$P\left(H_i, Y|\lambda\right) = \prod_{i=1}^{T} P\left(H_t, Y_t|\lambda\right) \tag{18}$$

In the formula, H_t represents the invalid power of the signal, and Y_t represents the frequency component of the signal.

The product of equation (18) points may be extremely small, resulting in overflow of Floating-point arithmetic numbers. Therefore, we usually take the logarithm of the above equation to obtain the logarithmic Likelihood function as follows:

$$\log\left(P\left(H_i, Y|\lambda\right)\right) = \sum_{t=1}^{T} \log\left(P\left(H_i, Y_t|\lambda\right)\right) = \sum_{t=1}^{T} \log\left(\omega_{Yt} P\left(H_t|\lambda_{Yt}\right)\right) \tag{19}$$

where, ω_{Y_t} represents the current estimated value of model parameters, λ_{Y_t} represents the generation identification feature parameter, $P\left(H_t|\ \lambda_{Y_t}\right)$ represents the probability density function of the t Gaussian model, and T represents the number of GMM.

2 E-step, calculate the mathematical expectation of the logarithmic likelihood function.

3 M-step: Maximise the function $Q(\lambda, \hat{\lambda})$ in E-step, take the partial derivative of ω_i, μ_i, and $\sum_i$ and make their values zero to obtain the re estimated values of the model parameters. The weight, mean, and variance re estimation formulas are as follows:

$$\begin{cases} \omega_i^{(k+1)} = \dfrac{1}{T}\sum_{t=1}^{T} \gamma_{ti}^{(k)}, i = 1, 2, ..., M \\ \mu_i^{(k+1)} = \dfrac{\sum_{t=1}^{T} \gamma_{ti}^{(k)} X_t}{\sum_{t=1}^{T} \gamma_{ti}^{(k)}}, i = 1, 2, ..., M \\ \sum_i^{(k+1)} = \dfrac{\sum_{t=1}^{T} \gamma_{ti}^{(k)} \left(X_t - \mu_i\right)^T}{\sum_{t=1}^{T} \gamma_{ti}^{(k)}}, i = 1, 2, ..., M \end{cases} \tag{20}$$

In the equation, $\gamma_{ti}^{(k)}$ represents the critical band curve, and X_t represents the model parameters to be selected.

Repeat steps (2) and (3) until the model converges and outputs the model parameters.

In the process of training the GMM model using the characteristic parameters of track and field athletes' action signals, if the training data is insufficient or the track and field athletes' action signals are severely interfered with, there may be a phenomenon of very small covariance, which will seriously affect the performance and recognition results of the system. Therefore, when using the EM algorithm to train the GMM model, a minimum value of 0.02 is set for covariance in this paper.

3.3 Parameter initialisation of Gaussian mixture model

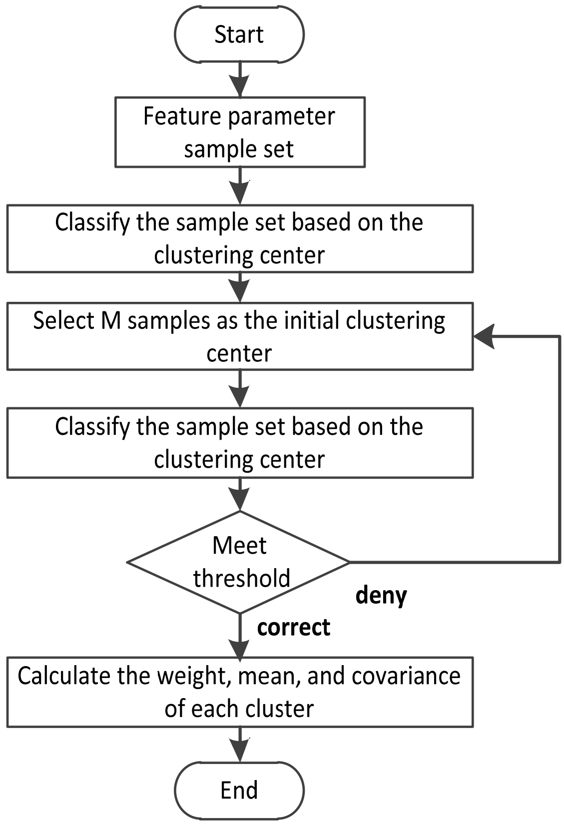
Before using EM algorithm to estimate model parameters, K-means algorithm is used to initialise parameters of GMM, and the operation is carried out through the specific process shown in Figure 6. K-means algorithm can divide the data into K clusters according to the clustering characteristics of the data, and use the centre of these clusters as the initial mean parameter of the GMM. Through this initialisation process, a good starting point can be provided for subsequent EM algorithms, thereby more accurately estimating the parameters of the model.

