An Improved Gait Recognition Method Using Modified Gait Energy Image

Jing Li
College of Computer Science and Information Engineering Harbin Normal University
Harbin, China
Email: xinyu.cheng [AT] 163.com

Abstract—**Gait recognition is a valuable technology for remote and concealed identity authentication, widely applied in intelligent video monitoring. Existing gait recognition algorithms fall into two categories: appearance-based methods and model-based methods. While gait features differ from static biometric features like faces or fingerprints, they exhibit significant and robust characteristics over a gait cycle. Gait Energy Image (GEI) is a commonly used feature in gait recognition, synthesizing gait images into a single representative image. In this study, we propose an improved gait recognition method that addresses the impact of viewpoint variations, clothing, and carried objects. The method uses modified GEI (MGEI) and view detection and combins two-dimensional principal component analysis and linear discriminant analysis for feature extraction. Experimental results demonstrate the effectiveness of the proposed method in reducing the influence of view variations and achieving robust gait recognition.**

Keywords-component; gait recognition;MGEI; view detection;principal component analysis;linear discriminate analysis

I. INTRODUCTION

Gait recognition is a technology that identifies individuals based on their unique gait characteristics. It offers the advantage of remote and concealed identity authentication, making it highly relevant in intelligent video monitoring. Existing gait recognition algorithms can be broadly classified into appearance-based [1,2] and model-based [3-5] methods. The former utilizes the spatiotemporal shape and movement characteristics of gait sequences, while the latter employs a structure model to measure key parameters such as gait period, frequency, and direction. Unlike static biometric features like faces [6,7] or fingerprints [8], gait features are dynamic and require a complete gait cycle to demonstrate their significance and robustness [9]. Extracting meaningful, robust, and space-efficient gait features from a set of gait image sequences poses a challenge in gait recognition. Gait Energy Image (GEI) [10] is a commonly used feature that synthesizes gait images into a single image, effectively reducing data volume while retaining essential features such as gait outline, frequency, and phase [11]. Literature [12] proposes two gait feature extraction methods: $(2D)^2$ Principal Component Analysis and Weighted $(2D)^2$ Principal Component Analysis [13-14], and achieves gait recognition using the nearest neighbor distance. $(2D)^2$PCA and $W(2D)^2$PCA have stronger adaptability to perspective changes than some traditional feature extraction methods such as principal component analysis (PCA), linear discriminant analysis (LDA) and so on. However, the variety of view, clothing and carried objects still have a significant impact on gait recognition. Building upon the GEI, this study proposes a gait recognition method that addresses the impact of viewpoint variations, clothing, and carried objects. The proposed method introduces modifications to the gait energy image, extracts entropy features, performs view detection using the nearest neighbor criterion, and combines two-dimensional principal component analysis and linear discriminant analysis for feature extraction. These innovations aim to reduce the influence of view variations and improve the robustness of gait recognition.

II. GAIT ENERGY IMAGE

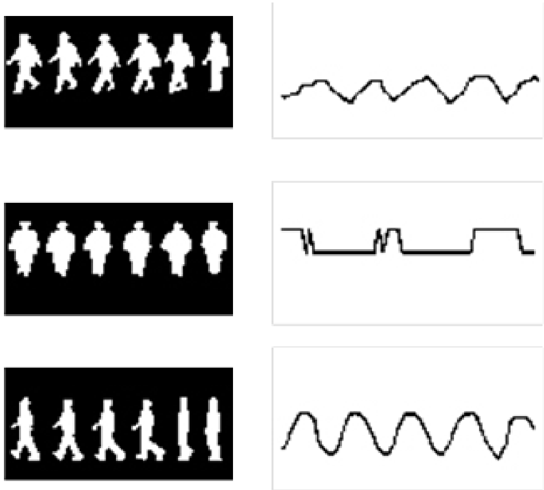
The gait energy image is a representation of gait that combines multiple gait images into a single image using a weighted averaging method. It is defined as follows:

$$G(x,y) = \frac{1}{T}\sum_{t=1}^{T} B_t(x,y) \tag{1}$$

T represents the length of the gait cycle, $B_t(x,y)$ denotes the brightness value of a pixel point (x, y) at time t, and the background area has a brightness value of 0 while the target area has a brightness value of 255.

The gait cycle is typically determined by the aspect ratio of the human side shadow, as described in detail in [15]. The aspect ratio of the human side shadow changes periodically with different gait postures, as illustrated in Figure 1. When the legs are close together, the aspect ratio is the smallest. As the legs move apart, the aspect ratio increases, and when the legs come together again, the aspect ratio reaches a minimum. This aspect ratio variation occurs periodically during the gait cycle.

For this study, the CASIA B gait database [16], provided by the Institute of Automation, Chinese Academy of Sciences, is used. This database contains binary images of each frame of the gait sequence. Therefore, the preprocessing steps for gait image sequences, such as motion detection, image segmentation, and image filtering, are not discussed in this paper. The gait energy image is obtained from the binary gait images using (1). For specific steps, please refer to [15]. The subsequent processing steps are performed on the gait energy image.



(a) Gait sequence (b) Aspect ratio variation curve

Figure 1. Gait aspect ratio variation curve

III. METHOD OF THIS PAPER

In order to enhance the robustness of the gait recognition algorithm against variations in view, clothing, and carried objects, this paper proposes a method that modifies the gait energy image and incorporates view detection to minimize the impact of these factors on gait recognition.

