Pedestrian reidentification based on feature fusion and metric learning

Hongbin Luo,
School of Artificial Intelligence, Gansu University of Political Science and Law,
Lanzhou, 730070
China

Abstract—**The pedestrian recognition in public environment is influenced by the pedestrian environment and the dynamic characteristic boundary factors, so it is easy to produce the tracking error. In order to improve the ability of pedestrian re-identification in public environment, we need to carry out feature fusion and metric learning, and propose pedestrian re-identification based on feature fusion and metric learning. The geometric grid area model of pedestrian recognition in public environment is constructed, the method of fuzzy dynamic feature segmentation is used to reconstruct the dynamic boundary feature point of pedestrian recognition in public environment, the method of bottom-up modeling is used to design the dynamic area grid model of pedestrian recognition in public environment, the design of dynamic area grid model is three-dimensional grid area, the grayscale pixel set of pedestrian recognition dynamic constraint under public environment is extracted, the boundary feature fusion is carried out according to the distribution intensity of grayscale, the image fusion and enhancement information processing of pedestrian recognition under public environment, and the method of 3D dynamic constraint is used to realize the local motion planning of pedestrian recognition under public environment, and the recognition feature fusion and learning of pedestrian recognition under public environment is realized according to the result of contour segmentation. The simulation results show that the method is used for pedestrian recognition again in public environment, and the fuzzy judgment ability of pedestrian dynamic edge features is strong, which makes the error controlled below 10 mm, and the fluctuation of pedestrian recognition again is more stable, the recognition accuracy is higher and the robustness is better.**

Keywords—**Feature extraction, feature fusion, metric learning, motion planning, pedestrian, recognition.**

I. INTRODUCTION

WITH the development of image processing technology, using image processing method to realize pedestrian dynamic tracking recognition in public environment, combining pedestrian attitude information and environmental information to analyze pedestrian dynamic characteristics in public environment, extracting pedestrian edge contour feature points and attitude details in public environment, realize the dynamic tracking recognition and attitude running feature judgment of pedestrian in public environment [1]. The study of pedestrian re-identification in public environment has good application value in the fields of sky-eye monitoring, criminal investigation and anti-riot. The basis of pedestrian dynamic tracking recognition in public environment is to extract and judge the boundary features of pedestrians in public environment, according to the 3D action characteristics of pedestrians in public environment, to plan the dynamic constraints of pedestrian re-identification in public environment, to establish the 3D dynamic space area of pedestrian re-identification in public environment, to carry out feature fusion and metric learning in 3D space area, to improve the dynamic tracking and adaptive recognition performance of pedestrians in public environment, and the research of feature fusion and measurement learning under the dynamic constraints of pedestrian re-identification in public environment has received great attention [2], [3].

In the research of pedestrian recognition, after receiving the pedestrian image manually labeled or automatically detected from the surveillance video, the system first needs to extract the feature expression vector which has a certain robustness to the influence of light and visual angle from the image, then uses a certain measurement model to calculate the distance or similarity between the two images, and then determines whether the two images contain the same or not A pedestrian. Because there may be a large number of pedestrians to match in the video surveillance scene, the distance or similarity between

the given retrieval image and all candidate images needs to be calculated. By sorting the results, the candidate images with smaller distance (or higher similarity) can be placed in the front position. From this point of view, pedestrian recognition can also be considered as a fine-grained image retrieval task. So similar to image retrieval, in the research literature of pedestrian recognition, the given pedestrian image for retrieval is usually called the probe image, and the image to be matched is called the gallery images. Because the extraction of feature descriptors and the distance / similarity model are the two key links in the pedestrian recognition system, the current pedestrian recognition algorithm is basically around these two links. Among them, some work focuses on the design and extraction of feature descriptors, and some work emphasizes learning a specific distance or similarity model from training data to improve the accuracy of cross camera pedestrian identity matching. With the great success of deep learning in reinforcement learning, natural language processing and computer vision, many researchers also try to use deep learning method to achieve cross camera pedestrian identity matching. The feature representation in pedestrian recognition can be divided into two categories: low level visual features and middle level attributes with semantics. Generally speaking, the dimension of low-level visual features is relatively high but the expression ability is relatively weak; Although the middle-level attributes have the advantages of low dimension and strong expression ability, it is difficult to define, expand and detect. In the research of pedestrian recognition, the application of low-level visual features is more extensive than that of middle-level attributes. In order to capture the common features of the same pedestrian in the appearance of different cameras, and have a certain anti-interference ability to light, occlusion, angle of view and human body deformation, the commonly used low-level visual features in pedestrian recognition mainly include color, texture, shape, etc. Because these features can only capture one aspect of the image information, it is a very heavy and time-consuming work to collect and label pedestrian recognition data set.

Traditionally, the method of feature fusion and metric learning for pedestrians in public environment mainly adopts self-correlation quantization analysis method and matching filter fuzzy learning method to construct image fusion model reflecting pedestrian's 3D dynamic feature quantity in public environment, extract the set feature quantity of pedestrian re-recognition under public environment, establish 3D dynamic flow field for pedestrian feature fusion and metric learning under public environment, and obtain some research results. Zheng *et al.* introduced the posture as auxiliary information into the spatial transformation module of posture guidance [4], extracting the affine transformed pedestrian image from the partial image and the overall image and aligning it with the standard posture, and then using convolution. The neural network learns relevant features to realize partial pedestrian re-identification, but this method also

