Journal of Network Intelligence ©2024 ISSN 2414-8105
Taiwan Ubiquitous Information Volume **9**, Number **2**, May **2024**

Gaussian Mixture Model Based Player Tracking Technique in Basketball Sports Video

Xin-Hui Jia*

Department of Art and Sports,
Huanghe Science and Technology University, Zhengzhou 450061, P. R. China
201508113@hhstu.edu.cn

Cawlton Evans

Faculty of Engineering,
Dalhousie University, Halifax B3H 4R2, Canada
vu2082@163.com

*Corresponding author: Xin-Hui Jia
Received July 13, 2023, revised September 19, 2023, accepted November 11, 2023.

Abstract. *As sporting events become more and more competitive, many coaches are beginning to use information technology to improve the professionalism of their athletes. Monitoring and recognition of moving objects in sports video is currently a more common method. However, the uniqueness of basketball sports video poses a great challenge to target tracking techniques, especially the interference of camera shake and noise. This works suggests an athlete tracking method based on an improved Gaussian mixture model as a solution to the concerns mentioned above. Firstly, the problem of stereo visual perception in basketball sports videos is investigated. Secondly, the background modelling method based on the traditional Gaussian mixture model is analysed, and the background model parameter estimation method is constructed using the Student-t distribution. The parameter estimation is completed by the expectation-maximum algorithm and the parameter space is partitioned. Then, in order to further improve the accuracy of target tracking, a particle filtering optimization algorithm based on genetic algorithm was proposed in order to eliminate the particle degradation. Finally, two video sequences of NBA regular season games were used for target tracking tests. The experimental results show that the proposed improved Gaussian mixture model has better target tracking results compared with the traditional Gaussian mixture model and the generalised Gaussian mixture model. The tracking accuracy of the athletes is higher, which validates the effectiveness and advancedness of the proposed model.*
Keywords: Basketball game; Video target tracking; Gaussian mixture model; Student-T distribution; Particle filtering

1. **Introduction.** In real life, with the continuous development of computer vision technology, various new applications are emerging, such as face recognition, driving assistance, robot control and video target tracking. Motion target detection and tracking systems are intelligent recognition systems utilizing computerized image processing, pattern identification, computer vision, and additional technologies [1,2]. Video target tracking technology was first applied in the military field and is gradually being rapidly promoted in the civilian sector. Target tracking technology can observe the properties and state of the tracked target, so as to obtain the changes of the tracked target at different moments [3,4]. By analysing these changes, the target can be tracked in real time. In general, video target

tracking requires the analysis of video sequences in order to complete the detection of moving targets, including background modelling, target detection and target tracking, in order to obtain various motion parameters of the tracked target, such as acceleration, velocity, and position [5,6].

Generally, basketball coaches rely on manual observation and experience to coach their players' technical movements. However, this approach is inaccurate and inefficient. By collecting a large amount of video information from high-level athletes in their daily training and analysing this information effectively, the effectiveness of the athletes' training can be greatly improved [7,8]. With the increasing level of competition in sporting events, many coaches are using information technology to improve the professionalism of their athletes.

Motion target detection and tracking in sports video is a relatively common approach today. Motion target detection and tracking belongs to the field of machine vision research. Target detection is the analysis of the attributes of a moving target within a video and is a prior condition for target tracking. The technique of determining the location of a moving object of interest in each individual frame of a video sequence is referred to as "target tracking." However, the presence of a large number of moving targets in a sports video (with interference between them) and the large variation in the speed of the targets makes it difficult to implement motion target tracking.

Motion target detection and tracking technology has been an important research area of constant interest. At this stage, there are three main motion target detection methods [9,10]: 1) optical flow method; 2) inter-frame difference method; 3) background subtraction method. Among them, the optical flow method has a higher detection accuracy, but the complexity of the operation is greater and the sensitivity to illumination is higher, so the scope of application is narrower. The inter-frame difference method has better denoising ability, but is more dependent on the video sequence. The detection accuracy will be reduced when the motion speed varies widely [11]. Due to its good applicability, the background subtraction method has received a lot of attention and has become the mainstream method in motion target detection techniques. The background subtraction method is able to segment moving objects well and meets the needs of most video target tracking applications [12,13]. The accuracy of the background subtraction method is not affected by the speed of motion, and its detection performance is directly determined by the performance of background modelling. In real application scenarios, factors such as illumination changes [14], background interference, shadow interference and camera shake can adversely affect moving target detection.

1.1. **Related Work.** Background subtraction based on probability statistics is currently the most popular method and is divided into 2 main types: 1) parametric and 2) non-parametric. Wang et al. [15] proposed an improved parameter estimation method for background modelling based on probability statistics and applied it to different image processing scenarios. Chu et al. [16] proposed a non-parametric method based on k-means. Although the non-parametric method has better applicability, it can only be used for shorter video sequences, so most studies prefer to use parametric methods. Li et al. [17] proposed an embedded Mean Shift based particle filter target tracking. Delail et al. [18] proposed an improved particle filter target tracking algorithm for colour adaptive fusion. Quang et al. [19] proposed a target tracking technique based on particle filtering and Laplace transform. All of these methods use a hybrid optimization strategy to improve the performance of target tracking by combining advanced optimization algorithms with particle filtering algorithms in order to compensate for the shortcomings of the background subtraction method.

In the field of sports video analysis, motion target tracking technology plays an even more integral role. In practical terms, multiple cameras set up high up on the playing field are used to acquire motion video of the game. Firstly, after image processing, we can obtain various reference data. Then, based on the height and angle of the cameras, the exact coordinates of the target in the actual space in the sports video can be calculated. Secondly, by processing and analysing the captured images through computer vision techniques, digital image processing and pattern recognition, the various statistical data required can be obtained. Finally, by tracking the athletes in real time, the trajectory of their movements can be analysed.

Tracking results can be used to correct subtle movement errors in training, thus improving the athletes' training results and game performance, a function that cannot be achieved by traditional manual methods. However, in sports video, the colours of the athletes and the background field may be similar, and the athletes are also obscured from each other. The uniqueness of sports videos poses a great challenge to target tracking techniques. Therefore, the aim of this study is to propose an accurate and efficient algorithm for basketball player tracking technology and to verify its feasibility.

A number of researchers have conducted long-term research on the detection and tracking challenges of sports video. Wu et al. [20] proposed a target detection and tracking method based on optical flow recognition. Although this method effectively deals with the problem of light occlusion in the playing field area, its large amount of operations (long detection time) leads to poor early tracking results. Hou and Li [21] use the adjacent frame difference technique for motion target detection, but the motion target points are easily lost during the tracking process, so it is only applicable to sports events with simple scenes. All of the above methods have major drawbacks and cannot accurately extract the attributes of the motion targets. This is because the attributes of motion targets in sports videos are affected by interference such as light and noise, so a more stable mathematical model is needed to achieve the attribute extraction of motion targets in order to enhance the effectiveness of target tracking. The Gaussian mixture model uses gradient descent to optimise the parameters, which reduces the complexity of the algorithm and effectively removes the influence of noise on image segmentation [22,23]. In addition, the Gaussian mixture model has a good adaptive capability.

Due to the above advantages of Gaussian mixture model, this paper tries to apply it to motion target tracking in basketball videos. However, in practical applications, the probability distribution of pixels in video sequences does not always conform to the Gaussian distribution due to factors such as lighting variations and camera shake [24,25]. Student-t distribution has better robustness in noisy test datasets due to its special statistical properties [26,27], and thus may significantly aid in improving the precision of target tracking.

Therefore, in this paper, an improved Gaussian mixture model based on parameter estimation of Student-t distribution is proposed. First, the traditional background modelling method based on a finite mixture model is analysed. Then the background model parameter estimation method is constructed using the Student-t distribution. The parameter estimation is completed by the expectation-maximum algorithm and the parameter space is partitioned.

1.2. **Motivation and contribution.** In addition, no a priori information exists for athlete tracking in basketball video sequences, so a filter must be built and reasonably initialised to enable state detection of a moving target (athlete) when it enters the scene (basketball court). In addition, the filter needs to be continuously updated to allow real-time tracking of the moving target. Therefore, a particle filter optimisation algorithm based on a genetic algorithm is proposed in this paper. The simulation test results show that the proposed improved Gaussian mixture model has better target tracking compared to the traditional Gaussian mixture model and the generalised Gaussian mixture model. The tracking accuracy of the athletes is higher, which verifies the effectiveness and advancedness of the proposed model.

