Multiple feature fusion-based video face tracking for IoT big data

Zhifeng Liu1 | Jiayu Ou1 | Wenxiao Huo1 | Yejin Yan1 | Tianping Li2

1School of Physics and Electronics, Shandong Normal University, Jinan, Shandong, China
2Key Laboratory of Medical Physics and Image Processing in Shandong Province, Shandong Normal University, School of Physics and Electronics, Jinan, Shandong, China

Correspondence
Tianping Li, School of Physics and Electronics, Shandong Normal University, Jinan, Shandong 250300, China.
Email: sdsdltp@sdu.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Numbers: 61472220, 61572286

Abstract
With the advancement of Internet of Things (IoT) and artificial intelligence technologies, and the need for rapid application growth in fields, such as security entrance control and financial business trade, facial information processing has become an important means for achieving identity authentication and information security. However, in the process of acquiring facial feature information, face information is easily affected by factors, such as object occlusion, lighting changes, and similar backgrounds. In this paper, we propose a multifeature fusion algorithm based on integral histograms and a real-time update tracking particle filtering (PF) module. First, edge features and colour features are extracted, weighting methods are used to weight the colour histogram and edge features to describe facial features, and fusion of colour features and edge features is made adaptive by using fusion coefficients to improve face tracking reliability. Then, the integral histogram is integrated into the PF algorithm to simplify the calculation steps of complex particles and improve operational efficiency. Finally, the tracking window size is adjusted in real-time according to the change in the average distance from the particle centre to the edge of the current model and the initial model to reduce the drift problem and achieve stable tracking with significant changes in the target.
The results show that the algorithm improves video tracking accuracy, simplifies particle operation complexity, improves the speed, and has good anti-interference ability and robustness compared with extracting a single feature.

**KEYWORDS**
integral histogram, multifeature fusion, particle filtering, template drift, video face tracking

## 1 | INTRODUCTION

Face tracking technology is an important computer vision research area and is widely used in the fields of the Internet of Things (IoT) and artificial intelligence technologies. Face tracking is the process of predicting the motion information in subsequent frames based on the motion information in the initial frame of a face to determine the face trajectory and its morphological changes. Traditional face tracking mainly matches for a single feature, extracting a single colour feature, edge feature, texture feature or motion information, with low robustness.

With the rapid development of the IoT in the 21st century, target detection and target tracking are widely used in the IoT and artificial intelligence technologies, intelligent surveillance, Intel-Link traffic, human–machine interfaces, and so forth, and play a crucial role in electronic intelligence, security, military and other fields. Therefore, the current research on multiple features has far-reaching significance, and facial recognition has the advantages of simple acquisition, difficult to copy, and stealing complex; however, real-time video face tracking is still a very challenging problem in the field of computer vision due to the influence of lighting environment and speed conditions in actual motion, which makes the tracking target blurred and of poor quality, and is sensitive to external material occlusion, acquisition angle and image size. How to quickly and accurately detect faces and implement real-time video face tracking algorithms has become an important research process for face information processing.

A series of studies have been conducted on particle filter-based target tracking technology and achieved significant results, but there are still shortcomings. The PF-FA algorithm uses Hue, Saturation, and Value (HSV) colour features to model the tracking target and introduces the algorithm to the resampling stage of particle filtering (PF). Although the error is reduced compared with the standard PF algorithm, the algorithm has the disadvantages of slow convergence, particle scattering, and easy to be trapped in the local optimum, which makes the error has an incremental trend. Literature fuses colour and Local Binary Pattern (LBP) texture features to track valid blocks when partial occlusion is detected, and when complete occlusion is detected, there is a large uncertainty about the position of the target after reemergence, and the position estimation is easily misestimated. Literature reextracts the occlusion features by Speeded Up Robust Features (SURF), but when the resolution is low, the SURF feature points of the target are sparse and prone to matching failure, and the method is vulnerable to target deformation. Li et al. took advantage of the discriminative scale of the discriminative-scale space tracking (DSST) algorithm to improve the effect of long-time tracking and scale adaptive tracking, but the problem of occlusion still cannot be solved well. To solve the occlusion
problem, an occlusion judgement is added to the DSST algorithm, and a detector is used to
detect the target position when the target is occluded, improved tracking accuracy under
occlusion, but tracking performance needs to be improved.

In this paper, the video face tracking method is improved and refined, in which colour
features and edge features are adaptively fused to improve the face tracking reliability, the
integral histogram method improves efficiency, the size of the tracking window is adjusted
in real-time according to the average distance from the particle centre to the edge of the current
model and the initial model and the change in the tracking target, the tracking module is
updated to reduce the drift problem, and finally, the face video tracking is tested using a data
set to test the method. The experimental data show that the method accuracy is further im-
proved, the particle operations complexity is simplified, and it has good stability and robustness
under the influence of video with object occlusion, lighting changes, and similar backgrounds
and finally achieves accurate real-time human face tracking.

