

Research on the Pedestrian Re-Identification Method Based on Local Features and Gait Energy Images

Xinliang Tang¹, Xing Sun¹, Zhenzhou Wang¹, Pingping Yu¹, Ning Cao^{2,*} and Yunfeng Xu³

Abstract: The appearance of pedestrians can vary greatly from image to image, and different pedestrians may look similar in a given image. Such similarities and variabilities in the appearance and clothing of individuals make the task of pedestrian re-identification very challenging. Here, a pedestrian re-identification method based on the fusion of local features and gait energy image (GEI) features is proposed. In this method, the human body is divided into four regions according to joint points. The color and texture of each region of the human body are extracted as local features, and GEI features of the pedestrian gait are also obtained. These features are then fused with the local and GEI features of the person. Independent distance measure learning using the cross-view quadratic discriminant analysis (XQDA) method is used to obtain the similarity of the metric function of the image pairs, and the final similarity is acquired by weight matching. Evaluation of experimental results by cumulative matching characteristic (CMC) curves reveals that, after fusion of local and GEI features, the pedestrian re-identification effect is improved compared with existing methods and is notably better than the recognition rate of pedestrian re-identification with a single feature.

Keywords: Local features, gait energy image, weight, independent distance metric, cross-view quadratic discriminant analysis.

1 Introduction

Pedestrian re-identification is a research hotspot in the field of intelligent video image analysis. The objective of this task is to identify cross-camera pedestrian targets in non-overlapping monitoring networks [Liu, Tao, Song et al. (2015); Li, Zhou, Yu et al. (2019); Zhou, Ke and Luo (2019)]. Pedestrian re-identification technology is widely used in intelligent video surveillance, criminal investigations, home security management and other fields. In the video surveillance environment, pedestrian re-identification mainly depends on pedestrian appearance and clothing, and pedestrians with similar clothing or

¹ Hebei University of Science and Technology, Shijiazhuang, 050000, China.

² School of Internet of Things and Software Technology, Wuxi Vocational College of Science and Technology, Wuxi, 214028, China.

³ Rutgers Business School-Newark, Washington Park, Newark, NJ 07102, USA.

* Corresponding Author: Ning Cao. Email: ning.cao2008@hotmail.com.

Received: 24 February 2020; Accepted: 18 April 2020.

who have changed clothing can easily affect the identification results. Therefore, pedestrian re-identification has become a challenging and valuable topic in the field of computer vision.

Pedestrian re-identification involves extracting the features of a person, calculating the distance between each feature vector with a distance measurement function, and then describing the similarity between persons by the distance between features to achieve pedestrian re-identification. Cheng et al. [Cheng, Cristani and Stoppa (2011)] used a pictorial structure to divide the human body into different parts and extract features from each part. Hervé et al. [Hervé and Ondřej (2012)] estimated the salient region of each image, learned the similarity function of the salient regions, and sorted the results. Prosser et al. [Prosser, Zheng, Gong et al. (2010)] used a support vector machine to measure the similarity of predefined color and texture feature vectors. Liao et al. [Liao, Hu, Zhu et al. (2015)] established a feature expression method using color and texture histograms and proposed a locally monotonic (LOMO) model for calculating local color texture features and obtaining an ultrahigh dimensional feature vector. Matsukawa et al. [Matsukawa, Okabe, Suzuki et al. (2016)] proposed a Gaussian of Gaussian (GOG) descriptor that adopted a hierarchical Gaussian operator to divide the image into different regions described by multiple Gaussian distributions representing color and texture information. Each Gaussian distribution represented a small image block, and the features of each image block were combined to produce the feature vector of the pedestrian image. Wang et al. [Wang, Gong, Zhu et al. (2014)] first extracted the motion features of a pedestrian by calculating the gait energy image (GEI) [Hofmann, Bachmann and Rigoll (2012)] of each pedestrian video sequence and then fused HOG3D spatiotemporal features before finally conducting similarity measurements with discriminating video frame ranking models. Deng et al. [Deng, Wang, Cheng et al. (2017)] established a comprehensive feature of human gait dynamics by combining the spatiotemporal and kinematic gait characteristics of the human body. When the walking conditions of the human body change, this method can still achieve excellent performance. However, this method requires a high-quality gait sequence, the calculation amount for establishing the model is large, and it is difficult to achieve the ideal effect. Mogan et al. [Mogan, Lee and Tan (2017)] proposed a time gradient pattern method for gait recognition that not only described the spatial contour shape but also implicitly captured contour deformation over time. Based on the distribution of feature space and the relationship between sample points, Zheng et al. [Zheng, Zheng and Yang (2018)] proposed an algorithm combining verification loss and classification loss, but the algorithm did not consider image alignment and local features of pedestrian images. Zhao et al. [Zhao, Tian and Sun (2017)] used key points of the human body as prior knowledge to divide it into several fixed rigid structures, thereby strengthening local feature learning; however, the algorithm did not consider the integrity of the images of individuals and background features, resulting in low recognition accuracy. Sun et al. [Sun, Liang, Yang et al. (2017)] proposed a method for dividing the feature image horizontally and fusing multi-granularity features to complete the task of pedestrian re-identification; however, the approach did not consider pedestrian image misalignment.