has the problems of high computational cost, low anti-interference ability, and easy tracking error. Liu and Wan proposed a pedestrian re-identification feature extraction method based on the attention mechanism [5], which used random erasure method to enhance the data of the input pedestrian image, by constructing a top-down attention mechanism network, and combining the attention. The mechanism network is embedded in the ResNet50 network to extract the salient features of the entire pedestrian, and the salient features of the entire pedestrian are similarly measured and sorted to obtain the result of pedestrian re-recognition. The boundary feature extraction effect of this method is poor. Pei *et al.* proposed a pedestrian re-recognition method that integrates viewpoint mechanism and pose estimation [6]. The pose estimation algorithm Openpose is used to locate the pedestrian joint points, and the pedestrian image is viewed to obtain viewpoint information, and the viewpoint information and the position of the pedestrian joint points are obtained. Local area recommendation, generate local images of pedestrians, input global images and local images into CNN to extract features at the same time, and use feature fusion network to fuse global and local features to obtain more robust feature representations. This method has a better contour segmentation effect. Difference. Shen *et al.* proposed multi-resolution feature attention fusion pedestrian re-recognition [7]. With the help of attention mechanism, based on the backbone network HRNet (high-resolution network), four different branches are constructed through interlaced convolution to extract multi-resolution Pedestrian image features not only extract different granular features of pedestrians, but also interact with different branch features to represent pedestrians. The extraction of feature points in this method is not detailed enough.

In view of the above problems, this paper proposes pedestrian re-identification in public environment based on feature fusion and metric learning. Firstly, the geometric mesh area model of pedestrian re-identification in public environment is constructed, the dynamic boundary feature points in pedestrian re-identification in public environment are reconstructed by fuzzy dynamic feature segmentation method, the gray pixel set under the dynamic constraint of pedestrian re-identification in public environment is extracted, the boundary feature quantity fusion is carried out according to the distribution intensity of gray pixel, the fusion and information enhancement of pedestrian re-identification image in public environment is realized. Finally, the simulation test shows the superior performance of this method in improving the feature fusion and measurement learning ability under the dynamic constraints of pedestrian re-identification in public environment.

II. BASIC DEFINITIONS

A. Geometrical Grid Area Model for Human Reidentification

Due to the low resolution of pedestrian images collected from different perspectives, the changes of illumination, posture, background and occlusion, how to extract the distinguishing pedestrian features is a key problem in pedestrian recognition research. In the process of pedestrian feature extraction, the features with stability and strong discrimination are often extracted from the appearance features of pedestrians. Because the global color histogram does not contain spatial information, in the process of extracting pedestrian features, in order to add spatial information, it is necessary to segment the pedestrian image into different local areas, and then extract the color histogram features in each local area. Different color histogram features can be extracted in different color spaces, including RGB, HSV, lab and so on. In these color histograms, HSV has strong robustness to the transformation of light, so HSV is used to extract the color features of pedestrian image.

Because the process of finding intersection point is to take every edge of one model to solve with all the edges of another model, when facing the situation that the intersection point is a vertex, we further improve and propose a new solution: For each edge of a model, in the process of traversing the intersection point anticlockwise, we only pay attention to its starting point and not its end point, similar to a left closed right open set Interval, if the intersection of the current edge and the edge to be intersected is the vertex of both sides at the same time, and the intersection is the starting point of both sides at the same time, then this intersection is considered as an effective intersection. We modify its flag bit to be both intersection and vertex. When the intersection is the vertex of both sides at the same time, but the starting point of one edge is the ending point of the other, we do not regard the intersection as an effective intersection. Note that when the intersection point is a vertex, it is not only necessary to judge whether the intersection point is an effective intersection point, but also not to generate a new node, just to modify the flag bit of the effective intersection point, and record that it is both a vertex and an intersection point, as shown in Fig. 1.
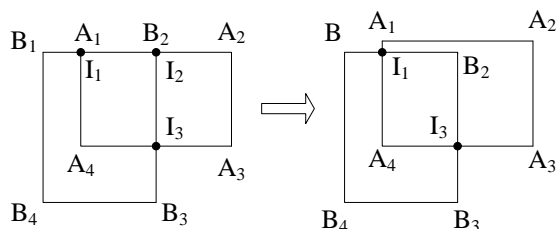

Fig. 1 Fretting diagram with overlapping edges

When meeting the node location, the current node should be recorded first, then the next node of the current node change node should be traversed, and then it should be regarded as the current node and traversed in turn. The original traversal rule is to directly record the change node of the node at the intersection position, and then traverse the next node of the change node. Although the two nodes are the same node, the path is indeed different. But although it seems that the idea of traversal is clear, there are still two choices of path when the intersection point is a fixed point. After the traversal rule is corrected, IA1 in chain A is also selected as the starting point of the composite. The new composite chain is as follows:
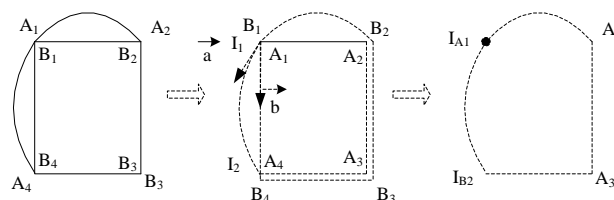

Fig. 2 New composite chain

1. Judge whether there is degradation between the two models. If there is degradation, go to step 2; Otherwise, go to step 3;

2. Jog one of the models by overlapping edges, and keep the original vector information, and turn to step 3;

3. The effective intersection points of the two models are obtained, and the intersection points are generated and inserted into the contour chains of the two models according to the insertion rules;

4. Find the starting point of the compound by the ray discrimination method;

5. Traverse the chain list of two profiles from the starting point according to the traversal rules to get the compound profile chain.