The rest of this paper is organized as follows. In Section 2, two important PF algorithm
processes for video face tracking are briefly described. In Section 3, the improved
resampling-based PF algorithm proposed in this paper, as well as the methods for building
the image integration histogram and fusing multiple features are described in detail.
Section 4 presents the face tracking system, compares the experimental results qualitatively
and quantitatively. Finally, the paper is summarized, and future tracking algorithms and
future work are envisioned.

2 | PARTICLE FILTERING ALGORITHM

PF is characterized by high accuracy, as it is not limited by noise and system models. Problems that cannot be solved by traditional analysis methods can be solved with the help of PF simulation, and in recent years, PF algorithms have achieved great success in the field of target tracking technology. PF is a nonlinear filtering method based on Monte Carlo simulation, using Bayesian inference and importance sampling theory as the basic framework. The Bayesian inference process uses a collection of particles to represent the posterior probability distribution of the random states of the target and to estimate the state of the nonlinear system, and some estimation criteria are used to estimate the state values of the target. The algorithm can be considered optimal when it has the smallest mean-square error. The essence of importance sampling is to select larger weights for particles with high confidence and thus determine the probability of becoming a target based on the specific distribution of particle weights.

In this paper, we study the tracking of a single face, and the algorithm obtains the max-
imum likelihood observation through the ideal likelihood observation peak distribution and
uses the given system model to estimate the present moment state from the previous moment
state to obtain the closest estimate of the true target state, enhance the observation effect and
the adaptive capability of the target tracking, and predict the posterior probability $p(x_t | y_{1:t})$ of the target state at time $t$. The PF steps are divided into four processes: initial state, target prediction, state correction and resampling.

Initial state: A large number of particles are used to simulate the uniformity of the particle distribution.

Prediction phase: PF generates a large number of particles based on the probability distribution of $x_{t-1}$, and the prior probability density of the state is predicted using the state
transfer equation and the control volume to make a preliminary estimate of the system at the next moment through the existing prior knowledge. The prediction equation is shown as

\[ p(x_t | y_{1:t-1}) = \int p(x_t | x_{t-1}) p(x_{t-1} | y_{1:t-1}) \, dx_{t-1}. \]  

(1)

Correction stage: The posterior probability density is introduced using the similarity between the possible and real target states and the observation equation, and the closer it is to the real state, the larger the particle weight is, otherwise, the smaller the weight is that is assigned. The update equation is shown as

\[ p(x_t | y_{1:t}) = \frac{p(y_t | x_t) p(x_t | y_{1:t-1})}{p(y_t | y_{1:t-1})}. \]  

(2)

Using a recursive approach to reduce the complexity when calculating the weights, the signal is processed by sequential importance sampling, and the distribution condition of the importance function at moment \((t-1)\) is used as the importance function at moment \(t\). The particle weight \(\{w^1_t, w^2_t, \ldots, w^N_t\}\) recursive form can be expressed as

\[ w^i_t \propto \frac{p(x_{0:t-1}^i | y_t)}{q(x_{0:t-1}^i | y_t)} = \frac{p(y_t | x_t^i) p(x_t^i | x_{t-1}^i) p(x_{t-1}^i | y_{t-1})}{q(x_{t}^i | x_{t-1}^i, y_t) q(x_{t-1}^i | y_{t-1})} = w_{t-1}^i p(y_t | x_t^i). \]  

(3)

Resampling: After sequential importance sampling, in the iterative process the diversity of particles is lost due to the difference between the importance function and the posterior distribution. To alleviate the degradation problem, based on adjusting the importance function, the particles are filtered according to the particle weights, and the importance weight of the particles is replicated instead of the number of low-weight particles for propagation to obtain the required true state. In the PF algorithm, it is not necessary to consider resampling at each iteration, but the effective number of particles \(N_{\text{eff}}\) is cited to determine the degradation size of the algorithm. If \(N_{\text{eff}} < N_{\text{th}}\), the resampling algorithm is used, \(N_{\text{th}}\) is the threshold value, which is generally taken as \(2N/3\), and the value is usually approximated by \(\hat{N}_{\text{eff}}\).

\[ \hat{N}_{\text{eff}} = \frac{N}{\sum_{i=1}^{N} \omega(x_{0:t})^2}. \]  

(4)

The posteriori density can be expressed as

\[ p(x_t | y_{1:t}) \approx \sum_{i=1}^{N} \bar{w}_t^i \delta(x_t - x_t^i). \]  

(5)

Finally, the resampled particles are brought into the state transfer equation again for the prediction process. The general framework structure of the PF algorithm is shown in Figure 1.
3 | IMPROVED PARTICLE FILTERING ALGORITHM

Although the PF algorithm has outstanding advantages in practical applications, the traditional resampling algorithm only resamples the particles according to certain particle weight rules, without considering the particles distributed in high probability regions, and if Equation (4) is applied to approximate the size of the weights, this will reduce the number of measured particle samples and invalidate the small weight samples, and there are disadvantages, such as decreasing weights, lack of samples during resampling, and inability to choose the optimal probability density, leading to particle degradation and new large estimation errors. In response to the negative effects created by resampling techniques, an improved resampling algorithm is proposed to replace the traditional particle resampling algorithm, improve the PF algorithm performance, adaptively adjust the particles distributed in the high likelihood region according to the accuracy factor value reflecting the measurement noise, increase existing knowledge and the same region of likelihood, and improve the algorithm robustness.