The main innovations and contributions of this paper include.

(1) To effectively remove the effects of jitter and noise on image segmentation, the Gaussian mixture model was improved using Student-t distribution parameter estimation in order to match the probability distribution of pixels in the video sequence. The parameter estimation was accomplished by an expectation-maximum algorithm, and the target segmentation of the parameter space was performed.

(2) In order to further improve the accuracy of target tracking, a particle filter optimisation algorithm based on a genetic algorithm is proposed. The particle filter and degradation weights are used to detect whether a moving target appears in the observation model, and a genetic algorithm is introduced to improve the particle filtering algorithm in order to eliminate particle degradation.

2. Stereo visual perception in basketball sports videos.

2.1. Camera calibration.

Stereo visual perception in sports video is essentially a problem of spatial localisation of a sports target. Determining the spatial position of the target is for the transformation from image plane coordinates to spatial three-dimensional coordinates, which involves camera calibration and visual model.

Camera imaging is based on the same principle as small-aperture imaging, i.e. perspective projection. There are three coordinate systems commonly used in camera calibration: the image coordinate system [28,29], the camera coordinate system and the world coordinate system. The task of coordinate system conversion is accomplished by establishing the relationship between the image coordinate system and the world coordinate system. In the two-dimensional coordinate system of an image, the coordinates of each pixel *(u,v)* represent the number of columns and rows of the pixel in the array, as shown in Figure 1. *(u,v)* is the coordinate system of the image in pixels.

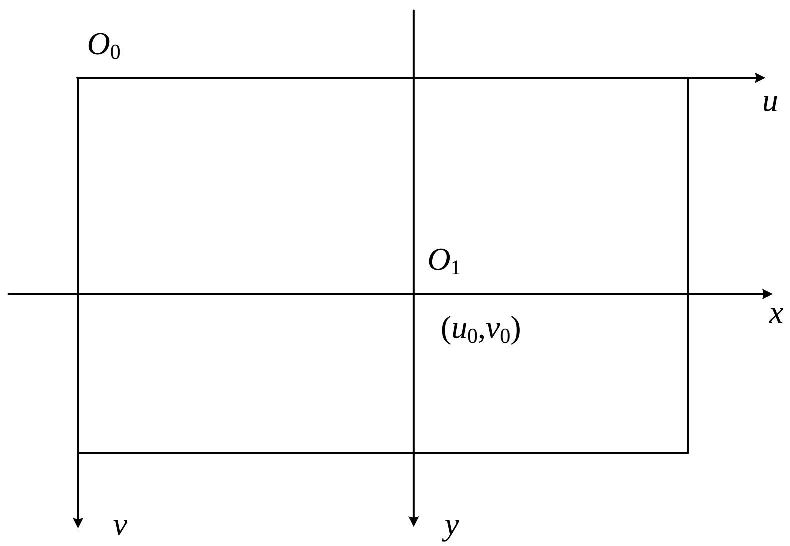


Figure 1. Image coordinate system for camera calibration

The world coordinate system is a coordinate system *(X, Y, Z)* defined in the 3D world. The transformation relationship from the image coordinate system to the world coordinate system is shown as follow:

$$s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_x & 0 & u_0 & 0 \\ 0 & \alpha_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \mathbf{M_1 M_2 X_w} = \mathbf{M X_w} \tag{1}$$

$$\mathbf{M} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \tag{2}$$

where α_x is the scale factor on the u-axis and α_y is the scale factor on the v-axis. $\mathbf{M}$ is the projection matrix, $\mathbf{M_1}$ is the internal camera parameter matrix and $\mathbf{M_2}$ is the external camera parameter. The camera calibration is to determine the transformation relationship between the image coordinate system and the world coordinate system, so as to solve for the internal and external parameters.

2.2. **Binocular stereo vision.** In biological vision systems, the vast majority of organisms have two eyes. Therefore, in a computerised stereo vision system, two images of the same scene from different angles need to be acquired using a camera [30,31]. Then, using the principle of three-dimensional reconstruction, the three-dimensional shape of the scene is reconstructed by the computer, so that the spatial position information of the object can be recovered. The principle of binocular stereo vision is shown in Figure 2.

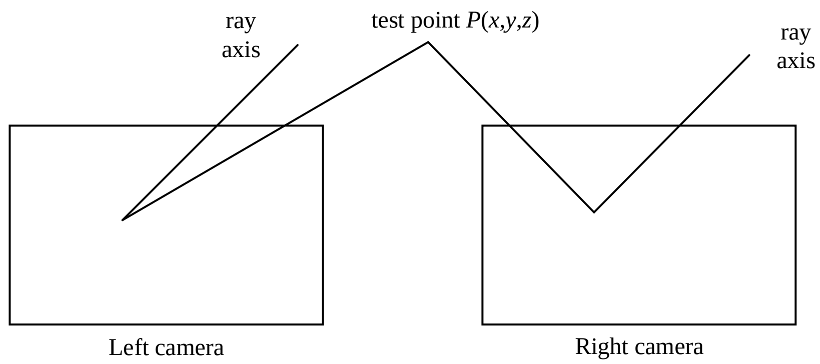


Figure 2. Principle of binocular stereo vision

2.3. **Structured light stereo vision.** When an optical projector is used instead of one of the cameras in binocular stereo vision, the optical projector projects certain light patterns, such as light planes and grid lines. Constraints need to be placed on the spatial position of the scene object to obtain a unique coordinate value for the point on the scene object [32], which results in structured light 3D vision. The principle of structured light stereo vision is shown in Figure 3.

3. **Basketball player tracking based on gaussian mixture model.**

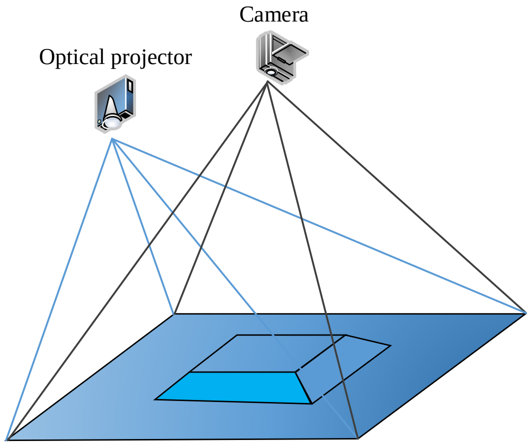


Figure 3. Structured light stereo vision principle

3.1. **Parameter estimation for the finite mixture model.** A single probability distribution has only a single peak [33]. However, the probability distributions of most practical application scenarios consist of multiple peaks. Therefore, probability distributions with multiple peaks should be used to effectively simulate real-world application scenarios. The finite mixture model is a linear superposition of the probability distribution functions of each cluster obtained statistically. The finite mixture model [34,35] is defined as shown as follow:

$$f(x) = \sum_{k=1}^{K} \pi_k P(x|\theta_k) \tag{3}$$

where $P(x \mid \theta_k)$, $k = 1, 2, \ldots, K$ is the probability distribution function of the k-th cluster, θ_k is the corresponding vector, π_k is the weighting coefficient of the k-th cluster. k is the total number of components of the finite mixed model.