The rest of this article is organized as follows. In Section 2, color and texture features are extracted as local features, while gait features are extracted in Section 3. In Section 4,

local and gait features are fused, and similarity measurement functions are learned. In Section 5, experiments are carried out to verify the effectiveness of the proposed algorithm. Finally, the last section concludes the paper.

2 Local feature extraction

In this paper, it is assumed that a pedestrian is composed of various parts such that a convolution structure is adopted to locate pedestrian body nodes [Syed, Tahir, Robina et al. (2019)]; the body structure is then divided according to the node positions, while the images are aligned. Then, the extracted nodes divide the pedestrian into four parts: the head, the trunk and arms, and the left and right legs. Fig.1 shows the body area division diagram. Each part is scaled to a fixed pixel value, and the pixel size of each part is unified; the image of each part can be described as $P = \{P_1, P_2, P_3, P_4\}$. Finally, the color and texture features of each part are extracted.

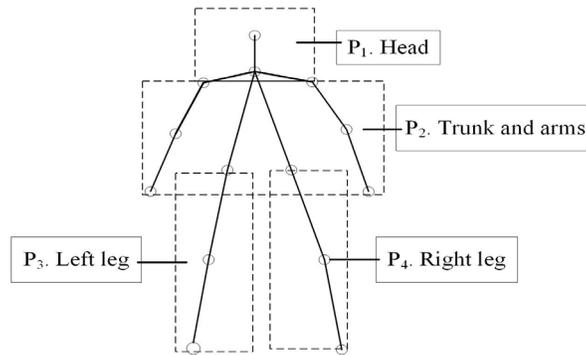


Figure 1: Body area division diagram

2.1 Color feature extraction

Hue-Saturation-Value (HSV), Lightness-A-B (LAB) and Luminance-Chroma: Blue-Chroma: Red (YCbCr) are three complementary color spaces; therefore, describing a given object from different angles can better reflect the differences between samples. This paper extracts features in the HSV, LAB and YCbCr color spaces. Due to illumination, clothing material and other factors, the color of images can vary greatly for the same pedestrian between different cameras and for different pedestrians with the same camera, which affects the extracted color [Wang, Huo, Yu et al. (2019)]. Therefore, before extracting the color feature histogram, the image is first processed with the retinex image enhancement algorithm [Nejad and Shiri (2019)] to reduce the influence of illumination on color [Yang, Yang, Yan et al. (2014)].

According to the image division standard shown in Fig. 1, a pedestrian image is divided into 4 area blocks. For each channel of each color space of each area block, a 16-dimensional color feature histogram is extracted, and all the color feature histograms are then connected in series to generate the color features of the pedestrian image. A 576-dimensional ($4 \times 3 \times 3 \times 16$) color feature can be extracted for each pedestrian image. If the original 576-dimensional features are used to directly learn the similarity metric function,

2×576×576 parameters must be estimated, which is too much. Therefore, in the experiment, principal component analysis (PCA) is used to reduce the feature dimension of the pedestrian image to a specific dimension.

2.2 Texture feature extraction

To better describe the local information of the pedestrian image, the scale invariant ternary pattern (SILTP) is used to extract the texture information of the pedestrian image. The SILTP [Liao, Zhao, Kellokumpu et al. (2010)] is an improved local binary pattern (LBP) description operator with a good anti-interference ability for noise in the region. When the detection region is dark, covered by shadows or contains a lot of noise, the operator has strong adaptability. Moreover, the SILTP operator exhibits scale invariance, which makes it robust to brightness changes; thus, the SILTP feature will be only slightly affected, even if the illumination suddenly changes from dark to bright. Fig. 2 shows the encoding process of the SILTP operator in this algorithm.