3.1 Specific steps for resampling

Preprocessing stage: Choose time \( \{x_k\}_{i=1}^{N} \) to obtain the required particle set, obtain the particle pair according to the formula \( w_{k+1}^i > w_k^i \), reset the formula, reweight the new particles to ensure that the particles meet the needed particle weights, and through this process, obtain the new particle pair.

Particle classification: Set the weight \( w_l \) and the threshold \( w_h \) to satisfy the relation \( 0 < w_l < w_h \) by which the particles are divided into two major parts.

The range class A particle values: \( \{x_k, w_k^i\}; w_k^i \leq w_l, w_k^i \geq w_h, i = 1, 2, ..., N \}. The range class B particle values: \( \{x_k, w_k^i\}; w_l < w_k^i < w_h, i = 1, 2, ..., N \}. Gravity particles and small weight particles form class A particles, and stable medium weight particles form class B particles. The sum of two particles is the sum of all particles.
Handling particles: Resampling class A particles using linear equations:

\[ x_n = x_s + KL(x_a - x_s), \] (6)

\[ L = \left[ \frac{1}{N_{iw}} \right]^{\frac{1}{K}}, \] (7)

where \( x_n \) represents the new particle, and the heavy particle with more times and repetitions is selected; \( x_s \) is the small and medium weight particle; \( K \) is the step coefficient, and the value of \( K \) is determined according to the actual needs during the experiment. A moderate step size is chosen in the straight line \( L \), the new weight particles are generated by \((x_a - x_s)\) and compared with the original particle weight. If the new particle weight is small, we reduce \( L \) by 1/2, effectively improve the weights of small weight particles to ensure that the particles meet the required particle weights and solve the particle degradation problem, while generating new particles according to the new \( L \).

### 3.2 Simulation experiments and data analysis

Simulation experiments verify the effectiveness of the improved resampling technique for state estimation and tracking nonlinear systems using the traditional resampling particle filter (TRPF) algorithm and improved resampling particle filter (IRPF) algorithm, respectively. The experiments use the unary nonstationary growth model (UNGM), and the state model and observation model equations of the simulation object are as follows:

\[ x_n = 0.5x_{n-1} + \frac{25x_{n-1}}{1 + x_{n-1}^2} + 8\cos[1.2(n - 1)] + u_n, \] (8)

\[ y_n = \frac{1}{20}x_n^2 + v_n, \] (9)

where \( u_n \) and \( v_n \) are zero-mean Gaussian noise, the simulation takes the number of particles as 100, the step size \( n = 50 \), the coefficient \( K = 0.4 \), and the measurement noise variance is 1 for 50 iterations.

The specific state estimation curves are shown in Figure 2. Under the same simulation environment, the TRPF algorithm with traditional resampling has a poor estimation effect because of the small number of particles, the concentration of the likelihood function is higher than that of the state transfer density function, and the particles are severely depleted, which leads to a poor match between the state estimation and the real state. However, the state estimation of the improved PF algorithm IRPF proposed in this paper can match the real state well with many effective samples and high estimation accuracy, which can effectively suppress sample degradation and ensure particle diversity compared with the TRPF algorithm.

To better compare the performance of the improved resampling PF algorithm in terms of state estimation, the root mean-square error (RMSE) is used as a measure, and the RMSE is calculated as

\[ RMSE = \left[ \frac{1}{T} \sum_{k=1}^{T} \left( x_k - \hat{x}_k \right)^2 \right]^{\frac{1}{2}}, \] (10)
where $T$ represents the time step of an experiment, and $x_k$ and $\hat{x}_k$ represent the true and estimated values at moment $K$. The smaller the RMSE value, the better the effect.

Figure 3 shows the comparison of RMSE results for 50 independent experimental simulations.

A comparative plot of the RMSE results shows that the TRPF algorithm using traditional resampling has a significantly higher RMSE and greater fluctuations than the IRPF algorithm in this paper, which uses a shorter state estimation time. This section is reworked to address the problems arising from resampling, and the final results show that the improved PF algorithm has the smallest mean-square error and is closer to the true state estimate with good robustness.
4 | FACE TRACKING SYSTEM

4.1 | Power model

The tracking algorithm is based on a moving target model, and the autoregressive process model \( AR(p) \) has been widely used to create such a dynamic model. The \( AR(p) \) model is abbreviated as

\[
x_t = \varphi_0 + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \cdots + \varphi_p x_{t-p} + \varepsilon_t.
\]

The target displacement, noise, velocity and acceleration are all attributes of the moving target. Therefore, a second-order regression process is used to describe the motion law. The \( AR(2) \) model is

\[
x_t = \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \varepsilon_t,
\]

where \( \{x_t, t = 0, \pm 1, \pm 2, \ldots \} \) is the time series, \( \varphi_1 \) and \( \varphi_2 \) are the drift coefficient matrices, \( \{\varepsilon_t\} \) is the white noise series, and \( \varphi(B)x_t = \varepsilon_t \), where \( \varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 \). Let the linear transfer function be \( \psi(B) \). The smooth solution of the second-order regression process is

\[
x_t = \psi(B)\varepsilon_t.
\]

These parameters can be obtained empirically or by training video sequences, and the weight coefficients \( x_t \) of the function \( \psi(B) \) can be found. The feasibility of the model is guaranteed, and whether the particle propagation is reasonable needs to be verified in the particle update process.