3.2. **Background modelling based on Gaussian mixture models.** For Equation (3), if $P(x \mid \theta_k)$ is a Gaussian distribution, then $f(x)$ is the Gaussian mixture model. Let x_i denote the i-th pixel of the image, $i = 1, 2, \ldots, N$. The probability density function of the current pixel has a linear combination of class K, and Ω_j is the j-th class, $j = 1, 2, \ldots, K$. The probability density function of x_i is shown as follows:

$$p(x_i) = \sum_{j=1}^{K} \pi_{ij} p(x_i|\Omega_j) \tag{4}$$

where π_j represents the scale factor for each Gaussian distribution [36,37], for $0 \leq \pi_{ij} \leq 1$ and $\sum_{k=1}^{K} \pi_{ij} = 1$. In the Gaussian mixture model, each Gaussian distribution $p(x_i \mid \Omega_j)$ is expressed as follows:

$$p(x_i|\Omega_j) = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(x_i - \mu_j)^2}{2\sigma_j^2}\right) \tag{5}$$

where μ_j is the mean of the Gaussian distribution and σ_j is the covariance of the Gaussian distribution. According to Equation (3), we can obtain the logarithmic likelihood function.

$$L(\theta) = \sum_{i=1}^{N} \ln\left(\sum_{j=1}^{K} \pi_{ij} p(x_i|\Omega_j)\right) \tag{6}$$

The posterior probability of the i-th pixel can be expressed as follow:

$$p(\Omega_j|x_i) = \frac{\pi_{ij}p(x_i|\Omega_j)}{\sum\limits_{k=1}^{K} \pi_{ik}p(x_i|\Omega_k)} \tag{7}$$

Generally, empirical solution is used to realize log-likelihood function.

$$E(\theta) = -L(\theta) = -\sum_{i=1}^{N} \ln\left(\sum_{j=1}^{K} \pi_{ij}p(x_i|\Omega_j)\right) \tag{8}$$

In order to obtain the largest log likelihood function, the Equation (8) must be minimized.

$$\varepsilon(\theta^{old}|\theta^{new}) = -\sum_{i=1}^{N}\sum_{j=1}^{K} p^{old}(\Omega_j|x_i) \cdot \ln\left(\pi_{ij}p^{new}(x_i|\Omega_j)\right) \tag{9}$$

where ε represents the error function. In order to get the maximum log-likelihood function, ε must be minimized, and the optimized parameters should be brought into Equation (6).

3.3. **Improvement of the Gaussian mixture model.** We need to approximate the pixel distribution as a normal distribution for Gaussian mixture background modelling. However, for basketball sports video sequences, the camera shake and the presence of noise during filming lead to a distribution of pixel grey values with more narrow band characteristics. Taking a 2-minute video sequence of a basketball game as an example, we selected a point with a relatively homogeneous distribution of grey values (position point (5,115) in the pixel coordinate system) and analysed its grey scale histogram, as shown in Figure 4. The sample number of basketball video sequences is 285. The Gaussian fit to the histogram and the t-distribution fit are shown in Figure 5.

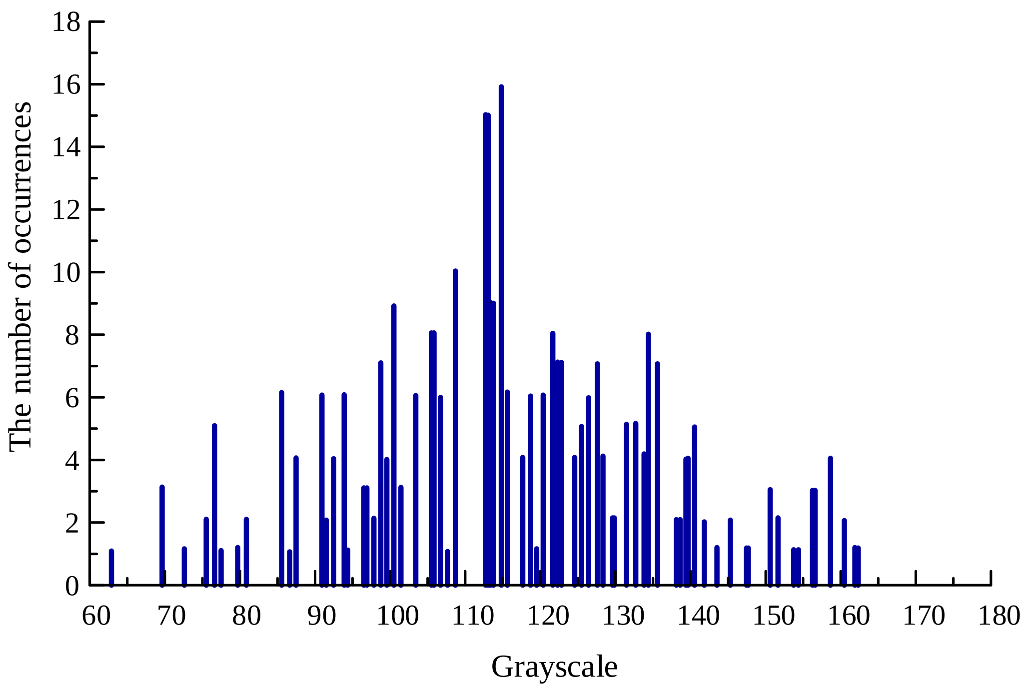


Figure 4. Histogram distribution of the grey scale of the pixel points (5,115)

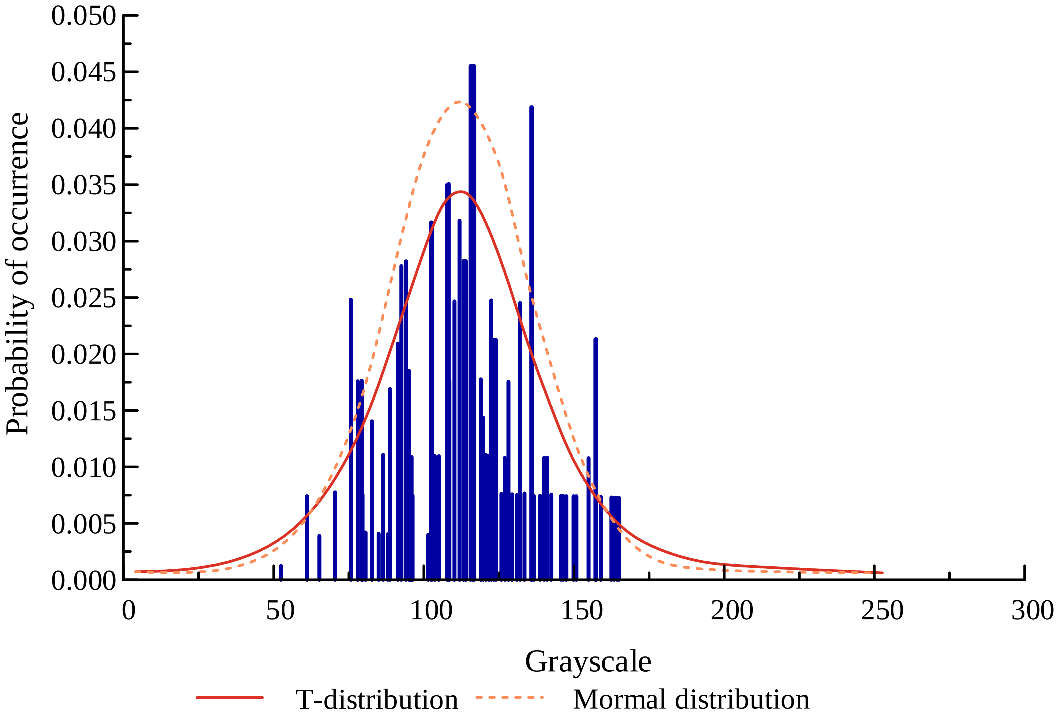


Figure 5. Gaussian fit of the histogram and t-distribution fit

It can be seen that the probability density model based on the t distribution can better fit the actual distribution of the sequence images than the Gaussian distribution (normal distribution) [38,39]. Therefore, the Gaussian mixture model was improved using Student-t distribution parameter estimation, thus effectively removing the effects of jitter and noise on image segmentation.

We set $X = \{X_1, \ldots, X_N\}$ to represent the independent distribution data with dimension d, which is divided into K types. The Student-t distribution can be expressed as $t(\mu_k, H_k, \nu_k)$, where ν is the degree of freedom of the Student-t distribution, μ is the mean of the Student-t distribution, and H_k is the symmetric and positive definite matrix [40]. The smaller the value of ν, the longer the tail length of the Student-t distribution as shown in Figure 6. When the degree of freedom ν is infinite, the variance of Student-t distribution and variance of Student-t distribution and Gauss distribution is the same.