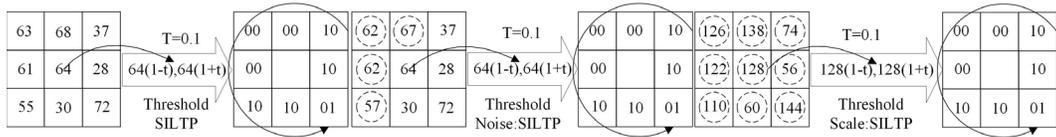


Figure 2: Encoding process of the SILTP operator

Assuming that the pixel position of an image is (x_c, y_c) , the SILTP encoding method is as follows:

$$SILTP_{Q,R}^t(x_c, y_c) = \bigoplus_{k=0}^{Q-1} s_t(I_c, I_k) \tag{1}$$

$$s_t(I_c, I_k) = \begin{cases} 01 & I_k > (1+t)I_c \\ 10 & I_k < (1-t)I_c \\ 00 & \text{others} \end{cases} \tag{2}$$

where I_c is the grayscale value of the image center pixel point, I_k is the grayscale value of the pixel point corresponding to the Q neighborhood with radius R , \bigoplus connects the binary values of all neighborhoods into character strings, and t is the threshold range. As shown in Fig. 2, the image is binary coded according to Eqs. (1) and (2), and the center point is coded as 0010010100001001 in counterclockwise order. Fig. 2 reveals that, even if the image contains noise or scale changes within a certain range, the SILTP coding value remains unchanged, which indicates that the operator is robust to problems such as illumination changes.

According to the image division standard shown in Fig. 1, for each area block, the SILTP features in two modes with a threshold of 0.3, a neighborhood of 8 with a radius of 1, and a neighborhood of 16 with a radius of 2 are extracted. The dimensions in each mode are 81-dimensional. Then, four area block features are connected in series at the same time, and a 648-dimensional (4×2×81) texture feature can be extracted from each pedestrian image. The PCA algorithm reduces the feature to a specific dimension.

3 Gait feature extraction

The pedestrian re-identification results are affected to a certain extent by the appearance and clothing of pedestrians [Zhao, Ouyang and Wang (2013)]. Gait features have advantages in terms of long-distance non-invasiveness, difficulty in disguising and hiding, and feature extraction and recognition, which are not available in other biometrics, and can be used for low-resolution images. GEIs provide abundant static and dynamic information for gait recognition. The LBP has simple calculations and high accuracy, but it has incomplete features and low accuracy for complex images. HOG features can suitably describe the edge information of images. To fully consider the local deformation of the human target image and to better describe the boundary contour information of the human body, LBP and HOG features are extracted from the GEI to describe the GEI features and compare the similarities. Therefore, in this paper, the LBP and HOG operators are used to extract pedestrian GEI features as gait features.

3.1 Image preprocessing

In this paper, the background modeling method based on the Gaussian mixture model is used to model the background, the foreground is extracted by the background subtraction method, and the obtained binary image is morphologically processed. The image preprocessing flowchart is shown in Fig. 3.

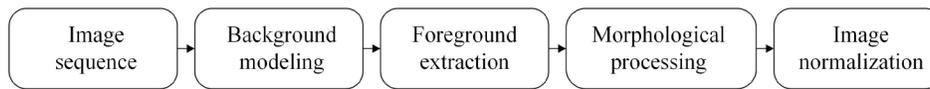


Figure 3: Image preprocessing flowchart

The pedestrian binary image obtained after preprocessing is shown in Fig. 4.



Figure 4: Image preprocessing

In gait images, the position of the human body region changes continuously during movement. Subsequent image feature extraction is facilitated by normalizing the image to a uniform size, which ensures that the normalized human body region is located at the center of the image and that human body alignment is realized in all images.

Normalization: The image is traversed from left to right and from top to bottom to identify the four boundaries $x_{min}, x_{max}, y_{min}, y_{max}$ of the human body region. In the subsequent analysis, the center point of the human body region will be more readily found according to $x = (x_{min} + x_{max}) / 2, y = (y_{min} + y_{max}) / 2$, and images are cropped to a size of 160×90 . The cropped images retain the complete aligned human body regions, while the pixel size of the image has been reduced.

3.2 Gait cycle extraction

Gait is considered a regular cyclic movement, and extracted gait information is based on several binary images of the human body during a given period. In this paper, the optical flow method [Chen, Zha, Li et al. (2018)] is adopted to calculate the gait cycle.

The optical flow method describes the instantaneous motion shape, including the shape of the moving object (space) and the shape of the motion (time). Different gaits can be distinguished from the periodic difference in the motion shape. In this method, calculations rely solely on the relative motion of adjacent frames without constructing any model or acquiring the image background in advance. The results provide not only motion information about the moving objects but also rich information about the three-dimensional structure of the scene.