4.2 | Observation model

In PF, observation is feature based. This section focuses on facial feature extraction by combining colour features and edge features to track the target face.

1. **Observation model based on colour features**: The colour histogram shows the proportion of different colours in an image. Colour information has the features of translation, rotation and computational simplicity compared with geometric features and is currently widely used in face detection and tracking problems. In this paper, human skin colour is used as a coarse detection and localization method for human faces, and the colour distribution is described more robustly by using HSV colour space modelling and computing histograms using only face elements. For skin colour detection, the skin colour points are judged according to a threshold value, so that the distribution of the added particles and the update of the weights are more adapted to the algorithm. The colour histogram is shown in Figure 4B. The effect of illumination variation is reduced to some extent.

The target region weighted colour histogram is constructed as

\[
P_l(n) = C_l \sum_{i=1}^{N} K_E \left( \frac{x_i - x_0}{a} \right) \delta [h(x_i) - n],
\]
\[ C_l = \frac{1}{\sum_{i=1}^{N} K_E \left( \frac{x_i - x_0}{a} \right)} \]  

where \( C_l \) is a normalization constant, \( N \) is the total number of pixels in the target region, and \( K_E(\cdot) \) is an Epanechnikov kernel contour of radius \( a \in [1, M] \), indicating that the closer a pixel is to the centre of the target, the more likely it is that it belongs to the target, and the Epanechnikov kernel is described as follows:

\[ K_E(x) = \begin{cases} C(1 - \|x\|^2), & \|x\| < 1, \\ 0, & \|x\| \geq 1, \end{cases} \]  

where \( \delta_a \) is the Kronecker delta function, function \( h(x_i) \) maps the pixel positions to the corresponding histogram face elements, and \( n \) is the histogram segment number index value.

When the colour distributions of the reference target template \( q_l(n) \) and the candidate target template \( p_l(n) \) are calculated, the Bhattacharyya distance \( d_l \) can be used to measure the similarity of the two distributions as follows.

\[ d_l = \sqrt{1 - \rho_l[p_l(n), q_l(n)]}, \]  

where \( \rho_l[p_l(n), q_l(n)] \) is the Bhattacharyya dispersion coefficient:

\[ \rho_l[p_l(n), q_l(n)] = \sum_{n=1}^{M} \sqrt{p_l(n)q_l(n)}. \]  

The Bhattacharyya coefficient measures the correlation, \( M \) is the total number of colour-weighted histogram bins, and \( x_0 \) is the centre of the observed region. The face colour observation likelihood probability density function \( p(y|x) \) is defined as

\[ p(y|x) = \frac{1}{\sqrt{2\pi \sigma_l}} e\left(-\frac{d_l^2}{2\sigma_l^2}\right), \]  

FIGURE 4 Colour histogram [Color figure can be viewed at wileyonlinelibrary.com]
where $\sigma_l$ is the Gaussian variance, which is taken as 0.2 in the experiment. The larger the value of Equation (19), the greater the colour similarity between the candidate target and the face template, and the greater the possibility that it is a real face target.

2. **Observation model based on edge features**: Another key information cue of the face is the edge feature,\(^3\) which is characterized by insensitivity to illumination changes and similar face colour backgrounds, and the edge direction histogram reflects the edge and texture features of the face.

The histogram of edge directions is shown in Figure 5B. First, the original image is greyed out where the edge points are extracted by the operator operation, and the shape features of the image are obtained using the statistical edge point direction histogram. In this paper, the Sobel operator is used to detect the edge contours, the amplitude of the gradient $g$ is calculated by Equation (20), and the gradient direction $\theta$ is calculated by Equation (21).

\[
g = \sqrt{g_x^2 + g_y^2},
\]

\[
\theta = \arctan\left(\frac{g_y}{g_x}\right),
\]

where $g_x$ and $g_y$ are the horizontal and vertical gradients of the image, respectively.

\[
\rho_m [p_m(n), q_m(n)] = \sum_{n=1}^{M} \sqrt{p_m(n)q_m(n)}.
\]

The face edge likelihood function is defined as

\[
p(y_m|x) = \frac{1}{\sqrt{2\pi} \sigma_m} e^{-\frac{1-\rho_m}{2\sigma_m^2}},
\]

where $\sigma_m$ is the Gaussian variance, which is taken as 0.3 in the test.
4.3 Strategy of this paper

On the basis of the improved resampling-based PF, a multifeature fusion algorithm based on an integral histogram and a real-time update tracking particle filter module is proposed for video face tracking, which not only reduces the algorithm running time and drift problem but also improves the tracking accuracy and robustness to achieve a real-time tracking effect considering the long computation time.