The specific algorithm process is as follows:

1 Randomly select M sample points from the feature parameter training sample data as the initial centres of each cluster, denoted as $\left(c_1^{(1)}, c_2^{(1)}, ..., c_M^{(1)}\right)$.

2 Calculate the distance between each sample point H_t and M cluster centres separately, classify the sample points into the cluster with the smallest distance from them, set m as the number of cluster centres, then $\left|H_t - c_i^{(m)}\right| \le \left|H_t - c_j^{(m)}\right|, i \ne j,$ and $i, j = 1, 2, \ldots, M$ are satisfied.

Figure 6 Initialisation process



3 Calculate the new cluster centre based on the clustering results, using the following formula:

$$c_i^{(m+1)} = \frac{1}{N_t} \sum_{H_t \in c_i^{(m)}} H_t, i = 1, 2, \ldots, M \tag{21}$$

In the formula, c_i represents the parameter of the Gaussian density function; N_t represents the weight coefficient.

4 Calculate the change in the cluster centre. If $\left| c_i^{(m+1)} - c_i^{(m)} \right| \geq \delta$, repeat step (2) until the change in the cluster centre is less than δ, and the clustering is complete.

After clustering, calculate the initial parameter λ_0 of the GMM as follows:

$$\begin{cases} \omega_i = \dfrac{N_t}{T} \\ \mu_i = \dfrac{1}{N_t} \sum\limits_{H_t \in c_i} H_t \\ \sum\limits_i = \dfrac{1}{N_t} \sum\limits_{H_t \in c_i} \left(H_{td} - \mu_{id}\right)^2, d = 1, 2, ..., D \end{cases} \tag{22}$$

In the formula, H_{td} is the d^{th} dimensional component of feature sample H_t, and μ_{id} is the d^{th} dimensional component of μ_i.

3.4 Athlete movement recognition

The movement characteristic parameters of track and field athletes are set up corresponding GMM for N measured movement, and a model library (λ_n, $n = 1, 2, \ldots, N$) with N GMM is formed. In the process of track and field athletes' action recognition, it is to match the characteristic parameter (H_t') of an unknown athlete's action signal to be recognised in the N actions of the model library with the GMM of N actions in the track and field athletes' action model library. The movement n corresponding to the Maximum a posteriori estimation is the recognition result. According to the Bayesian formula, the Posterior probability can be obtained as shown in equation (23).

$$P(\lambda_n, H_t') = \frac{P(\lambda_n) P\left(H_t' | \lambda_n\right)}{P(H_t')} \tag{23}$$

In the equation, H_t' represents the centre frequency of the critical band, and the identification result is:

$$n^* = \arg \max_{1 \le n \le N} \left[P(\lambda_n, H_t') \right] \tag{24}$$

The Prior probability $P(\lambda_n)$ of the feature parameter H_t' to be identified that belongs to the n motion generally regards the Prior probability of each motion as equal, that is:

$$P(\lambda_n) = \frac{1}{N}, n = 1, 2, ..., N \tag{25}$$

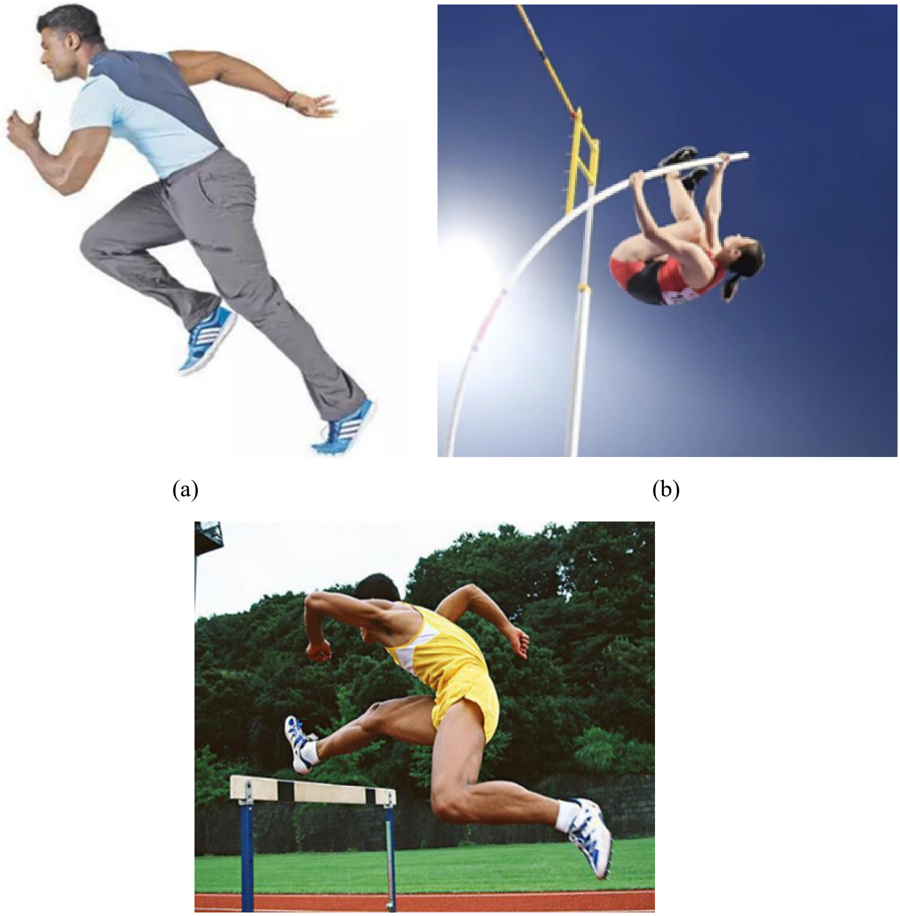
For feature parameter H_t', during the matching process with the models of n motion actions in the model library, $P(H_t')$ is a fixed value. Therefore, equation (22) can be simplified as:

$$n^* = \arg \max_{1 \le n \le N} \left[P\left(H_t' | \lambda_n\right) \right] \tag{26}$$

From the above equation, it can be seen that the final recognition result is the model λ_n corresponding to the maximum value of $P(H_t'|\lambda_n)$.

Based on the above steps, the track and field athletes' action recognition based on the GMM is completed.

Figure 7 Experimental part dataset, (a) action 1 (b) action 2 (c) action 3 (see online version for colours)



(a) (b) (c)

4 Experimental analysis

4.1 Experimental data

In order to verify the performance of the track and field athletes' action recognition method based on the GMM in this paper, experimental analysis was carried out. The experimental research was mainly conducted in the MATLAB environment, using the publicly available competition video dataset from the existing track and field athlete action database as the experimental subjects. The experimental dataset is shown in Figure 7. The experimental parameter settings are shown in Table 1.

Table 1 Experimental parameters

Parameter	*Index*
Programming language	MATLAB
Support vector machine library	MATLAB built-in support vector machine functions
Code writing software platform	Python 4.8
Operating system	Windows10
Memory	5TG
Processor	Intel Xeon series
Graphics card	AMD Radeon

4.2 Performance index

The performance indicators of motion recognition for track and field athletes are generally evaluated using complexity and recognition accuracy.

Complexity calculation formula:

$$L = A\sum_{i=1}^{n} D_i + 2A^2 \tag{27}$$

Recognition accuracy: the accuracy refers to the proportion of samples correctly recognised by the model on the entire dataset. It is one of the most common and intuitive indicators used to evaluate overall recognition performance. The closer it is to 1, the more correctly identified results are in the recognition results, that is, the more accurate the recognition results are. Its calculation formula is:

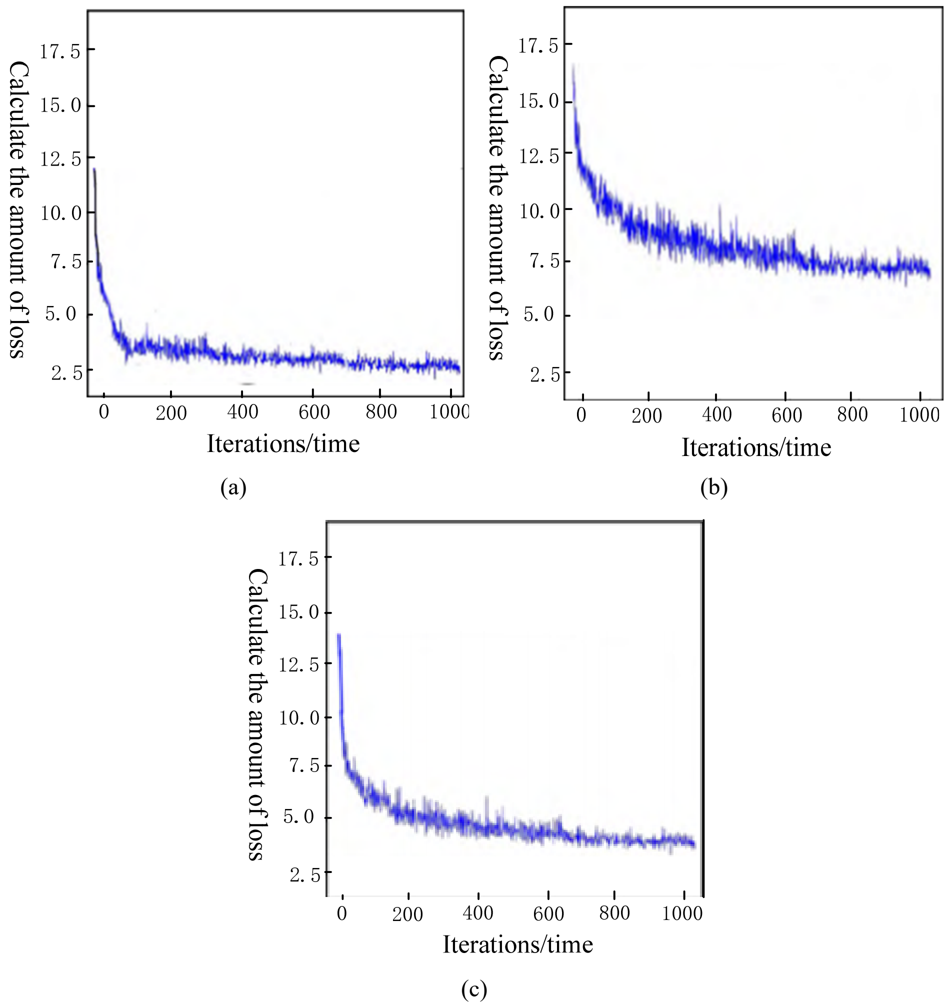
$$P = \frac{TP}{(TP + FP)} \tag{28}$$

Among them, true positive (TP) represents the number of correctly identified athlete actions, and false positive (FP) represents the number of incorrectly identified athlete actions.

4.3 Test and analysis

Using literature Han et al. (2021) and literature Han and Fan (2019) as comparative methods, computational complexity experiments can evaluate the performance of different methods in terms of computational cost and iteration times. By comparing the computational complexity of different methods, one can understand their advantages and disadvantages in resource utilisation and efficiency, and effectively select methods suitable for practical applications. The experimental comparison results are shown in Figure 8.

Figure 8 Comparison results of computational complexity, (a) proposed method (b) Han et al. (2021) method (c) Han and Fan (2019) method (see online version for colours)



According to Figure 8, it can be seen that the calculation loss of this method has stabilised and the loss rate is relatively low in less than 200 cycles; the computational loss of the comparison method is significantly higher than that of the proposed method, and it only stabilises after more than 200 iterations, indicating that the computational complexity of this design method is relatively low and effective.

When selecting recognition accuracy for testing, the experimental comparison results are shown in Table 2.

From the results shown in Table 2, it can be seen that the recognition accuracy of the method proposed in this paper is higher than 90%, with a maximum of 98%. The highest recognition accuracy of the methods proposed in Han et al. (2021) and Han and Fan (2019) is 89% and 88%, respectively. It can be seen that the recognition accuracy of the method proposed in this paper is significantly higher than that of the comparison method. The method presented in this article has a high accuracy in identifying the movements of track and field athletes. Prove that the technical level and application value of the method proposed in this article are high.

Table 2 Comparison results of recognition accuracy

Number of experiments/time	*Han et al. (2021) method/%*	*Han and Fan (2019) method/%*	*Liang (2021) method/%*
10	89	81	98
20	87	84	97
30	88	88	96
40	89	87	94
50	84	83	94
60	85	84	95
70	87	81	98
80	89	82	95
90	88	88	96
100	84	84	97

5 Conclusions

Through the identification and analysis of athletes' movements, coaches and athletes can understand the advantages and disadvantages of sports technology. In this regard, the method of track and field athletes' movement identification based on the GMM is studied. The experimental results indicate that:

1. Compared with the methods in Han et al. (2021) and Han and Fan (2019), the identification method used in this article has extremely low computational loss, indicating low complexity and short time consumption;
2. When using the identification method in this article, its recognition accuracy is compared to the method in Han et al. (2021). The highest recognition accuracy of the method in Han and Fan (2019) is 98%, and the overall recognition accuracy is high and stable, indicating that the proposed method has high recognition accuracy and high recognition performance.