The Student-t distribution model of each pixel in the image sequence is shown as follow:

$$f(x) = \sum_{k=1}^{K} \pi_k S(X_i, \mu_k, H_k, v_k) \tag{10}$$

where $S(X_i, \mu_k, H_k, v_k)$ denotes the probability density function of component t in a cluster

$$S(X_i, \mu_k, H_k, v_k) = \frac{\Gamma\left(\frac{v_k+d}{2}\right) |H_k|^{-\frac{1}{2}}}{(\pi v_k)^{\frac{d}{2}} \Gamma(\frac{v_k}{2}) \{1 + \Delta / v_k\}^{\frac{v_k+p}{2}}} \tag{11}$$

where $\Gamma(\cdot)$ denotes a gamma function.

$$\Delta = (X_t - \mu_k) H_k^{-2} (X_t - \mu_k)^T \tag{12}$$

In order to improve the convergence rate, this paper uses the maximum expectation algorithm to complete the model parameter estimation. We set Θ represents a set of parameters for a finite mixed model and $P(\Theta)$ represents the prior probabilities of the

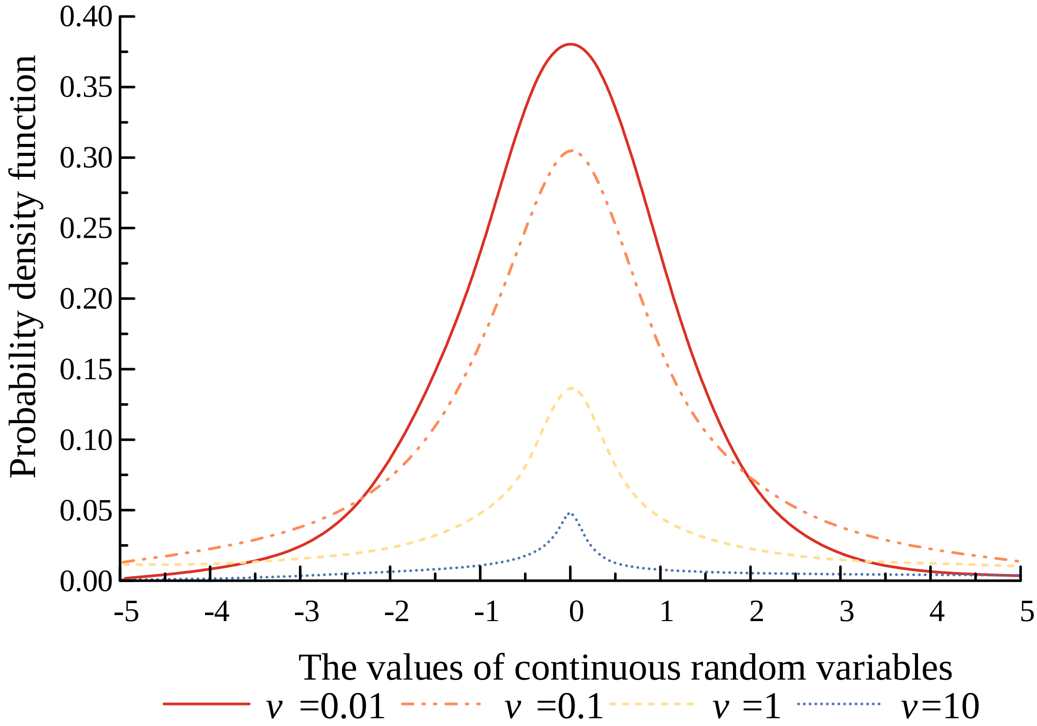


Figure 6. The influence of different v on Student-t distribution

parameters. Divide Θ into two parts $\{\Theta_1, \Theta_2\}$, where $\Theta_1 = \{\pi_k, \mu_k, H_k\}$ and $\Theta_2 = \{\nu_k\}$. Firstly, the μ_k is updated by using Θ_1, and then the degree of freedom is estimated after repeated m times.

The concrete steps are shown as follows:

Step 1: Set the initial condition is $\{\pi_k^0, \mu_k^0, H_k^0, \nu_k^0\}_{k=1}^K$, $t = 1$.

Step 2: In the t-th iteration step, the following is used to calculate $\mu_{ik}^{(t)}$ and $\pi_{ik}^{(t)}$.

$$\mu_{ik}^{(t)} = \frac{v_k^{(t)} + d}{v_k^{(t)} + \Delta(X_i, \mu_{ik}^{(t)} H_k^{(t)})} \tag{13}$$

$$\tau_{ik}^{(t)} = \frac{\pi_k^{(t)} S_k(X_i|\theta_k^{(t)})}{\sum_{j=1}^{K} \pi_j^{(t)} S_j(X_j|\theta_j^{(t)})} \tag{14}$$

where $\Theta_k^{(t)} = \{\mu_k^{(t)}, H_k^{(t)}, \nu_k^{(t)}\}$ represents the set of parameters for component t.

Step 3: Set $\Theta_2 = \Theta_2^{(t)}$ fixed, then we use the following to update $\mu_{ik}^{(t+1)}$, $\pi_{ik}^{(t+1)}$ and $H_{ik}^{(t+1)}$.

$$\pi_k^{(t+1)} = \frac{1}{N} \sum_{i=1}^{N} \tau_{ik}^{(t)} \tag{15}$$

$$\mu_k^{(t+1)} = \frac{\sum_{i=1}^{N} X_i \tau_{ik}^{(t)} \mu_{ik}^{(t)}}{\sum_{i=1}^{N} \tau_{ik}^{(t)} \mu_{ik}^{(t)}} \tag{16}$$

$$H_{ik}^{(t+1)} = \frac{\sum_{i=1}^{N} (X_i - \mu_{ik}^{(t+1)})(X_i - \mu_{ik}^{(t+1)})^T \tau_{ik}^{(t)} \mu_{ik}^{(t)}}{\sum_{i=1}^{N} \tau_{ik}^{(t)}} \tag{17}$$

Step 4: Repeat Steps (3)–(5) for a total of m times.

Step 5: Use $(\Theta_1^{(t+1)}, \Theta_2^{(t)})$ to replace $\Theta_1^{(t)}, \Theta_2^{(t)}$ and use Equation (12) to calculate $\mu_{ik}^{(t)}$.

Step 6: Set $\Theta_1 = \Theta_1^{(t)}$ fixed and update $\nu_{ik}^{(t+1)}$.

$$1+\log(0.5v_k)+\frac{1}{\sum_{i=1}^{N} \tau_{ik}^{(t)}} \left(\sum_{i=1}^{N} \tau_{ik}^{(t)} (\log \mu_{ik}^{(t)} - \mu_{ik}^{(t)}) \right) -\Phi(0.5v_k)+\Phi(0.5(v_k+p))-\log(0.5(v_k+p)) = 0 \tag{18}$$

$$\Phi(x) = \frac{d \ln \Gamma(x)}{dx} \tag{19}$$

Step 7: Calculate $\left|\Theta_1^{(t)} - \Theta_1^{(t+1)}\right|$, if it is less than ε, the iteration will be finished, otherwise proceed to Steps 2-6. We set $\varepsilon = 10e^{-5}$.

3.4. **Motion target tracking.** How to achieve accurate tracking of moving targets in complex backgrounds has been a hot issue for researchers. Particle filtering algorithms derived from Monte Carlo ideas can be effectively applied to target tracking. Therefore, this paper proposes a particle filtering optimization algorithm based on genetic algorithm. As a bionic evolutionary algorithm, genetic algorithm has the characteristics of high global search ability and fast convergence. Therefore, this paper adopts genetic algorithm to improve the particle filtering algorithm, so as to eliminate the phenomenon of particle degradation. In addition, the HSV distribution model [41] is used to construct a target observation model, and then the particle filter and degradation weights are used to detect the presence or absence of moving targets.

Let the tracking target be x and the target state be x_k, x_k^n denote the n-th sample of the k-th frame in the motion video. Firstly, the HSV distribution model of the tracking target is constructed by the barotropic distance (similarity comparison function).

$$D[p, q] = \sqrt{1 - \rho[p^u, q^u]} \tag{20}$$

where $\rho[p^u, q^u]$ denotes the baroclinic coefficient. The expression for the observation model is shown as follows:

$$P(Y_k \mid X_k) = \frac{1}{\sqrt{2\pi}} \exp\left(-\gamma D^2(p, q)/2\right) \tag{21}$$

where γ represents the control factor.