In the optical flow method, it is reasonable to calculate the gait cycle by the periodic variation in offset ν in the whole gait sequence. Therefore, $T(t)$ and expressions of the ν component of the gait flow field t are used to calculate the gait cycle.

$$T(t) = \sum_{x=1}^W \sum_{y=1}^H f_{\nu}(x, y, t) \quad (3)$$

where x and y represent the serial numbers of the rows and columns, respectively, W and H represent the numbers of rows and columns, respectively, and $f_{\nu}(x, y, t)$ represents the optical flow component ν of the t and $t+1$ frames of the sequence at coordinate (x, y) . Fig. 5 shows the periodic variation of $T(t)$. The frames between every pair of adjacent peaks or valleys belong to the same period. The image in the frame where $T(t)$ is located is extracted as a gait cycle.

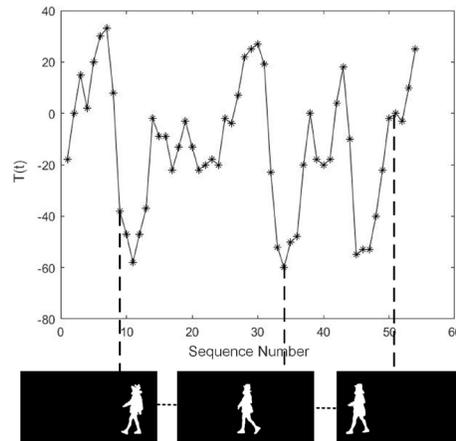


Figure 5: Periodic variation in $T(t)$

3.3 Gait energy image extraction

The GEI is a commonly used feature in gait detection; it is easy to extract and reveals gait features such as speed and morphology. It is defined as follows:

$$G(x, y) = \frac{1}{N} \sum_{t=1}^N I_t(x, y) \tag{4}$$

where N represents the length of an image sequence of a gait cycle, t is the number of frames, and $I_t(x, y)$ is the pixel value at pixel point (x, y) . GEIs are calculated by Eq. (4) for all images within a gait cycle. GEIs [Luo, Yang and Liu (2016)] of different people are shown in Fig. 6. The brighter the position in the image, the more pixels appear. In brighter areas, the changes during walking are smaller, and in darker areas, the changes are larger. If the GEI is directly analyzed as a feature, then the amount of data is very large; thus, PCA is performed on the GEI to reduce the dimensionality of the image.

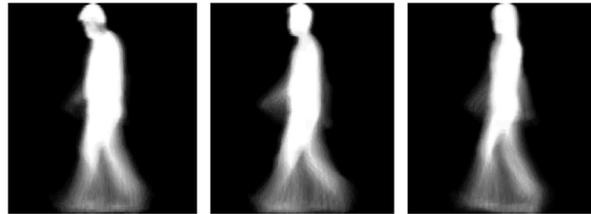


Figure 6: Gait energy images of different people

3.4 LBP feature extraction of gait energy images

Given a GEI, the LBP operator will form a series of LBP codes by calculating the values of all pixel points of the image; then, the image features that are represented by the feature histogram are extracted, namely:

$$H_i = \sum_{x,y} I(f(x, y) = i), i = 0, 1, 2, \dots, 2^N \tag{5}$$

where $I(X) = \begin{cases} 1 & f(x, y) = i \\ 0 & f(x, y) \neq i \end{cases}$.

In this paper, the LBP feature image is divided into equally sized and non-overlapping blocks. The LBP feature histograms of all blocks are calculated and serially connected to establish the histogram of LBP feature images, i.e., the LBP features.

The image and histogram of a GEI after LBP feature extraction are shown in Fig. 7.

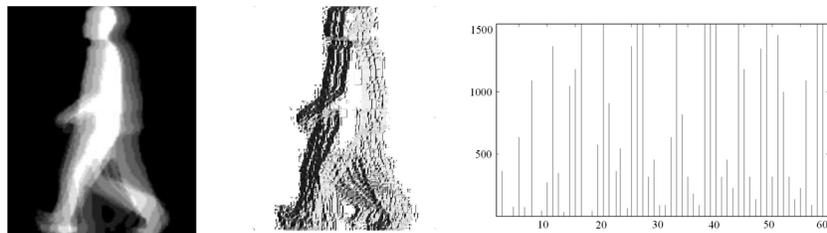


Figure 7: Histogram after LBP feature extraction

3.5 HOG feature extraction of gait energy images

The implementation process of the HOG feature extraction algorithm is as follows:

(1) Normalization of the color space. Normalization eliminates interference from uncontrollable elements such as the shooting environment, background, and device in the target image.