1. Adaptive multifeature fusion: Video face tracking finds the region in the current frame image that has the most similar template features to the tracked target. Nummiaro extracted colour to verify the tracking effect, and the experiment showed that the single colour information is insensitive to object occlusion but very sensitive to illumination. If only a single feature method is used to track the target, it is often not ideal for the final tracking effect. In this paper, we combine the edge features, which are insensitive to illumination changes and similar face colour background and propose a strategy of multifeature information fusion based on improved resampling PF to calculate the particle weights. First edge features and colour features are extracted in the video data set, and a weighting method is used to weight the colour and edge features histogram to describe the facial features. Then, the fusion coefficient is used to adaptively fuse the colour features and edge features. Adaptive PF removes the excess number of particles to improve the computational speed and the reliability of the face observation model to make up for the respective deficiencies of colour features and edge features, which achieves complementary information between features, and the impact on the tracking results is smaller when the target changes, achieves video target tracking in complex backgrounds. The observation model $p(y_t | x_t)$ extracts colour features and edge features. The whole face likelihood function is defined as

$$p(y_t | x_t) = \theta_l p(y_t | x_t) + \theta_m p(y_m | x_t),$$

where $\theta_l + \theta_m = 1$, the weight $\theta_m$ is large when the edge information is reliable in tracking; otherwise, vice versa. $p(y_t | x_t)$ is the likelihood function of the colour feature, and $p(y_m | x_t)$ is the likelihood function of the edge feature.

The particle weights are updated as

$$w^t_i \propto w^{t-1}_i p(y_t | x^t_i) = w^t_i [\theta_l p(y_t | x_t) + \theta_m p(y_m | x_t)].$$

2. The integral histogram of the image: In the particle filter tracking algorithm, when there are many particles, the computation time consumed will be larger as the number of particles increases. To solve this problem, the integral histogram is integrated into the PF algorithm, which adds and subtracts the integral histogram of the four vertices of the rectangular region where the particles are located, which replaces the inefficient statistical work of the ordinary algorithm, simplifies the calculation steps of complex particles, solves the problem of long histogram calculation time caused by too many particles and improves the operational efficiency.

To verify the effectiveness of the integral histogram, we analyse the effect of the ordinary histogram and the integral histogram on the computational speed with different numbers of particles.
As seen in Figure 6, the computational time of the ordinary histogram is temporarily shorter when the number of particles is low, and as the number of particles increases, that is, the area of the region where the particles are located increases, the integral histogram used in this paper converges faster and the computational time is shorter. The strategy of using the integral histogram in this paper is more robust than the ordinary histogram because the number of particles selected in this paper is higher.

3. Calibration model: Since the face state continues to change during the tracking process, current face information is reflected in real-time by updating the template, and frequent updating of the template; this causes the real target in the template to be replaced by the background and the drifting phenomenon \(^{40}\) will occur, which eventually leads to tracking failure. To mitigate drift, \(^{41}\) in this paper, we propose a template correction technique to correct the template matching result according to the change in the average distance from the particle centre to the edge of the current model and the initial model, adjust the size of the tracking window in real-time, and stably track face targets with significant dimensional changes. The template adaptive correction method used in this paper is as follows:

\[
H_{\text{new}} = \frac{1}{\tau}H_{\text{old}} + \left(1 - \frac{1}{\tau}\right)H_{\text{current}},
\]

where \(H_{\text{new}}\) is the new reference histogram, \(H_{\text{old}}\) is the initial reference histogram and \(H_{\text{current}}\) is the current reference histogram. \(\tau\) is the constant \((0 \leq 1/\tau \leq 1)\), and the smaller the \(\tau\) is, the more effective the histogram update is. The algorithm can also update the tracking module in real-time according to the changes in the tracking target and make adaptive adjustments to the tracking window, thus improving the accuracy and precision of real-time tracking.
4.4 Experimental results

To verify the tracking effectiveness of the method in this paper, three sets of more complex experimental videos, including object occlusion, illumination changes and similar backgrounds, are used to track two different algorithms, namely, the colour-only-based tracking algorithm and the tracking algorithm in this paper, visual verification of the improved algorithm's tracking superiority in different complex contexts. We manually select a rectangular window as the matching template in the window and set \( N = 100 \). The experimental platform is based on Visual Studio 2010 opencv2.4.8 and uses Euclidean distance to measure the target tracking results. The data set is used for the visual tracker benchmark, and the complete benchmark contains 100 sequences from the recent literature. Three video sequences were selected for testing, and the information of the video sequences is shown in Table 1. At the same time, four different video frames were randomly selected from each sequence, and these four frames were used as the object of the experiment for comparison and analysis.