For profile A and profile B, four vertices can be used as the starting point of the composite, but A2 (B2) and A3 (B3) are not intersections. We choose one of these two points as the starting point of the composite. If A3 is taken as the starting point, A3 is the current node, not the intersection, record and add it to the compound contour chain; Then the current node pointer points to the next of A3, that is, a2a2 is not the intersection, record and add it to the compound contour chain; Next, the current node pointer points to the next of A3, that is, an A is the intersection, so vector judgment should be carried out: Two contours make a starting point at the vertex respectively as the direction of the current vertex Quantity: If the edge of the starting point in the contour vector is a straight line segment, the vector direction is the direction of the straight line segment; If the edge of the starting point in the contour vector is an arc segment, the vector direction is the tangent vector direction of the arc (in this paper, the counter clockwise direction is always used), as shown in Figs. 3-12. If the current node is a!, we use 5x1. If the current node is a surname, we use Sxl to record the current node and add it to the composite contour chain when the result of vector cross multiplication is

less than 0. If the result of vector cross multiplication is greater than 0, we record the change node of the current node and add it to the composite contour chain. In this way, the path selection when the intersection point is the vertex is selected.

In order to realize the feature fusion and metric learning under the dynamic constraints of pedestrian re-identification in public environment, it is necessary to first construct the geometric grid area model of pedestrian re-identification in public environment, and use the bottom-up modeling method to design the dynamic area grid model of pedestrian re-identification in public environment. The dynamic area grid model is designed as 3D grid area [8], the spatial distribution of grid area is V_2, and the geometric grid of pedestrian re-identification in public environment is obtained according to the uniform grid partition method.

$$V_2 = \{p(x,y) \mid x \in (0, width), y \in (0, height), x, y \in N\} \quad (1)$$

where, $width \times height$ is the domain of the definition of coordinates. In the grid composed of $width$ as row and $height$ as column, pedestrian re-identification and feature fusion and metric learning in the public environment are carried out. the grid model is shown in Fig. 3.
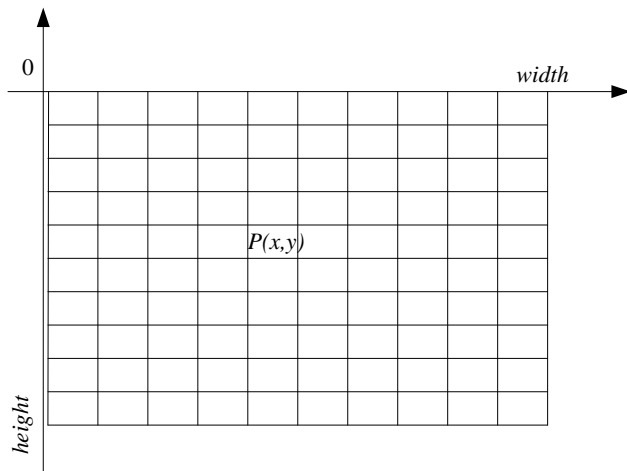

Fig. 3 Geometric mesh model for pedestrian re-identification in public environment

According to the geometric grid model of pedestrian re-identification under the common environment shown in Fig. 1, the shortest dynamic tracking distance from the boundary point G along the regional path u to v is:

$$pheromone(x, y) = (kind, capability) \quad (2)$$

In the boundary feature recognition and reconstruction of the grid area, according to the pheromone in the environment, $u^{(2)} = \left(u_1^{(2)}, u_2^{(2)}, u_3^{(2)}, u_4^{(2)}\right)$ as the pixel sequence, the quantitative tracking recognition is carried out, and the geometric feature quantity reflecting the dynamic attitude information of pedestrians in the public environment is

extracted. The formula $u = \left(y_0, z_0, \lambda, \phi\right)^T$ is used to calculate the moving rules and boundary features of pedestrians in the public environment, and the autocorrelation feature matching method is used to obtain the gray pixel set of the geometric feature parts identified by the pedestrian dynamic tracking in the public environment as $c = \left(\overline{X^{(3)}}, u_1^{(3)}, u_2^{(3)}\right)$. Using the pheromone rule to reconstruct the feature of pedestrian dynamic tracking in public environment, the coordinate value of c at any point on the geometric axis is obtained, the boundary characteristic quantity of pedestrian in moving public environment is X, and the vector quantization characteristic vector $\vec{a} = \left(\cos u_3^{(3)} \cos u_4^{(3)}, \sin u_3^{(3)} \cos u_4^{(3)}, \sin u_4^{(3)}\right)$ under the vector of variable shape base is obtained. According to the geometric mesh region model [9]-[11], the boundary block structure diagram of pedestrian recognition in public environment is shown in Fig. 4.
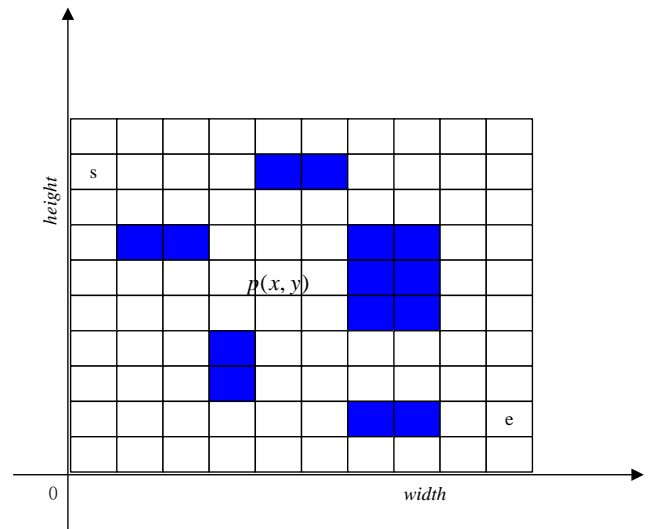