No a priori information exists for tracking targets in the basketball video studied in this paper, so a filter must be built and reasonably initialised. In addition, the filter needs to be continuously updated to allow real-time tracking of moving targets [42,43]. In order to obtain the estimated centre of the desired tracking target, white pixels need to be processed using a particle filter.

Let s denote the state vector of the particle filter.

$$s = \{x, y, v_x, v_y\} \tag{22}$$

where x and y indicate the horizontal and vertical coordinates of the centre of the sample image, respectively. v_x and v_y denote the velocities of the coordinates, respectively.

The estimated centre of the motion target is obtained by means of a weight calculation.

$$E = \sum_{n=1}^{N} \omega_n * s_t^n \tag{23}$$

where s_t^n is the state of the particle n at time t, ω_n is the weight of the particle n, and N is the total number of particles.

$$s_t = \mathbf{D}s_{t-1} + w_{t-1} \tag{24}$$

$$\mathbf{D} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{25}$$

where s_t represents the state of the particle at time t, $\mathbf{D}$ represents the key elements of the model, and w_{t-1} represents Gaussian random white noise.

When the tracking target is first detected, the weights of each particle can be derived using the likelihood function.

$$\omega_n = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(cp-tp)^2}{2\sigma^2}}, tp \in [150, 255] \tag{26}$$

where σ is the standard variance, cp is the true colour and tp represents the target colour value. We need to perform a normalisation operation on the weights.

$$f_{\mathrm{NOR}} = \frac{1}{\sum_{n=1}^{N} \omega_n} \tag{27}$$

In general, if the magnitude of a particle's weights is close to 1, this indicates that the weights are degenerating. This paper determines whether a moving target is present in the scene by degenerating the weights Φ.

$$\Phi = \frac{1}{\sum_{n=1}^{N} \omega_n^2} \tag{28}$$

where ω_n is the weight of the particle n.

3.5. **Genetic algorithm-based particle filtering optimization.** In order to solve the particle degradation problem and improve the global search capability, this paper optimizes the particle filtering algorithm using the selection, crossover and mutation operations of the genetic algorithm, with the following implementation steps.

Step 1: Generate the particle swarm $\{x_0^i\}.i = 0, 1, ..., N$ based on the initial conditions and set the weights of all particles to $1/N$;

Step 2: Perform genetic operations, including selection and crossover, to obtain a new set of particles $\{\tilde{x}_k^i\}.i = 0, 1, ..., N$;

Step 3: Update the weight values of the particles;

Step 4: Re-sample using the roulette method and choose to save the sample with the larger weight;

Step 5: Prediction of the current sample using a state transfer model.

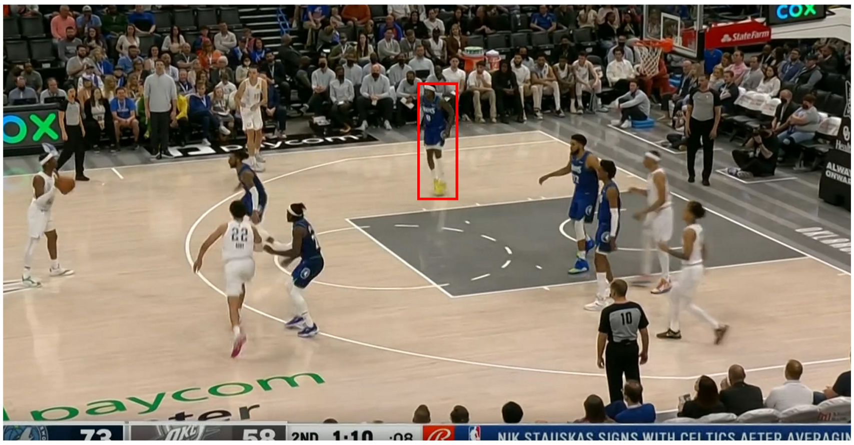
$$\begin{bmatrix} x_k \\ y_k \end{bmatrix} = \begin{bmatrix} x_{k-1} \\ y_{k-1} \end{bmatrix} + \begin{bmatrix} z_{k-1}^x \\ z_{k-1}^y \end{bmatrix} \tag{29}$$

Step 8: If the maximum number of iterations or the stop condition (error value) is reached, stop iterating in sequence, otherwise skip to Step 2.

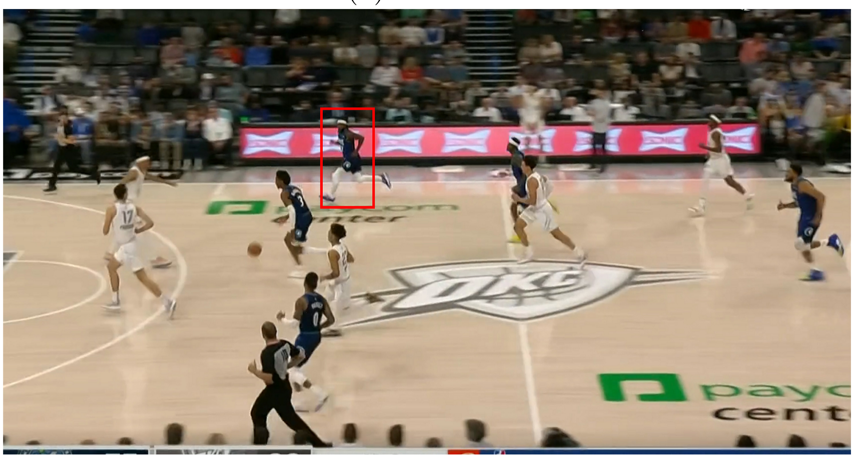
4. Experimental results and analysis.

4.1. Experimental environment and parameter settings. In order to validate the performance of the proposed improved Gaussian mixture model, specific tests were performed on two 800-frame video sequence of a basketball game and compared with the traditional Gaussian mixture model and the generalized Gaussian mixture model [44]. The image size of the basketball game video sequence was 320 240. The experimental environment used a Windows 10 operating system, an Intel(R) Core(TM) i7 CPU (@3.60 GHz), 8 GB of RAM, and a GeForce GTX 970 graphics card. The motion video target tracking simulation used the open source vision computing library OpenCV and Matlab 7.1. the Student-t distribution parameter estimation algorithm was implemented using the MATLAB programming language. The crossover probability was set to 0.6, the number of particles to 100 and the maximum number of iterations to 150.

4.2. Motion target tracking results. Both basketball game video sequences are from the NBA regular season, video sequence 1 is the Timberwolves vs Thunder game from March 5, 2022. Video sequence 2 is the Spurs vs. Hornets game from March 6, 2022. The frame sizes of the videos are 320 240 (pixels). An example of target tracking for video sequence 1 is shown in Figure 7.



(a) 50 frames



(b) 150 frame

Figure 7. Example of target tracking for video sequence 1

4.3. **Performance evaluation indicators.** To quantitatively evaluate the performance of motion target tracking, the background detection error (B_error), fall-out detection error (F_error) and whole recognition rate (WRR) are used.

$$\text{B\_error} = \frac{BN}{B} \tag{30}$$

$$\text{F\_error} = \frac{FN}{F} \tag{31}$$

$$\text{WRR} = \frac{BP + FP}{B + F} \tag{32}$$

where B denotes the number of pixels in the background, and F denotes the number of pixels in the foreground area. BN and FN denote incorrect detection, while BP and FP denote correct detection. Smaller values for B_error and F_error are better, while larger values for WRR are better.

4.4. **Tracking performance analysis.** The results of the two video sequences were averaged to obtain comparative results of the detection rates of the three different tracking models, as shown in Table 1 and Figure 8. It can be seen that the proposed model in this paper has the lowest B_error and F_error (52.21% and 0.98%), while having the highest WRR (63.71%).

Table 1. Comparison results of detection rate.