(2) Gradient calculation. Gradients in the horizontal and vertical directions of a pixel point (x, y) in the image are calculated as follows:

$$\begin{cases} G_x(x, y) = I(x+1, y) - I(x-1, y) \\ G_y(x, y) = I(x, y+1) - I(x, y-1) \end{cases} \quad (6)$$

Then, the gradient and direction of the obtained pixel point are as follows:

$$G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \quad (7)$$

$$\theta(x, y) = \arctan\left(\frac{G_x(x, y)}{G_y(x, y)}\right) \quad (8)$$

(3) Histogram of gradient direction. The image is divided into many cells (in this paper, $8 \times 8 = 64$ pixels are used as one cell, and there is no overlap between adjacent cells), and all the gradients in each cell are counted, while all the gradient directions are sorted into 9 bins. The gradient amplitudes of each bin are aggregated, and the histogram is finally normalized to obtain the characteristics of each cell.

(4) Normalization of overlapping block histograms. The HOG features of the whole image can be obtained by grouping several cells into one block (in this paper, every 4 cells form a block, thus allowing overlap between blocks), and the feature histograms of all cells in a block are connected in series to form the block feature histogram, while all the block feature histograms are connected in sequence.

The HOG feature image and GEI feature histogram of an image are shown in Fig. 8.

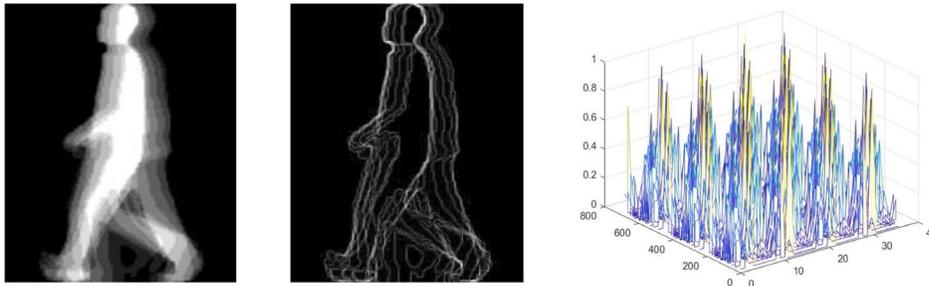


Figure 8: Extracted HOG features from the GEI

4 Feature fusion and similarity measurement function

The cross-view quadratic discriminant analysis (XQDA) algorithm [Liao, Hu, Zhu et al. (2015)] is an excellent metric learning algorithm. Compared with similar algorithms, this algorithm has a high and robust matching rate and a short calculation time. Therefore, in this paper, XQDA is used to measure similarity.

The XQDA method simultaneously performs feature subspace and metric learning and considers the effect of dimension reduction on the metric learning results while reducing

the feature dimension. In XQDA, the original feature $x_i, x_j \in R^d$ is mapped to a low-dimensional subspace by learning the mapping matrix $M \in R^{d \times r}$ ($r < d$). The similarity function is defined in Eq. (9):

$$\delta(x_i, x_j) = (x_i - x_j)^T M (\Sigma_I^{-1} - \Sigma_E^{-1}) M^T (x_i - x_j) \quad (9)$$

where $\Sigma_I' = M^T \Sigma_I M, \Sigma_E' = M^T \Sigma_E M$, and Σ_I and Σ_E are the covariance matrices of the intraclass and interclass sample difference distributions, respectively. When solving and decomposing the characteristic value of matrix $\Sigma_I^{-1} \Sigma_E$, matrix M is composed of eigenvectors corresponding to the first r maximum characteristic values, with $W = \Sigma_I^{-1} - \Sigma_E^{-1}$, and W is the measurement matrix.

In this paper, the color, texture, LBP and HOG features of the GEI of a pedestrian are extracted. To fully reflect the differences in feature expression in the different spaces and to fully utilize the complementary effect of the different features, the color, texture, LBP and HOG features of the GEI are separately extracted to learn the measurement matrix W_1, W_2, \dots, W_4 with the independent measurement learning method. The measurement criteria based on the different features are then obtained and used to weight and fuse the similarity of test samples in different spaces to obtain the final similarity.