Fig. 4 Boundary block structure for pedestrian re-identification in public environment

B. Pedestrian Re-recognition Imaging and Fuzzy Dynamic Feature Segmentation in Public Environment

The fuzzy dynamic feature segmentation method is used to reconstruct the dynamic boundary feature points in pedestrian re-identification in public environment [12]. Under dynamic constraints, several boundary constraints of pedestrian dynamic tracking in public environment are as follows:

$$X_{RL} = R \times \theta_{RL} \quad (3)$$

$$X_{RR} = R \times \theta_{RR} \quad (4)$$

$$X_P = X_{RM} + L \sin \theta_P \quad (5)$$

$$\dot{X}_P = \dot{\theta}_P L \cos \theta_P + \dot{X}_{RM} \quad (6)$$

$$Y_P = L \cos \theta_P \quad (7)$$

$$\dot{Y}_P = -\dot{\theta}_P L \sin \theta_P \quad (8)$$

$$X_{RR} + X_{RL} = 2X_{RM} \qquad (9)$$

According to the characteristic of pedestrian movement in public environment [13], the boundary layer fusion and 3D dynamic information reconstruction of pedestrian in public environment are carried out in 3D moving space V_2, and the image output of pedestrian recognition in public environment is as follows:

$$I(i, j) = \sum_{k=1}^{P} I_{(k)}(i, j) \times 2^{k-1} \qquad (10)$$

Among them, $I_{(k)}$ is a block-matched pixel set in the 3D boundary grid region, using the similarity feature analysis method to obtain the gray-scale histogram of pedestrian 3D image in public environment:

$$x_i(t) = \sum_{k=1}^{p} \sum_{l=0}^{2} \varphi_{kl}[w_{i1}^l, \cdots, w_{in}^l][x_1(t-k), \cdots, x_n(t-k)]^T$$
$$-\sum_{k=1}^{q} \sum_{l=0}^{2} \theta_{kl}[w_{i1}^l, \cdots, w_{in}^l] \qquad (11)$$

where, φ_{kl} is the phase offset for multiple key point fusions. Using Harris corner detection, pedestrian recognition and reconstruction of boundary action feature points in public environment, the reconstruction results are shown in Fig. 3. According to the geometric invariance of each reconstructed feature point, feature fusion and metric learning are carried out [14], [15], and the geometric invariance moment of output is obtained as follows:

$$\left.\begin{aligned} EX^{(cs2)} &= \{x \mid x \in [0, h]\} \\ EY^{(cs2)} &= \rho^e \cos\theta^e \\ EZ^{(cs2)} &= \rho^e \sin\theta^e \end{aligned}\right\} \qquad (12)$$

Using the fuzzy dynamic feature segmentation method to reconstruct the dynamic boundary feature points in pedestrian re-identification in the public environment, the gray pixel set under the dynamic constraint of pedestrian re-recognition in the public environment is extracted [16], and the $\rho^e - R$ rotation invariant moment is used for vector translation in the direction of the x axis, thus realizing the pedestrian re-recognition imaging and the fuzzy dynamic feature segmentation in the public environment, as shown in Fig. 5.
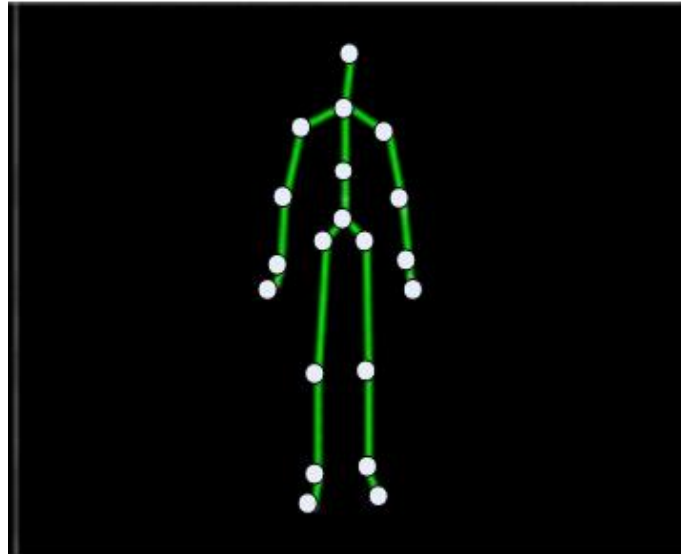

Fig. 5 Boundary feature distribution of pedestrian re-identification imaging in public environment

III. Feature Fusion And Measurement Learning Optimization For Pedestrian Reidentification In Public Environment

A. Reconstruction of Dynamic Boundary Features in Pedestrian Re-Identification

Based on the geometric mesh region model of pedestrian re-identification in public environment, the feature fusion and metric learning optimization design of pedestrian re-identification in public environment is carried out. In each belt area of pedestrian, each color channel of RGB is divided into 16 bin, 16 dimensional histogram is extracted respectively, and RGB color feature vector of 16 * 3 = 48 dimension is formed to represent RGB color feature, which is then normalized by L1 for feature fusion, as shown in Fig. 6.
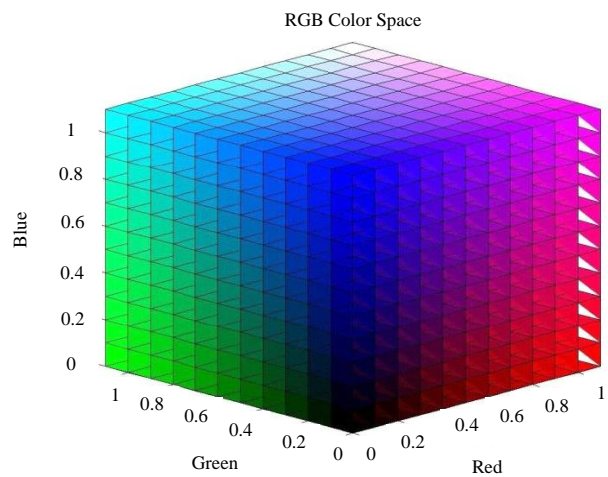