	B_error (%)	**F_error (%)**	**WRR (%)**
Traditional Gaussian mixture model	54.32	6.47	57.61
Generalised Gaussian mixture model	53.14	3.21	62.05
Improved Gaussian mixture mode	52.21	0.98	63.71

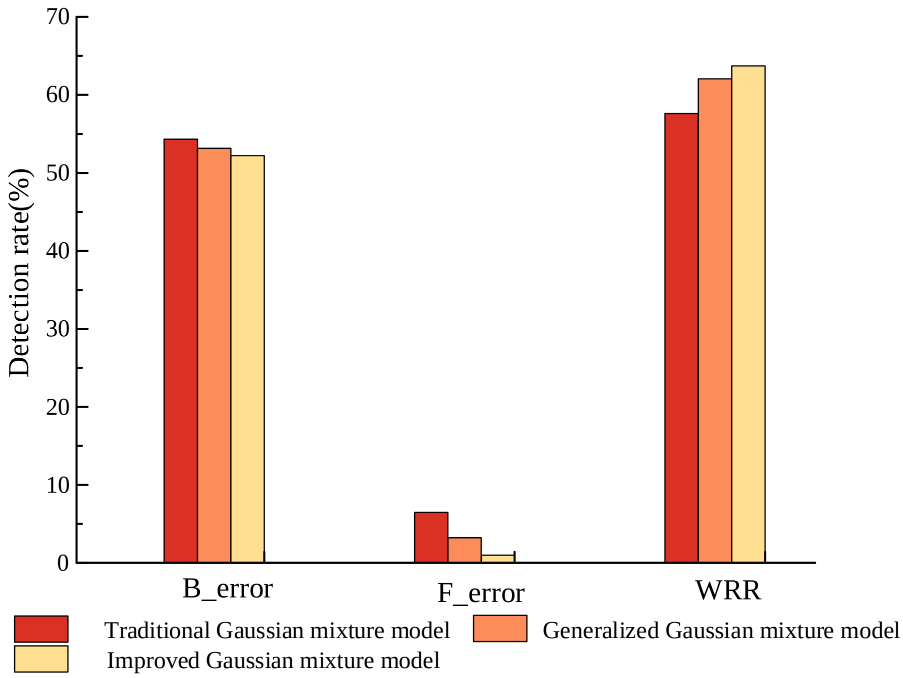


Figure 8. Target detection rates for three different models

To further quantify the performance of the target tracking algorithm, the average tracking error was used as an evaluation metric. For the tracker proposed in this paper, the maximum number of particles is 100. For two video sequences, three tracking algorithms (conventional Gaussian mixture model, generalised Gaussian mixture model and improved Gaussian mixture model) were tested and the results are shown in Figure 9 and Figure 10.

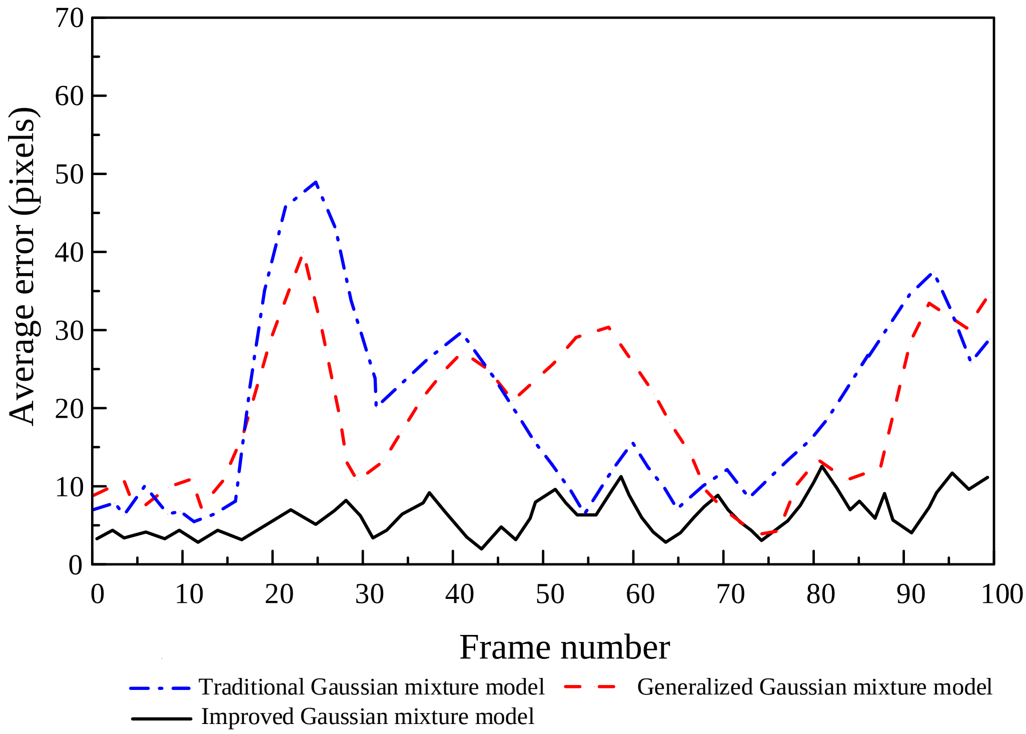


Figure 9. Tracking error analysis for video sequence 1

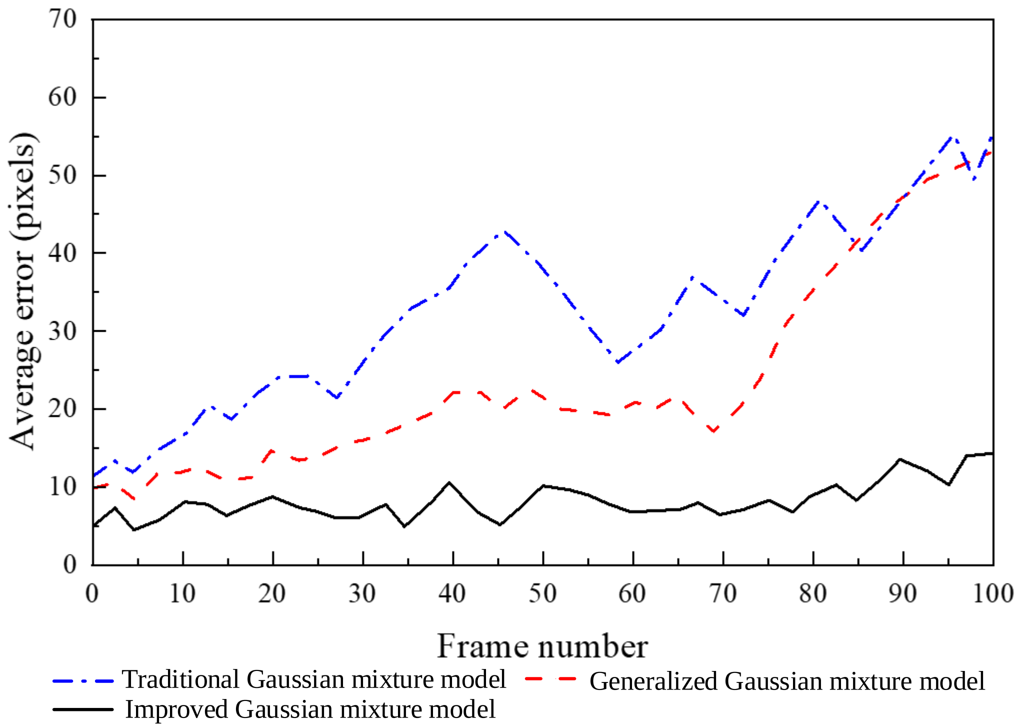


Figure 10. Tracking error analysis for video sequence 2

It can be seen that the improved Gaussian mixture model significantly outperforms the other two models in both video sequences. The average error accuracy of the improved Gaussian mixture model has been maintained at a relatively low level. In NBA basketball game sequence 1, the average error of the improved Gaussian mixture model was 38% and 30% lower compared to the traditional Gaussian mixture model and the generalised Gaussian mixture model respectively. In NBA basketball game sequence 2, the average error of the improved Gaussian mixture model was reduced by 57% and 62% compared to the traditional Gaussian mixture model and the generalised Gaussian mixture model respectively. The experimental results show that. The background model using the t-distribution method can obtain a fit that is closer to the distribution characteristics of the actual data and effectively removes the effects of jitter and noise on target tracking.