In summary, the proposed method can accurately identify the movements of track and field athletes, which is not only of great significance for track and field sports itself, but also has an important driving role for research in related fields such as sports biomechanics, sports medicine, and sports training. This method can provide more accurate and rich data support for research in these fields, helping to promote theoretical development and practical applications in these fields. Although this study focuses on motion recognition for track and field athletes, the proposed method is universal and can be applied to motion recognition in other sports, and can even be extended to a wider range of fields such as daily activity recognition and gesture recognition. Therefore, this study is not only of great significance for track and field sports, but also has an important driving role in the development of action recognition.

Acknowledgements

This work was supported by the Hunan Provincial Social Science Federation's project: Research on the Development Strategy of Road Running Events Based on Urban Marketing Theory, And the National Social Science Foundation: Research on the Networking of Supply Quality Entities for Marathon Events.

References

Akbar, M.N., Riaz, F., Awan, A.B. et al. (2022) 'A hybrid duo-deep learning and best features based framework for action recognition', *Computers, Materials, and Continua*, Vol. 2022, No. 11, pp.2555–2576.

Alsawadi, M.S., El-kenawy, E-S.M. and Rio, M. (2023) 'Using BlazePose on spatial temporal graph convolutional networks for action recognition', *Computers, Materials, and Continua*, Vol. 2023, No. 1, pp.19–36.

Barkoky, A. and Charkari, N.M. (2022) 'Complex network-based features extraction in RGB-D human action recognition', *Journal Of Visual Communication and Image Representation,* Vol. 1, No. 82, pp.1–9.

Chen, S., Jiang, Q. and Yan, X. (2020) 'Multimodal process monitoring based on transition-constrained Gaussian mixture model', *Chinese Journal of Chemical Engineering: English*, Vol. 28, No. 12, pp.3070–3078.

Ge, X., Xue, Z., Zhou, J. et al. (2022) 'An unsupervised clustering method for nuclear magnetic resonance transverse relaxation spectrums based on the Gaussian mixture model and its application', *Petroleum Exploration and Development*, Vol. 49, No. 2, pp.339–348.

Han, K., Zhang, Y. and Chen, X. (2021) 'Research on the recognition and evaluation of gymnasts' actions from the perspective of deep learning', *Modern Science Instruments*, Vol. 2021, No. 2, pp.234–238.

Han, Y. and Fan, B. (2019) 'Discussion on the method of identifying erroneous movements in sports training based on 3D modeling detection', *Modern Scientific Instruments*, Vol. 2019, No. 5, pp.16159–161165.

Li, B., Chen, J., Zhang, D. et al. (2022) 'Representation learning for compressed video action recognition via attentive cross-modal interaction with motion enhancement', *Computer Vision and Pattern Recognition*, Vol. 5, No. 7, pp.485–497.

Liang, L. (2021) 'An automatic detection system for training movements of aerobics athletes based on posture estimation', *Automation and Instrumentation*, Vol. 2021, No. 3, pp.119–122.

Lin, H. (2021) 'Research on intelligent recognition of motion actions based on laser sensors', *Journal of Laser,* Vol. 42, No. 7, pp.84–89.

Meng, D., Yang, S., Lin, T. et al. (2022) 'RBMDO using Gaussian mixture model-based second-order mean-value saddlepoint approximation', *Computer Modeling in Engineering and Sciences*, Vol. 2022, No. 8, pp.553–568.

Sun, L., Cao, Y., Wu, W. et al. (2020) 'A multi-target tracking algorithm based on Gaussian mixture model', *Systems Engineering and Electronic Technology: English Version*, Vol. 31, No. 3, pp.482–487.

Zhang, J., Ye, G., Tu, Z. et al. (2022) 'A spatial attentive and temporal dilated (SATD) GCN for skeleton-based action recognition', *CAAI Transactions on Intelligence Technology*, Vol. 7, No. 1, pp.46–55.

Zhang, Y., Ma, G., Tian, W. et al. (2020) 'Threshold-based adaptive gaussian mixture model integration (TA-GMMI) algorithm for mapping snow cover in mountainous terrain', *Computer Modeling in Engineering and Sciences*, Vol. 2020, No. 9, pp.1149–1165.

Zhao, J., Chong, S., Huang, L. et al. (2022) 'Action recognition based on CSI signal using improved deep residual network model', *Computer Modeling in Engineering and Sciences*, Vol. 2022, No. 3, pp.1827–1851.

Zhou, S., Chen, L. and Sugumaran, V. (2020) 'hidden two-stream collaborative learning network for action recognition', *Computers, Materials, and Continua*, Vol. 2020, No. 6, pp.1545–1561.