Fig. 6 RGB color space

This paper proposes pedestrian re-identification in public

environment based on feature fusion and metric learning. The fuzzy dynamic feature segmentation method is used to reconstruct the dynamic boundary feature points in pedestrian re-identification in public environment, as shown in Fig. 7.
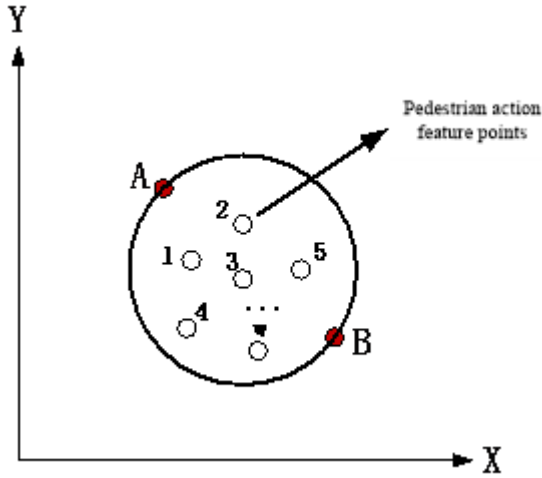

Fig. 7 Dynamic boundary feature reconstruction in pedestrian re-identification in public environment

The image fusion method is used to obtain the dynamic constraint boundary points of pedestrian recognition in public environment. Generate fusion feature f, the formula of feature fusion is as follows:

$$f(x,y) = vec(q(x,y) \otimes r(x,y))$$
$$= [q_1 r]^T [q_2 r]^T \cdots [q_n r]^T \qquad (13)$$

In the above formula, q, r is the characteristic quantity. In the spatial coordinate system, the dynamic feature data set of geometric dynamic constraint of pedestrian recognition in public environment:

$$(\theta^e, \rho^e) = EFA(\theta^*, \rho^*) \qquad (14)$$

Using h as the edge pixel set of pedestrian re-recognition in public environment, using 3D dynamic area planning method, using pixel point (θ^e, ρ^e) as the quantization center, using fuzzy dynamic feature segmentation method to reconstruct the dynamic boundary feature point of pedestrian re-identification in public environment, extracting the gray pixel set under the dynamic constraint of pedestrian re-identification in public environment:

$$\sigma(Z; D_X) = \sum_{i>j} \left| d_{ij}(Z) - d_X(x_i, x_j) \right|^2 \qquad (15)$$

In the above formula, the $d_{ij}(Z)$ is the Euclidean distance of the pixel point, and the $d_X(x_i, x_j)$ is the combined spatial distribution distance of the multiple key points [17], [18]. By using the gray-scale quantization method, the reconstruction area of the pedestrian dynamic tracking point is divided into two subblocks of $M \times N$. Finally, the reconstruction output of the pedestrian dynamic tracking point is:

$$G_{m,n} = \begin{pmatrix} g_{(m,n)}(1,1) & g_{(m,n)}(1,2) \\ g_{(m,n)}(2,1) & g_{(m,n)}(2,2) \end{pmatrix} \qquad (16)$$

$$m = 1,2,\ldots,M; n = 1,2,\ldots,N;$$

wherein,

$$g_{(m,n)}(u,v) = I_{(k)g}[2(m-1)+u, 2(n-1)+v] \qquad (17)$$

where, $u \in \{1,2\}; v \in \{1,2\}$ represents the global learning operator for pedestrian re-identification in a public environment, based on the reconstruction results for feature fusion and metric learning.

B. Local Motion Planning and Fuzzy Judgment Output for Pedestrian Re-identification in Public Environment

According to the distribution intensity of gray pixels, the boundary feature fusion is carried out, and the fusion output of the downlink image of the public environment is recorded as:

$$P_1 = \sum_{k=1}^{h} p_{(k)g}(i,j) \times 2^{k-1} \qquad (18)$$

$$P_2 = \sum_{k=1}^{h} p_{(k)g}^{*}(i,j) \times 2^{k-1} \qquad (19)$$

The 3D dynamic constraint method is used to realize the local motion planning of pedestrian recognition in public environment. The characteristic point distribution set of motion planning is as follows:

$$H(z) = P_1 \cdot \sum_{k=1}^{h} p_{(k)g}(i,j) \times 2^{k-1} / P_2 \cdot \sum_{k=1}^{h} p_{(k)g}^{*}(i,j) \times 2^{k-1} \qquad (20)$$

The Euclidean distance between the gray pixel set distribution feature points is calculated. For k adjacent points, the tracking quantization function of motion action points is as follows:

$$x_i(t) = [w_{i1}^{lk}, \cdots, w_{in}^{lk}][x_1(t-k), \cdots, x_n(t-k)]^T \qquad (21)$$

The Euclidean distance between the gray pixel set distribution feature points is calculated. For $[w_{i1}^{lk}, \cdots, w_{in}^{lk}]$ adjacent points, the tracking quantization function of motion action points is as follows:

$$SSIM(x,y) = [l(x,y)]^{\alpha} \cdot [c(x,y)]^{\beta} \cdot [s(x,y)]^{\gamma} \qquad (22)$$