Four measurement matrices W_A, W_B, W_C, W_D can be acquired by learning the color and texture features of local region blocks [Farenzena, Bazzani, Perina et al. (2010)] and the LBP and HOG features of GEIs. The similarities $\delta_A, \delta_B, \delta_C, \delta_D$ of the four features can be obtained by Eq. (10). The similarity of the local color and texture features is weighted to obtain the similarity δ_{local} of the local features, as shown in Eq. (10).

$$\delta_{local}(x_i, x_j) = \gamma \delta_A(x_i, x_j) + (1 - \gamma) \delta_B(x_i, x_j) \quad (10)$$

The LBP and HOG features of the GEI are weighted for similarity to obtain the similarity δ_{gei} of the gait, as shown in Eq. (11).

$$\delta_{gei}(x_i, x_j) = \eta \delta_C(x_i, x_j) + (1 - \eta) \delta_D(x_i, x_j) \quad (11)$$

Here, $\delta_n(x_i, x_j) = (x_i - x_j)^T M_n (\Sigma_I^{-1} - \Sigma_E^{-1}) M_n^T (x_i - x_j)$ $n = A, B, C, D$.

The local and pedestrian gait features are weighted and fused to establish the final similarity of the pedestrian, as shown in Eq. (12).

$$\delta(x_i, x_j) = \lambda \delta_{local}(x_i, x_j) + (1 - \lambda) \delta_{gei}(x_i, x_j) \quad (12)$$

Here, λ, γ, η are the similarity weights. The experimental results show that the recognition rate of the algorithm is optimal when $\lambda, \gamma, \eta \in [0.3, 0.4]$. In the following experiments, the weight values λ, γ, η are uniformly set to 0.4.

5 Experiment and analysis

5.1 Experimental data and evaluation indicators

In this paper, Dataset B [Zheng, Zhang, Huang et al. (2011)] from the CASIA gait database of the Institute of Automation of the Chinese Academy of Sciences is used to

carry out the experiments. The dataset contains 120 pedestrian data samples of 100 pedestrians in the normal walking state recorded from a 0-degree visual angle; 10 samples of pedestrians with similar clothing and 10 samples of pedestrians with similar postures were added. To analyze the identification performance of this algorithm, 70 samples were selected for training, and 50 samples were selected for testing.

The most commonly used method for evaluating the pedestrian re-identification rate is the cumulative matching characteristic (CMC) curve. The calculation method is as follows:

$$CMC(l) = \frac{1}{N} \sum_{i=1}^N \prod (rank(P_i) < l) \quad (13)$$

where l is the ranking value of the cumulative accuracy, i.e., the probability that positive samples exist in the identification results with a ranking lower than or equal to l ; N is the number of samples for matching; \prod is a symbolic function and is equal to 1 if certain conditions are met or 0 otherwise; $rank(\cdot)$ is the sample distance ranking calculation; P_i is the positive sample distance of gallery sample i ; and $rank(P_i)$ is the sorting operation of the positive samples.

5.2 Performance comparison of the algorithms with different values of weight λ

Eq. (13) indicates that the similarity of pedestrian images is δ . To analyze the influence of different weights δ on algorithm performance, Fig. 9 shows the results of a comparative experiment of algorithm performance for weights λ of the dataset used in this experiment. When λ is between 0.3 and 0.4, the similarity between the color and texture features and the gait features is higher due to the gait feature compensation in the case of similar pedestrian clothing and the color and texture compensation in the case of similar postures; in this way, the performance of the algorithm is optimized. Tab. 1 summarizes the algorithm recognition results of the local and GEI features in this paper based on a 0-degree visual angle.

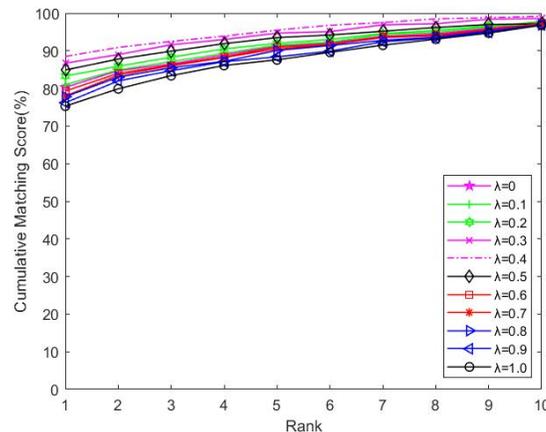


Figure 9: Performance comparison of the algorithms for different weights λ at a 0-degree visual angle

Table 1: Algorithm identification results of local and gait energy image features at a 0-degree visual angle

Features	Rank 1	Rank 5	Rank 10
Local	75.3%	88.6%	96.8%
Gait Energy Image	80.3%	90.2%	97.3%
Features in this paper	88.5%	93.5%	99.2%

5.3 Performance comparison of algorithms with different weights γ, η

Eq. (12) indicates that the similarity of pedestrian local features is weighted by the similarity of color and texture features, and the GEI features similarity is weighted by the LBP and HOG features similarities. Fig. 10 shows the performance comparison results for different weights γ, η of the dataset in this experiment. The experimental results reveal that the identification performance is optimal when γ, η is between 0.3 and 0.4.