1. **Test 1**: Test 1 investigates the effect of face tracking under object occlusion.

   Figure 7 shows the tracking results based on colour features only when the face is occluded by other objects. When unobscured, the face position can be accurately located. However, when the face is partially occluded by the book, the weights change drastically, the face region information is reduced, and the tracking algorithm based on colour features alone tends to lose the face in Frame 167 of Figure 7B. When the book moves to change the occlusion position, the localization ability is completely lost. When the occluded object is completely removed, it still fails to locate the real position of the face in time, and the tracking effect is poor.

   The tracking results of the method used in this paper are shown in Figure 8. In Frames 167 and 268 in Figure 8B,C, the tracking box can accurately locate the face region in the cases of object occlusion and when the occlusion moves, and the algorithm continues to track the face accurately after the occlusion is withdrawn, mainly because the improved algorithm can take full advantage of each feature and can track in real-time using template updates.

   The comparison curves of the RMSE between single and multifeatures are given in Figure 9. The tracking error of the algorithm in this paper is significantly lower than the tracking error based on colour features only. This is mainly due to the proposed fusion scheme, which enhances the tracking reliability.

2. **Test 2**: Test 2 investigates the effect of face tracking under different illumination conditions.

   Tracking based on colour features is better under different illumination only before the light changes, and when the changes are significant, the tracking results deviate from the real face

| Test | Sequence | Frame size | Sequence characteristics | Total frames | Total frames in this paper |
|------|----------|------------|--------------------------|--------------|---------------------------|
| Test 1 | FaceOcc2 | 480 × 360  | Object occlusion         | 812          | 310                       |
| Test 2 | Trellis  | 480 × 360  | Illumination variation   | 569          | 182                       |
| Test 3 | David    | 480 × 360  | Similar background       | 770          | 119                       |
position. In Frames 36 and 138 in Figure 10B,D, when the experiment is based on colour features only, tracking fails. This is because the tracking algorithm based on colour features only is very sensitive to illumination changes.

Figure 11 shows that the adaptive multifeature fusion strategy proposed in this paper can track faces well even when the illumination changes because the combined edge features are
insensitive to the change in face surface colour, which compensates for the effect of illumination changes on the tracking results and reduces error.

Figure 12 shows our quantitative analysis of the mean-square error tracking results. The results show that the adaptive fusion strategy and the real-time update tracking particle filter module in this paper have a smaller overall error value with fluctuation through comparative analysis of each frame, are insensitive to the change in a light colour and have a better tracking effect compared with the algorithm that simply extracts single features.

3. Test 3: Test 3 investigates the effect of face tracking in backgrounds similar to the face.

Figure 13 shows the tracking results based on colour features only when a background similar to the face is present. If only colour features are used for PF, the algorithm is insensitive to facial expressions and pose changes, but when an object similar to the face appears in the video, for example, in Frame 82 of Figure 13C, the face is disturbed by the hand, which causes the algorithm to misjudge, and the rectangular box cannot locate the correct face position and tracks the hand. The face can be tracked only when the similar background disappears.

Figure 14 shows the tracking effect of the method used in this paper, which can track faces stably even with a similar background. It can also be seen in Figures 9 and 12 that our improved algorithm has very small error, the reason is that the improved algorithm adopts a self-adjusting approach, observing which feature can play a greater advantage and giving it a greater weight, solving the defects caused by a single feature, and having a better anti-interference ability and robustness compared with the original algorithm’s feature averaging approach.

4. Test 4: Comparison with the popular algorithm regions with convolutional neural networks (R-cnn)
Tests 1–3 demonstrate the effectiveness of the improved method by comparing the improved adaptive multifeature fusion algorithm with the original algorithm in a cross-sectional manner. Test 4 further verifies the superiority of our algorithm by conducting a cross-sectional comparison of the improved algorithm with the popular algorithm R-cnn in recent years using the same data set and test equipment.
As shown in Figure 15, the improved PF algorithm of this paper and R-cnn are experimented with the video sequences in Test 3. By comparison, it is found that the tracking accuracy of the improved algorithm of this paper is higher than that of the R-cnn algorithm, and the computation time required by the algorithm of this paper is shorter and more robust as the number of frames increases.

A comparison of the experimental runtimes of the different algorithms is shown in Figure 16. The running speed based on colour features alone is the fastest, but the tracking accuracy is poor. The above video experiment results of the adaptive multifeature fusion tracking algorithm based on colour and edge features can accurately track the face target.
while ensuring a shorter runtime. Considering the long computational time, this paper uses an integral histogram and correction model to compute the facial features based on adaptive multifeature fusion,\(^{45}\) which not only reduces the algorithm running time and drift problem but also improves the tracking accuracy and robustness to achieve a real-time tracking effect.