The boundary geometric feature and sample set of pedestrian recognition in public environment can be obtained by using the method of dynamic pedestrian dynamic tracking point reconstruction. The overall flow chart is as follows:

```
          ┌─────────────┐
          │    Start    │
          └──────┬──────┘
                 │
        ┌────────▼─────────┐
        │ Constructing     │
        │ geometric mesh   │
        │ region           │
        └────────┬─────────┘
                 │  Image imaging processing
        ┌────────▼─────────┐
        │ Imaging and Fuzzy│
        │ dynamic feature  │
        │ Segmentation     │
        └────────┬─────────┘
                 │  Boundary feature Fuzzy judgement
                 │  Optimization
        ┌────────▼─────────┐
        │ Dynamic boundary │
        │ feature point    │
        │ reconstruction   │
        └────────┬─────────┘
                 │
        ┌────────▼─────────┐
        │ Local motion     │
        │ planning and     │
        │ fuzzy judgment   │
        │ output           │
        └────────┬─────────┘
                 │
          ┌──────▼──────┐
          │     End     │
          └─────────────┘
```
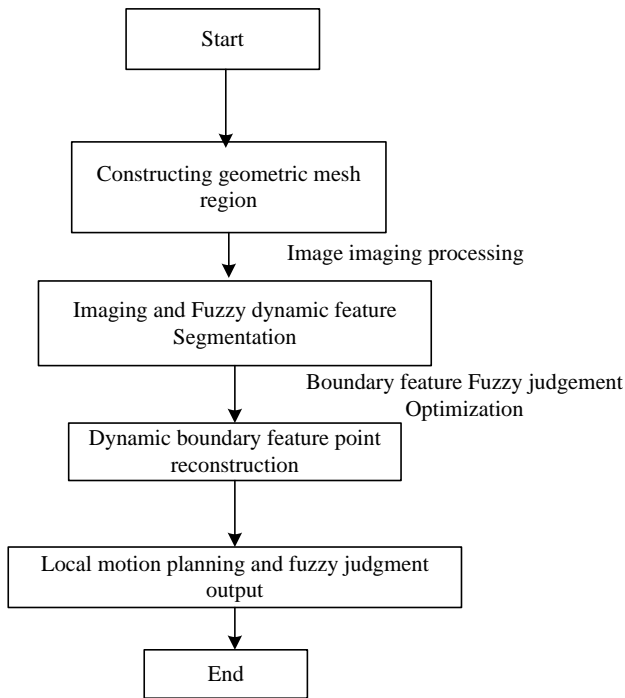
Fig. 8 Overall flow chart of pedestrian dynamic tracking in public environment

According to Fig. 8, the overall process of pedestrian dynamic tracking in public environment is divided into four steps. The first step is to construct a collection grid area. After image imaging processing, the second step is imaging and fuzzy dynamic feature segmentation, and boundary feature fuzzy judgment optimization is performed. Enter the third step of dynamic boundary feature point reconstruction, and finally the fourth step outputs local motion planning and fuzzy judgment.

IV. SIMULATION EXPERIMENT AND ANALYSIS OF RESULTS

A. Result Analysis

In order to test the application performance of the method in feature fusion and metric learning under the dynamic constraint of pedestrian re-recognition in public environment, a simulation experiment is carried out. The experiment is designed by Matlab, which takes two groups of pedestrians (MAN and WOMAN) in public environment as test objects, and sets the target template matching coefficient of pedestrian dynamic tracking in public environment to be 0.23, the histogram to M_s =12, and the template size to be 12 pairs of frames. The intensity of pixel level is 24dB, the center position is (0.25, 0.32), 10 frames of images are continuously read in, feature fusion and metric learning under dynamic constraints in pedestrian re-recognition in public environment are carried out, boundary feature quantity fusion is carried out according to the distribution intensity of gray pixels, fusion and information enhancement processing of pedestrian re-recognition images in public environment are realized, local motion planning of pedestrian re-recognition in public environment is realized by

adopting a three-dimensional dynamic constraint method, using this method and method [4]-[6], the results of feature fusion and metric learning under dynamic constraints are shown in Figs. 9 and 10. In the figure, from bottom to top are the algorithm of this paper, the method of [4]-[6].

Fig. 9 Feature fusion and metric learning tracking results of pedestrian recognition sequence MAN in public environment

Fig. 10 Feature fusion and metric learning tracking results of pedestrian recognition sequence WOMAN in public environment

The analysis results of Fig. 9 and Fig. 10 show that the pedestrians in the method in this paper are in the identification frame, and the top, bottom, left and right grids are not exceeded, and the features are accurately identified. The boundary of the method [4], [5], the recognition is poor, and all parts are outside the recognition, that is, the characteristics are not fully recognized. The range of the method is too large [6]. Therefore, the method in this paper is more accurate for the boundary feature judgment under the dynamic constraint of the pedestrian re-recognition in the public environment. Tracking recognition performance is better. Use different methods to

analyze the tracking error and test the center position error of pedestrian targets in the public environment. The test results are shown in Fig. 11.
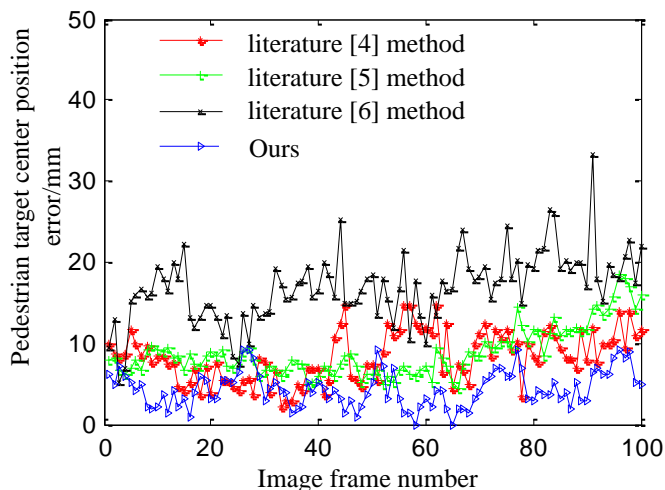