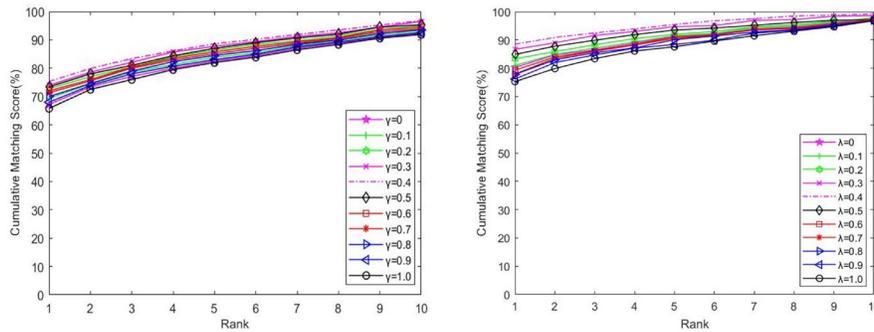


Figure 10: Performance comparison of the algorithms for different weights γ, η at a 0-degree visual angle

5.4 Result comparison

Tab. 2 summarizes the comparison of the experimental results and those from the gait energy image-region bounded by legs-local binary pattern (GEI-RBL-LBP) [Kumar and Nagendraswany (2014)] and gait energy image-principal component analysis-multiple discriminant analysis (GEI-PCA-MDA) [Rida, Almaadeed and Bouridane (2014)] pedestrian re-identification methods, with relevant parameters for 120 samples, including the 100 data samples from the CASIA Dataset B standard database and the 20 newly added data samples, which are set to optimal values. The experimental results reveal that the average identification effect of the proposed algorithm is the best compared with the other pedestrian re-identification methods.

Table 2: Rank-1 comparison of the results of the different methods

Methods	Rank 1
GEI-RBL-LBP	79.3%
GEI-PCA-MDA	83.3%
The method used in this paper	88.5%

In this paper, pedestrian samples with similar clothing and postures are included as interference samples, which increases the difficulty of identifying the dataset. In the proposed method, color, texture and gait features are fused; for pedestrians with similar clothing, the gait features of the pedestrians compensate for the defects of the color and texture features, and for pedestrians with similar postures, the color and texture features suitably compensate for the deficiencies of the gait features. Consequently, the pedestrian re-identification rate is improved.

6 Conclusion

This paper proposes a pedestrian re-identification method that combines color, texture and gait features, thereby preserving not only the color and texture features but also the gait features of pedestrians. The experimental results show that the method can suitably improve both the low identification rate caused by pedestrians with similar clothing and postures and the pedestrian re-identification effect. However, the weight coefficient used in similarity fusion in this paper is not general. The coefficients in this paper are obtained by trial and error, and the optimal weights of features cannot be adaptively determined in a direct manner. In the future, the optimal weights under different conditions will be able to be determined by learning.

Funding Statement: This research was funded by the Science and Technology Support Plan Project of Hebei Province (grant numbers 17210803D and 19273703D), the Science and Technology Spark Project of the Hebei Seismological Bureau (grant number DZ20180402056), the Education Department of Hebei Province (grant number QN2018095), and the Polytechnic College of Hebei University of Science and Technology.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- Cheng, D. S.; Cristani, M.; Stoppa, M.** (2011): Custom pictorial structures for re-identification. *Proceedings of the British Machine Vision Conference*, vol. 1, no. 6, pp. 1-11.
- Deng, M.; Wang, C.; Cheng, F.; Zeng, W.** (2017): Fusion of spatial-temporal and kinematic features for gait recognition with deterministic learning. *Pattern Recognition*, vol. 67, pp. 186-200.
- Farenzena, M.; Bazzani, L.; Perina, A.; Murino, V.; Cristani, M.** (2010): Person re-identification by symmetry-driven accumulation of local features. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2360-2367.
- Hervé, J.; Ondřej, C.** (2012): Negative evidences and co-occurrences in image retrieval: the benefit of PCA and whitening. *Computer Vision-ECCV*, vol. 7573, pp. 774-787.
- Hofmann, M.; Bachmann, S.; Rigoll, G.** (2012): 2.5D gait biometrics using the depth gradient histogram energy image. *IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems*, pp. 399-403.