5 | CONCLUSION

For target tracking in complex backgrounds, it is often difficult to achieve robust tracking with a single feature algorithm. In this paper, we use the fused colour and edge information to describe the observation information of the target, combine the fused observation model to the PF algorithm based on improved resampling, and introduce the integral histogram and real-time update tracking PF module to make the algorithm converge faster, have shorter computational time and track more stably. Experiments show that the method in this paper can precisely locate the face position under conditions of object occlusion, lighting change, similar colour background and so forth, and save time. For the face occlusion problem, the tracking algorithm used in this paper is most effective by updating the template.

However, the proposed algorithm has some limitations. For example, hardware-assisted methods need to be considered, the initial template of the face should be well defined or the target model should be anchored to the first frame. The target tracking algorithm faces many challenges, and in the future, we will research faster and more robust tracking algorithms.

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

AUTHOR CONTRIBUTIONS

Formal analysis: Zhifeng Liu and Jiayu Ou. Methodology: Zhifeng Liu and Wenxiao Huo. Supervision: Tianping Li. Writing—original draft: Zhifeng Liu and Yejin Yan.

ORCID

Zhifeng Liu http://orcid.org/0000-0001-6215-3852
Jiayu Ou http://orcid.org/0000-0002-4498-5185
Wenxiao Huo http://orcid.org/0000-0003-1995-9556
Yejin Yan http://orcid.org/0000-0003-4578-8148
Tianping Li http://orcid.org/0000-0002-6710-9436

REFERENCES

1. Cui Y, Zhang J, He Z, Hu J. Multiple pedestrian tracking by combining particle filter and network flow model. Neurocomputing. 2019;351:217-227. https://doi.org/10.1016/j.neucom.2019.03.077
2. Khalili A, Soliman AA, Asaduzzaman M. Quantum particle filter: a multiple mode method for low delay abrupt pedestrian motion tracking. Electron Lett. 2015;51:1251-1253. https://doi.org/10.1049/el.2015.1013
3. Lou J, Zhu W, Wang H, Ren M. Small target detection combining regional stability and saliency in a color image. Multimed Tools Appl. 2016;76:14781-14798. https://doi.org/10.1007/s11042-016-4025-7
4. Bos R, de Waele S, Broersen PMT. Autoregressive spectral estimation by application of the burg algorithm to irregularly sampled data. IEEE Trans Instrum Meas. 2002;51:1289-1294. https://doi.org/10.1109/tim.2002.808031
5. Beck G, Duong T, Lebbah M, Azzag H, Cérin C. A distributed approximate nearest neighbors algorithm for efficient large scale mean shift clustering. *J Parallel Distrib Comput*. 2019;134:128-139. https://doi.org/10.1016/j.jpdc.2019.07.015

6. Wang Z, Miao X, Huang Z, Luo H. Research of target detection and classification techniques using millimeter-wave radar and vision sensors. *Remote Sens*. 2021;13:1064. https://doi.org/10.3390/rs13061064

7. Xu Y, Jiang X, Li F. Improved anti-occlusion target tracking algorithm based on compressive particle filtering. In: *8th International Symposium on Computational Intelligence and Design (ISCID)*. Hangzhou, China: Institute of Electrical and Electronics Engineers. 2015:527-531. https://doi.org/10.1109/ISCID.2015.144

8. Lin S, Lin J, Chuang C. Particle filter with occlusion handling for visual tracking. *IET Image Process*. 2015;9(11):959-968.

9. Li T, Wu Q, Ding S, et al. A long-term tracking algorithm based on TLD and fDSST. *Electro-Opt Control*. 2019;26(4):44-48. 70.

10. Zhao Q, Zhang B, Feng W, Du Z, Zhang H, Sun D. Long-term real-time object tracking based on multi-scale local correlation filtering and global re-detection. *Computing*. 2020;102(5):1487-1501. https://doi.org/10.1007/s00067-020-00807-8

11. Tang X, Chen Z, Fu Y. Anti-occlusion and re-tracking of real-time moving target based on kernelized correlation filter. *Optoelectron Eng*. 2020;47(1):54-63. https://doi.org/10.12086/oe.2020.190279

12. Fan D, Gong J, Zhang S, et al. A transient composition tracking method for natural gas pipe networks. *Energy*. 2021;215:119131. https://doi.org/10.1016/j.energy.2021.119131

13. Chai Y, Shin S, Chang K, Kim T. Real-time user interface using particle filter with integral histogram. *IEEE Trans Consumer Electron*. 2010;56:510-515. https://doi.org/10.1109/tce.2010.5505963

14. Lee T-Y, Shen H-W. Efficient local statistical analysis via integral histograms with discrete wavelet transform. *IEEE Trans Visual Comput Graphics*. 2013;19:2693-2702. https://doi.org/10.1109/tvcg.2013.152

15. Wang X, Han C-Z. A multiple model particle filter for maneuvering target tracking based on composite sampling. *Acta Autom Sin*. 2014;39:1152-1156. https://doi.org/10.3724/sp.j.1004.2013.01152

16. Fu C, Lin F, Li Y, Chen G. Correlation filter-based visual tracking for UAV with online multi-feature learning. *Remote Sens*. 2019;11:549. https://doi.org/10.3390/rs11050549