Fig. 11 Test results of pedestrian center position error in public environment

The analysis Fig. 11 shows that the method in this paper carries out the dynamic tracking and boundary characteristic judgment of pedestrians in the public environment, and the error is low and the robust performance is better.

Each part of the DPM pedestrian detector is regarded as each core. Based on the spatial weighted color histogram and the hog feature, the average value of each core is shifted iteratively. The deformation penalty provided by DPM pedestrian detector is used to constrain the movement of each part of the core. The aggregation of multiple parts of the core is used to determine the latest position of the tracking target. In order to estimate the impact of the maximum number of iterations on the tracking performance, as shown in Fig. 12.
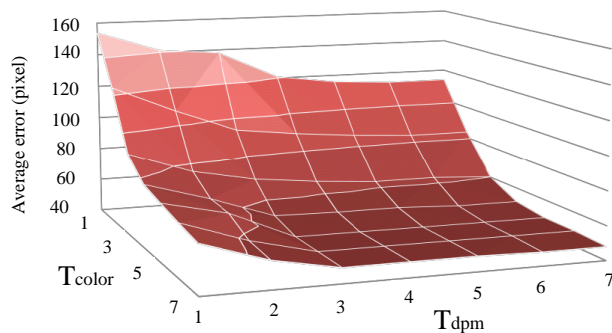

Fig. 12 Latest position side test results of tracking target

Taking full advantage of the rapidity of multi-core tracking and the robustness of DPM detector, the scheme is tested on some recorded video sequences (including ethms standard data set and videos recorded in this paper) and real-time mobile platforms (mobile robot Patsy and AR drone 2.0). The method in this paper can effectively aggregate multiple parts of the core, and then determine the latest location of the tracking target.

B. Discussion

According to the above experiments, it is obvious that the pedestrian re-identification method based on feature fusion and metric learning proposed in this paper has a very high degree of feature fusion of male and female pedestrians. No matter how the pedestrian moves, it can effectively follow the pedestrian and always keep the pedestrian in the identification frame. Inside. This is an intuitive display of tracking effects. On this basis, the pedestrian center position error and pixel error analysis are carried out to further determine the accuracy of the method in this paper. In terms of the center position error of the pedestrian target, the maximum error value of the method in this paper is 10 mm, the maximum error is 16 mm in [4], the minimum error is 3 mm, the maximum error is 21 mm in [5], and the minimum error is 6 mm. With the increase of the number of image frames, the error gradually increases and shows an upward trend. The maximum error is 35 mm in [6]. The minimum error is 6 mm, which has large fluctuations and is not stable enough. Through digital comparison, it can be seen that the error range of this method is between 0 and 10 mm, and the fluctuation range is small. At the same time, the maximum error is far lower than the error of other methods, even the error of is less than one-third of other methods, which further confirms that the target tracking effect of this method is better, and the error is reduced. In terms of pixels, the system has stabilized within 3 minutes of the method in this paper, and the average error of 7 colors is only 40 pixels, determine the latest location of the tracking target. Therefore, in general, the method in this paper can extract features better and effectively reduce tracking. Errors improve the recognition accuracy and enhance the recognition effect.

V. CONCLUSION

Fuzzy and accurate estimation of pedestrian boundary features in public environment under dynamic constraints is the basis for improving pedestrian recognition accuracy in public environment. Because of the large number of feature points in the public environment, the 3D dynamic area planning method can effectively shorten the reconstruction time by using the pixel point as the quantization center. In order to improve the accuracy of image reconstruction, the pedestrian movement area in 3D dynamic public environment can be further refined under the constraint of simplified algorithm. This paper proposes pedestrian re-identification in a public environment based on feature fusion and metric learning. Research shows that this method has better accuracy, better tracking performance and lower tracking error. In the experiment process, the maximum error obtained is only 10

mm. In some of the image frames, it can achieve zero error, and it has good performance. Application value. However, some feature representation methods still have some shortcomings, which are worthy of further study:

First, in the design of the underlying visual features, although a variety of different feature representation methods have been proposed, and also brought about the improvement of RELD performance, there is still a lack of comprehensive and detailed analysis and evaluation of the design and extraction process of the underlying visual features. We find that most of the common underlying visual features in RELD have similar properties or extraction processes (most of them are statistical features, such as color histogram, hog, LBP, covariance features, etc.), so we can use a unified extraction process to extract features, so we can also use a unified framework to reasonably evaluate the details of feature extraction, To help researchers better understand the role of each feature design link and inspire us to design better feature representation.

Secondly, the current mainstream feature representation methods, including low-level visual feature design and high-level semantic feature learning, use a single feature vector to represent the overall appearance of pedestrians. We believe that this single vector representation method is easy to lead to the loss of detail information, and appearance details are often the key to achieve accurate pedestrian matching; Therefore, we need to design feature representation methods that can better retain key information in pedestrian images or videos, such as feature representation based on feature sets or feature sequences.