4.5. **Complexity and Robustness Analysis.** In order to analyse the operational efficiency of the algorithms, the complexity of the proposed algorithms was evaluated. The training time required for the three different tracking algorithms was counted. In the tests the training time was calculated for each pixel point in the image. Finally, a simple summation operation was performed for the training time of all pixel points. The results of the training time comparison for each pixel point are shown in Table 2.

Table 2. Comparison of processing speed.

	Video sequence 1(s)	**Video sequence 2(s)**
Traditional Gaussian mixture model	0.018	0.021
Generalised Gaussian mixture model	0.047	0.051
Improved Gaussian mixture model	0.086	0.091

It can be seen that the improved Gaussian mixture model requires iterative computation of degrees of freedom, making the steps to build the model increased. As a result, the training speed of the improved Gaussian mixture model is somewhat reduced compared to the other 2 models. Subsequent research will focus on addressing this issue.

The robustness of Gaussian mixture models based on conventional particle filtering depends mainly on the number of particles. The higher the number of particles, the less robust it is. However, in the improved Gaussian mixture model, the number of particles has no significant impact on the tracking performance of the moving target. In other words, the improved Gaussian mixture model can achieve better tracking accuracy and higher robustness with a smaller number of particles. A comparison of the root mean square error of the three motion target tracking models was carried out in order to evaluate their robustness, and the results are shown in Figure 11.

The time required to process each image frame and the optimum number of iterations for the improved Gaussian mixture model are shown in Table 3. It can be seen that the improved Gaussian mixture model can successfully track one or more moving targets. Thus, the genetic algorithm-based particle filtering optimisation effectively solves the problem of particle degradation.

Combining the above experimental results, we can conclude that, compared with the other two Gaussian mixture models, the improved Gaussian mixture model shows excellent performance in terms of detection performance, robustness and tracking error, effectively overcoming the interference of camera jitter and noise in sports videos.

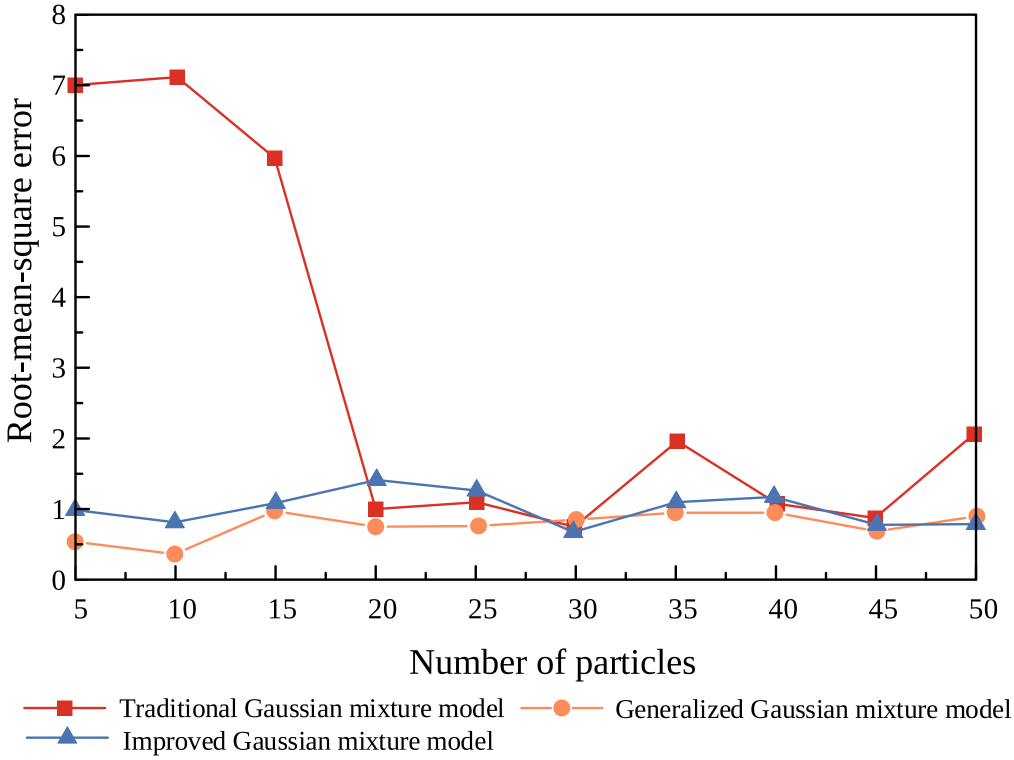


Figure 11. Number of particles versus root mean square error

Table 3. Running time and number of iterations.

Number of targets in the video sequence	**Time to process a single frame (s)**	**Optimal number of iterations**
Individual	0.035	5
Three	0.040	5
Multiple (under 10)	0.052	8

5. **Conclusion.** In this paper, an improved Gaussian mixture model based on parameter estimation of the Student-t distribution is proposed in order to better fit the probability distribution of pixels in sports video sequences. First, the traditional background modelling method based on a finite mixture model is analysed. The background model parameter estimation method was then constructed using the Student-t distribution. The parameter estimation was completed by the expectation-maximum algorithm, and the parameter space was partitioned. Secondly, a particle filter optimisation algorithm based on genetic algorithm is proposed in this paper to detect whether the moving target appears in the observed model by means of particle filters and degradation weights, and a genetic algorithm is introduced to improve the particle filtering algorithm in order to eliminate the particle degradation. The experimental results show that without any a priori information, the improved Gaussian mixture model can track the human target in the basketball game video better, verifying its feasibility and advancement. Compared with the traditional Gaussian mixture model and the generalised Gaussian mixture model, the improved Gaussian mixture model has better target tracking and better tracking accuracy of the athletes. However, the improved Gaussian mixture model requires iterative computation of degrees of freedom, which makes the training speed somewhat reduced. Further research on this issue will be carried out later.

REFERENCES

[1] Y. Wang, T. Wang, G. Zhang, Q. Cheng, and J.-Q. Wu, "Small Target Tracking in Satellite Videos Using Background Compensation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 10, pp. 7010-7021, 2020.

[2] S. S. Moghaddasi, and N. Faraji, "A hybrid algorithm based on particle filter and genetic algorithm for target tracking," *Expert Systems with Applications*, vol. 147, 113188, 2020.

[3] J. Gemerek, S. Ferrari, B. H. Wang, and M. E. Campbell, "Video-guided Camera Control for Target Tracking and Following," *IFAC-PapersOnLine*, vol. 51, no. 34, pp. 176-183, 2019.

[4] S. Kamkar, F. Ghezloo, H. A. Moghaddam, A. Borji, and R. Lashgari, "Multiple-target tracking in human and machine vision," *PLoS Computational Biology*, vol. 16, no. 4, 1007698, 2020.

[5] W. Tiancheng, "Unsupervised video multi-target tracking based on fast resampling particle filter," *The Journal of Supercomputing*, vol. 76, no. 2, pp. 1293-1304, 2018.

[6] B. Zhao, Y. Han, H. Wang, L. Tang, X. Liu, and T. Wang, "Robust Shadow Tracking for Video SAR," *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 5, pp. 821-825, 2021.

[7] E. C. d. Castro, E. O. T. Salles, and P. M. Ciarelli, "A New Approach to Enhanced Swarm Intelligence Applied to Video Target Tracking," *Sensors (Basel, Switzerland)*, vol. 21, no. 5, 1903, 2021.

[8] J. Bao, X. Zhang, T. Zhang, J. Shi, and S. Wei, "A Novel Guided Anchor Siamese Network for Arbitrary Target-of-Interest Tracking in Video-SAR," *Remote Sensing*, vol. 13, no. 22, 4504, 2021.

[9] L. Kang, R.-S. Chen, N. Xiong, Y.-C. Chen, Y.-X. Hu, and C.-M. Chen, "Selecting Hyper-Parameters of Gaussian Process Regression Based on Non-Inertial Particle Swarm Optimization in Internet of Things," *IEEE Access*, vol. 7, pp. 59504-59513, 2019.

[10] A. L. H. P. Shaik, M. K. Manoharan, A. K. Pani, R. R. Avala, and C.-M. Chen, "Gaussian Mutation–Spider Monkey Optimization (GM-SMO) Model for Remote Sensing Scene Classification," *Remote Sensing*, vol. 14, no. 24, 6279, 2022.