- Kumar, H. P. M.; Nagendraswany, H. S.** (2014): LBP for gait recognition: a symbolic approach based on GEI plus RBL of GEI. *International Conference on Electronics and Communication Systems*, pp. 1-5.
- Li, H. F.; Zhou, W. Y.; Yu, Z. T.; Yang, B.; Jin, H. P.** (2019): Person re-identification with dictionary learning regularized by stretching regularization and label consistency constraint. *Neurocomputing*, vol. 379, pp. 356-369.
- Liao, S. C.; Hu, Y.; Zhu, X. Y.; Li, S. Z.** (2015): Person re-identification by local maximal occurrence representation and metric learning. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2197-2206.
- Liao, S. C.; Zhao, G. Y.; Kellokumpu, V.; Pietikäinen, M.; Li, S. Z.** (2010): Modeling pixel process with scale invariant local patterns for background subtraction in complex scenes. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1301-1306.
- Liu, X.; Tao, D. C.; Song, M. L.; Zhang, L. M.; Bu, J. J. et al.** (2015): Learning to track multiple targets. *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 5, pp. 1060-1073.
- Matsukawa, T.; Okabe, T.; Suzuki, E.; Sato, Y.** (2016): Hierarchical Gaussian descriptor for person re-identification. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1363-1372.
- Mogan, J. N.; Lee, C. P.; Tan, A. W. C.** (2017): Gait recognition using temporal gradient patterns. *5th International Conference on Information and Communication Technology*, pp. 1-4.
- Nejad, M. B.; Shiri, M. E.** (2019): A new enhanced learning approach to automatic image classification based on Salp Swarm Algorithm. *Computer Systems Science and Engineering*, vol. 34, no. 2, pp. 91-100.
- Prosser, B.; Zheng, W. S.; Gong, S. G.; Xiang, T.** (2010): Person re-identification by support vector ranking. *Proceedings of the British Machine Vision Conference*, vol. 2, no. 5, pp. 1-11.
- Rida, I.; Almaadeed, S.; Bouridane, A.** (2014): Improved gait recognition based on gait energy images. *26th International Conference on Microelectronics*, pp. 40-43.
- Sun, Y. F.; Liang, Z.; Yang, Y.; Tian, Q.; Wang, S. J.** (2018): Beyond part models: person retrieval with refined part pooling (and a strong convolutional baseline). *Computer Vision-ECCV*, vol. 11208, pp. 501-518.
- Syed, M. S. S.; Tahir, A. M.; Robina, K.; Syed, S. H.; Faiz, A. S.** (2019): Human behavior classification using geometrical features of skeleton and support vector machines. *Computers, Materials & Continua*, vol. 61, no. 2, pp. 535-553.
- Wang, T. Q.; Gong, S. G.; Zhu, X. T.; Wang, S. J.** (2014): Person re-identification by video ranking. *Computer Vision-ECCV*, vol. 8692, pp. 688-703.
- Wang, Z. Z.; Huo, W.; Yu, P. P.; Qi, L.; Geng, S. S. et al.** (2019): Performance evaluation of region-based convolutional neural networks toward improved vehicle taillight detection. *Science Letter*, vol. 9, no. 18, pp. 1-19.

Yang, Y.; Yang, J. M.; Yan, J. J.; Liao, S. C.; Yi, D. et al. (2014): Salient color names for person re-identification. *Computer Vision-ECCV*, vol. 8689, no. 9, pp. 536-551.

Zhao, H. Y.; Tian, M.; Sun, S. (2017): Spindle net: person re-identification with human body region guided feature decomposition and fusion. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 907-915.

Zhao, R.; Ouyang, W.; Wang, X. G. (2013): Unsupervised saliency learning for person re-identification. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3586-3593.

Zheng, S.; Zhang, J. G.; Huang, K. Q.; He, R.; Tan, T. N. (2011): Robust view transformation model for gait recognition. *18th IEEE International Conference on Image Processing*, pp. 2073-2076.

Zheng, Z. D.; Zheng, L.; Yang, Y. (2018): A discriminatively learned CNN embedding for person re-identification. *ACM Transactions on Multimedia Computing Communications, and Applications*, vol. 14, no. 13, pp. 1-20.

Zhou, S. R.; Ke, M. L.; Luo, P. (2019): Multi-camera transfer GAN for person re-identification. *Journal of Visual Communication and Image Representation*, vol. 59, no. 1, pp. 393-400.