Thirdly, due to the complexity of the change of human appearance in the monitoring scene, most feature representation methods adopt the way of multi feature fusion to describe the pedestrian appearance; However, most of the commonly used feature fusion methods are based on heuristic fusion strategies, such as directly connecting different features in series, which cannot guarantee the optimal combination of features. Therefore, it is necessary to design a data-driven and adaptive feature fusion strategy to further improve the effectiveness of feature representation.

References

[1] H. He, and Y. Tan, "Automatic pattern recognition of ECG signals using entropy-based adaptive dimensionality reduction and clustering," Applied Soft Computing, vol. 55, pp. 238-252, 2017.

[2] M. Ye, Y. Qian, and J. Zhou, "Multitask sparse nonnegative matrix factorization for joint spectral-spatil hyperspectral imagery denoising," IEEE Transactions on Geoscience and Remote Sensing, vol. 53, no. 5, pp. 2621-2639, 2015.

[3] L. R. Hu, J. G. Wu, and L. Wang, "Application and method for linear projective non-negative matrix factorization," Computer Science, vol. 40, no. 10, pp. 269-273, 2013.

[4] Y. Zheng, J. Y. Zhao, Z. Wang, and Y. Zhang, "Partial pedestrain re-identification based on pose-guided alignment network," Computer Engineering, vol. 46, no. 5, pp. 253-259, 2020.

[5] Z. Y. Liu, and P. P. Wan, "Pedestrian re-identification feature extraction method based on attention mechanism," Journal of Computer Applications, vol. 40, no. 3, pp. 672-676, 2020.

[6] J. Z. Pei, Z. C. Xu, and P. Hu, "Person re-identification fusing viewpoint mechanism and pose estimation," Computer Science, vol. 47, no. 6, pp. 170-175, 2020.

[7] Q. Shen, C. Tian, J. B. Wang, S. S. Jiao, and L. Du, "Multi-resolution feature attention fusion method for person re-identification," Journal of Image and Graphics, vol. 25, no. 5, pp. 946-955, 2020.

[8] W. Zhao, X. J. Huang, and R. Y. You, "Removal of white noise from ECG signal based on morphological component analysis," Chinese Journal of Biomedical Engineering, vol. 23, no. 1, pp. 1-6, 2014.

[9] S. N. Ke, J. Gong, S. N. Li, Q. Zhu, X. T. Liu, and Y. T. Zhang, "A hybrid spatio-temporal data indexing method for trajectory databases," Sensors, vol. 14, no. 7, pp. 12990-13005, 2014.

[10] X. T. Ma, J. Q. Luo, and S. L. Wu, "Joint sorting and location method using TDOA and multi-parameter of multi-station," Journal of National University of Defense Technology, vol. 37, no. 6, pp. 78-83, 2015.

[11] X. W. Nie, "Radar signal pre-sorting based on K-means algorithm," Electronic Science and Technology, vol. 26, no. 11, pp. 55-58, 2013.

[12] C. H. Ju, and J. B. Zou, "An incremental classification algorithm for data stream based on information entropy diversity measure," Telecommunications Science, vol. 31, no. 2, pp. 86-96, 2015.

[13] Y. X. Lyu, C. Y. Wang, C. Wang, and C. Y. Yu, "Online classification algorithm for uncertain data stream in big data," Journal of Northeastern University (Natural Science Edition), vol. 37, no. 9, pp. 1245-1249, 2016.

[14] B. Sun, J. D. Wang, H. Y. Chen, and Y. T. Wang, "Diversity measures inensemble learning," Control and Decision, vol. 29, no. 3, pp. 385-395, 2014.

[15] A. Metwally, and C. Faloutsos, "V-SMART-Join, a scalable MapReduce framework for all-pair similarity joins of multisets and vectors," Proceedings of the VLDB Endowment, vol. 5, no. 8, pp. 704-715, 2012.

[16] S. Hashemi, and Y. Yang, "Flexible decision tree for data stream classification in the presence of concept change,
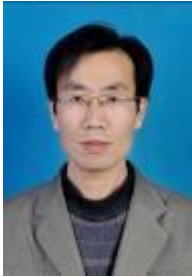
noise and missing values," Data Mining and Knowledge Discovery, vol. 19, no. 1, pp. 95-131, 2009.

[17] A. F. Oskooi, D. Roundy, M. Ibanescu, P. Bermel, J. D. Joannopoulos, and S. G. Johnson, "Meep: A flexible free-software package for electromagnetic simulations by the FDTD method," Computer Physics Communications, vol. 181, no. 3, pp. 687-702, 2010.

[18] W. Li, and X. Y. Li, "Simulation of data cluster recognition of abnormal trajectory for crowd monitoring image," Computer Simulation, vol. 36, no. 2, pp. 390-394, 2019.

Hongbin Luo, born in September 1979, male, Han nationality, with Master's degree, is associate professor. He received his Bachelo's degree in computer Science and Technology from Northwest Normal University in 2003, and received his Master's degree in Curriculum and Pedagogy from Northwest Normal University in 2010. He is working in Gansu University of Political Science and Law. His research interests include computer vision and deep learning. He has published more than 10 academic papers and presided over and participated in several provincial scientific research projects, such as "Research on Pedestrian Re-recognition Technology based on Deep Learning in Surveillance Video" and "Research on Intelligent Video Surveillance Technology".

Author Contributions

Hongbin Luo used the three-dimensional dynamic constraint method to realize the local motion planning of pedestrian recognition in public environment. According to the results of contour segmentation, the feature fusion and learning of pedestrian recognition in public environment are realized.

Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)