17. Zhang S, Zhuo L, Zhang H, Li J. Object tracking in unmanned aerial vehicle videos via texture features for adaptive real-time object tracking. *Adv Mater Res*. 2015;1109/lgrs.2016.2581207

18. Wang J, Yagi Y. Integrating color and shape-texture features for adaptive real-time object tracking. *IEEE Trans Image Process*. 2008;17:235-240. https://doi.org/10.1109/tip.2007.914150
27. Zhu M, Xi C, Li Y, et al. Truck active reversing control strategy based on modified particle filter and multi-sensors environment perception. *IET Intell Transp Syst*. 2019;13:1057-1068. https://doi.org/10.1049/iet-its.2018.5211
28. Bousetouane F, Dib L, Snoussi H. Improved mean shift integrating texture and color features for robust real time object tracking. *Vis Comput*. 2012;29:155-170. https://doi.org/10.1007/s00371-012-0677-0
29. Boucheix J-M, Lowe RK. An eye tracking comparison of external pointing cues and internal continuous cues in learning with complex animations. *Learn Instr*. 2010;20:123-135. https://doi.org/10.1016/j.learninstruc.2009.02.015
30. Li Z, Zhu H, Zhou C, et al. A color consistency processing method for HY-1C images of Antarctica. *Remote Sens*. 2020;12:1143. https://doi.org/10.3390/rs12071143
31. Zheng C, Chen Z, Huang D. Fault diagnosis of voltage sensor and current sensor for lithium-ion battery pack using hybrid system modeling and unscented particle filter. *Energy*. 2020;191:116504. https://doi.org/10.1016/j.energy.2019.116504
32. Wen J, Wang Z. Sensor-based adaptive activity recognition with dynamically available sensors. *Neurocomputing*. 2016;218:307-317. https://doi.org/10.1016/j.neucom.2016.08.077
33. Li Y, Lu Z, Gao Q, Li J. Particle filter and mean shift tracking method based on multi-feature fusion. *Null*. 2010;32:411-415. https://doi.org/10.3724/spj.1146.2008.01740
34. Schreiber D. Robust particle tracking with drift correction. *Pattern Recognit Lett*. 2007;28:1483-1491. https://doi.org/10.1016/j.patrec.2007.03.007
35. Wang Y, Chen F-X, Guo H-X. Kernel spatial histogram target tracking based on template drift correction. *Acta Autom Sin*. 2012;38:430-436. https://doi.org/10.3724/sp.j.1004.2012.00430
36. Dai Y, Liu B. Robust video object tracking via Bayesian model averaging-based feature fusion. *Opt Eng*. 2016;55:083102. https://doi.org/10.1117/1.oee.55.8.083102
37. Rodriguez P, Wohlberg B. Incremental principal component pursuit for video background modeling. *J Math Imaging Vis*. 2015;55:1-18. https://doi.org/10.1007/s10851-015-0610-z
38. Nummiaro K, Koller-Meier E, Van Gool L. An adaptive color-based particle filter. *Image Vis Comput*. 2003;21(1):99-110. https://doi.org/10.1016/S0262-8856(02)00129-4
39. Roy P, Chandra Pal S, Arabameri A, et al. Novel ensemble of multivariate adaptive regression spline with spatial logistic regression and boosted regression tree for gully erosion susceptibility. *Remote Sens*. 2020;12:3284. https://doi.org/10.3390/rs12203284
40. Mesas-Carrascosa F-J, Pérez-Porras F, Meroño de Larriva J, et al. Drift correction of lightweight microbolometer thermal sensors on-board unmanned aerial vehicles. *Remote Sens*. 2018;10:615. https://doi.org/10.3390/rs10040615
41. Dong-Yue C, Yuan-Chen Q. Robust object tracking based on adaptive feature selection. *Inf Technol J*. 2013;12:7325-7330. https://doi.org/10.3923/itij.2013.7325.7330
42. Liu L, Xi Z, Sun Q. Multi-sensor tracking and collaboration based on spatial particle filter. *J Visual Commun Image Representation*. 2019;59:316-326. https://doi.org/10.1016/j.jvcir.2018.12.050
43. Li T, Zhou P, Liu H. Multiple features fusion based video face tracking. *Multimed Tools Appl*. 2019;78:21963-21980. https://doi.org/10.1007/s11042-019-7414-x
44. Ge Z, Zhao L, Yao Y, Song Z, Gao F. Utilizing transition information in online quality prediction of multi-phase batch processes. *J Process Control*. 2012;22:599-611. https://doi.org/10.1016/j.jprocont.2012.01.004
45. Suchithra R, Ezhilsabareesh K, Samad A. Optimization based higher order sliding mode controller for efficiency improvement of a wave energy converter. *Energy*. 2019;187:116111. https://doi.org/10.1016/j.energy.2019.116111

**How to cite this article:** Liu Z, Ou J, Huo W, Yan Y, Li T. Multiple feature fusion-based video face tracking for IoT big data. *Int J Intell Syst*. 2022;37:10650-10669. https://doi.org/10.1002/int.22702