[11] B. Yang, and R. Nevatia, "Multi-Target Tracking by Online Learning a CRF Model of Appearance and Motion Patterns," *International Journal of Computer Vision*, vol. 107, no. 2, pp. 203-217, 2013.

[12] A. Vo, X. Feng, W. C. Smith, D. Zhu, M. Patel, D. Kozak, Y. Wang, J. Zheng, M. Ashraf, and X. Xu, "Analyzing ophthalmic suspension particle size distributions using laser diffraction: Placebo background subtraction method," *International Journal of Pharmaceutics*, vol. 598, 120401, 2021.

[13] W. Zeng, C. Xie, Z. Yang, and X. Lu, "A universal sample-based background subtraction method for traffic surveillance videos," *Multimedia Tools and Applications*, vol. 79, no. 31-32, pp. 22211-22234, 2020.

[14] A. Djerida, Z. Zhao, and J. Zhao, "Background subtraction in dynamic scenes using the dynamic principal component analysis," *IET Image Processing*, vol. 14, no. 2, pp. 245-255, 2020.

[15] G. Wang, L. Schultz, and J. Qi, "Statistical Image Reconstruction for Muon Tomography Using a Gaussian Scale Mixture Model," *IEEE Transactions on Nuclear Science*, vol. 56, no. 4, pp. 2480-2486, 2009.

[16] C.-C. Chiu, M.-Y. Ku, and L.-W. Liang, "A Robust Object Segmentation System Using a Probability-Based Background Extraction Algorithm," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 4, pp. 518-528, 2010.

[17] L. Jia-qiang, Z. Rong-hua, C. Jin-li, Z. Chun-yan, and Z. Yan-ping, "Target tracking algorithm based on adaptive strong tracking particle filter," *IET Science, Measurement & Technology*, vol. 10, no. 7, pp. 704-710, 2016.

[18] B. Al Delail, H. Bhaskar, M. J. Zemerly, and M. Al-Mualla, "Robust Likelihood Model for Illumination Invariance in Particle Filtering," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 10, pp. 2836-2848, 2018.

[19] P. Bui Quang, C. Musso, and F. Le Gland, "Particle filtering and the laplace method for target tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 1, pp. 350-366, 2016.

[20] S. Wu, R. Li, Y. Shi, and Q. Liu, "Vision-Based Target Detection and Tracking System for a Quadcopter," *IEEE Access*, vol. 9, pp. 62043-62054, 2021.

[21] J. Hou, and B. Li, "RETRACTED: Swimming target detection and tracking technology in video image processing," *Microprocessors and Microsystems*, vol. 80, 103535, 2021.

[22] G. J. McLachlan, and S. Rathnayake, "On the number of components in a Gaussian mixture model," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 4, no. 5, pp. 341-355, 2014.

[23] L. Li, R. J. Hansman, R. Palacios, and R. Welsch, "Anomaly detection via a Gaussian Mixture Model for flight operation and safety monitoring," *Transportation Research Part C: Emerging Technologies*, vol. 64, pp. 45-57, 2016.

[24] Z. Ban, J. Liu, and L. Cao, "Superpixel Segmentation Using Gaussian Mixture Model," *IEEE Transactions on Image Processing*, vol. 27, no. 8, pp. 4105-4117, 2018.

[25] P. An, Z. Wang, and C. Zhang, "Ensemble unsupervised autoencoders and Gaussian mixture model for cyberattack detection," *Information Processing & Management*, vol. 59, no. 2, 102844, 2022.

[26] N. Ding, H. Ma, H. Gao, Y. Ma, and G. Tan, "Real-time anomaly detection based on long short-Term memory and Gaussian Mixture Model," *Computers & Electrical Engineering*, vol. 79, 106458, 2019.

[27] J. Gao, H. Zou, F. Zhang, and T. Y. Wu, "An intelligent stage light-based actor identification and positioning system," *International Journal of Information and Computer Security*, vol. 18, no. 1/2, 204-218, 2022.

[28] F. Zhang, T.-Y. Wu, J.-S. Pan, G. Ding, and Z. Li, "Human motion recognition based on SVM in VR art media interaction environment," *Human-centric Computing and Information Sciences*, vol. 9, 40, 2019.

[29] F. Zhang, T.-Y. Wu, and G. Zheng, "Video salient region detection model based on wavelet transform and feature comparison," *EURASIP Journal on Image and Video Processing*, vol. 2019, no. 1, 58, 2019.

[30] I. Shahin, A. B. Nassif, and S. Hamsa, "Emotion Recognition Using Hybrid Gaussian Mixture Model and Deep Neural Network," *IEEE Access*, vol. 7, pp. 26777-26787, 2019.

[31] R. Li, Z. Wang, C. Gu, F. Li, and H. Wu, "A novel time-of-use tariff design based on Gaussian Mixture Model," *Applied Energy*, vol. 162, pp. 1530-1536, 2016.

[32] S. Yin, Y. Zhang, and S. Karim, "Large Scale Remote Sensing Image Segmentation Based on Fuzzy Region Competition and Gaussian Mixture Model," *IEEE Access*, vol. 6, pp. 26069-26080, 2018.

[33] S. Kim, R. Haschke, and H. Ritter, "Gaussian Mixture Model for 3-DoF orientations," *Robotics and Autonomous Systems*, vol. 87, pp. 28-37, 2017.

[34] X.-F. Han, J. S. Jin, M.-J. Wang, W. Jiang, L. Gao, and L.-P. Xiao, "Video fire detection based on Gaussian Mixture Model and multi-color features," *Signal, Image and Video Processing*, vol. 11, no. 8, pp. 1419-1425, 2017.

[35] A. Das, U. R. Acharya, S. S. Panda, and S. Sabut, "Deep learning based liver cancer detection using watershed transform and Gaussian mixture model techniques," *Cognitive Systems Research*, vol. 54, pp. 165-175, 2019.

[36] M. Chen, and S. J. Ludtke, "Deep learning-based mixed-dimensional Gaussian mixture model for characterizing variability in cryo-EM," *Nature Methods*, vol. 18, no. 8, pp. 930-936, 2021.

[37] T. Kawabata, "Gaussian-input Gaussian mixture model for representing density maps and atomic models," *Journal of Structural Biology*, vol. 203, no. 1, pp. 1-16, 2018.

[38] S. K. Mohanty, and S. K. Udgata, "SATPAS: SINR-based adaptive transmission power assignment with scheduling in wireless sensor network," *Engineering Applications of Artificial Intelligence*, vol. 103, 104313, 2021.

[39] Y. Chen, H. Zhang, Y. Zheng, B. Jeon, and Q. M. J. Wu, "An improved anisotropic hierarchical fuzzy c-means method based on multivariate student t-distribution for brain MRI segmentation," *Pattern Recognition*, vol. 60, pp.778-792, 2016.

[40] J. Cao, M. G. Genton, D. E. Keyes, and G. M. Turkiyyah, "Exploiting low-rank covariance structures for computing high-dimensional normal and Student-t probabilities," *Statistics and Computing*, vol. 31, no. 1, 11081, 2021.

[41] M. W. Schwarz, W. B. Cowan, and J. C. Beatty, "An experimental comparison of RGB, YIQ, LAB, HSV, and opponent color models," *ACM Transactions on Graphics*, vol. 6, no. 2, pp. 123-158, 1987.

[42] M. Compare, and E. Zio, "Predictive Maintenance by Risk Sensitive Particle Filtering," *IEEE Transactions on Reliability*, vol. 63, no. 1, pp. 134-143, 2014.

[43] Z. Zhao, P. Feng, J. Guo, C. Yuan, T. Wang, F. Liu, Z. Zhao, Z. Cui, and B. Wu, "A hybrid tracking framework based on kernel correlation filtering and particle filtering," *Neurocomputing*, vol. 297, pp. 40-49, 2018.

[44] R. Alroobaea, S. Rubaiee, S. Bourouis, N. Bouguila, and A. Alsufyani, "Bayesian inference framework for bounded generalized Gaussian-based mixture model and its application to biomedical images classification," *International Journal of Imaging Systems and Technology*, vol. 30, no. 1, pp. 18-30, 2019.