A. Modified gait energy image

To mitigate the influence of clothing and carried objects on gait recognition and emphasize the inherent characteristics of the gait energy image, a modification technique is employed [17]. The gait energy image is filtered in the frequency domain using a Gaussian distribution function. The filtering process can be expressed as:

$$G_F(u,v)^s = G(u,v)F(u,v) \tag{2}$$

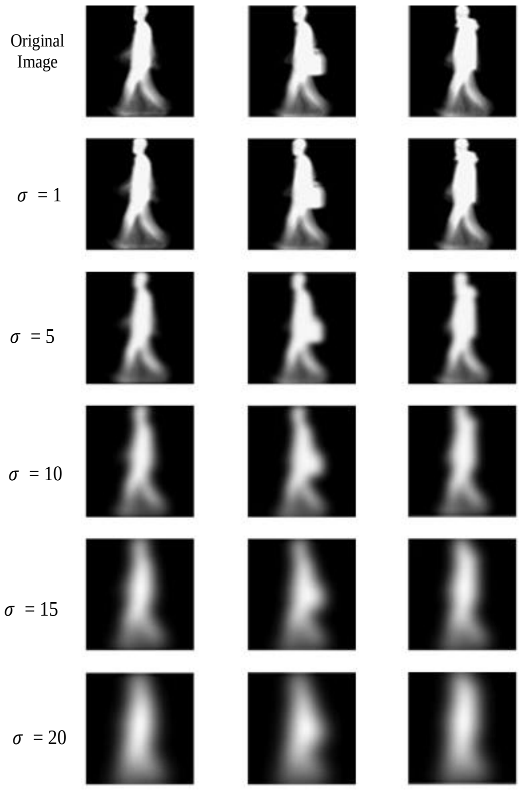
Here, $G_F(u,v)$ represents the filtered gait energy image, $G(u,v)$ represents the original gait energy image, and $F(u,v)$ represents the discrete Fourier transformation of the Gaussian distribution function. By applying the discrete Fourier transformation to $G_F(u,v)^s$, a filtered gait energy image at scale s, denoted as FGEI (Filtered Gait Energy Image), is obtained.

The Gaussian distribution function is defined as:

$$F(x,y) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}} \tag{3}$$

Here, σ represents the parameter controlling the spread of the Gaussian distribution.

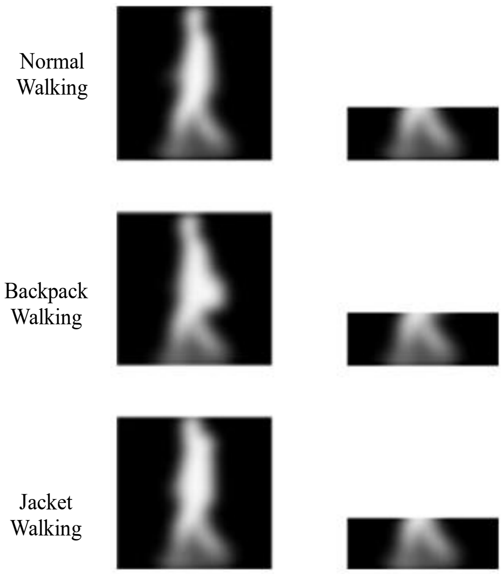
Figure 2 illustrates the FGEI at different scales. As the value of σ increases, the FGEI becomes less affected by clothing and carried objects, resulting in less pronounced differences in gait patterns. From Figure 2, it can be observed that the gait energy image at σ = 10 exhibits strong resistance to interference while still preserving noticeable gait differences. Therefore, this paper adopts σ = 10 as the parameter for the Gaussian distribution.



(a)Normal walking (b) Backpack walking (c)Jacket walking

Figure 2. Gait Energy Image Filtering Effect

As depicted in Figure 3(a), the upper body region of the human body is more susceptible to the influence of accessories such as backpacks and clothing, which can weaken the robustness of the gait energy image. To further mitigate the impact of clothing and carried objects on gait recognition, this paper focuses solely on the gait energy map of the lower limbs for gait recognition purposes. Specifically, the lower third part of the FGEI is extracted and used to construct a new gait energy image, referred to as the modified gait energy image (MGEI), as illustrated in Figure 3(b). From this point onward, the term "gait energy image" refers to the MGEI.



(a) Original gait energy image (b) Clipped gait energy image

Figure 3. Gait Energy Image Clipping Effects

B. *View Detection*

To address the issue of changes in view, the paper utilizes view detection to identify the view of an unknown gait sequence and find matching categories in the same view databases for individual identification, thereby reducing the impact of view changes on gait recognition.

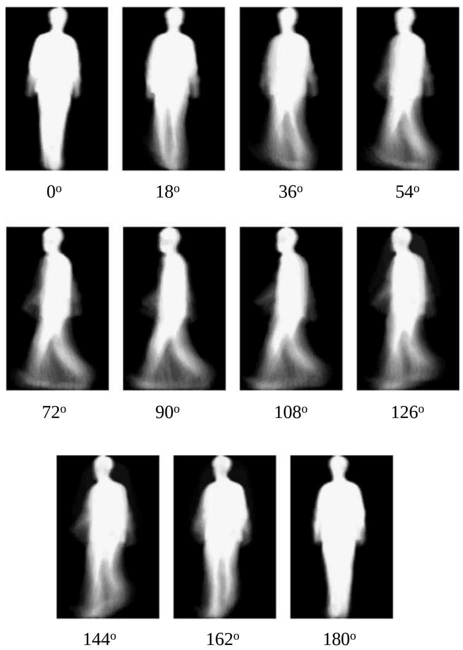


Figure 4. Gait energy images of 11 views

From the CASIA B database, gait energy images with 11 different views can be constructed, ranging from 0º to 180º as shown in Figure 4.

To obtain the gait energy image with a view of (180º, 360º), the paper employs an image mirroring method by applying the following transformation:

$$G^m(x, y) = G(-x, y) \quad (4)$$

Here, $G(x, y)$ represents the original gait energy image, and $G^m(x, y)$ is its mirror image.

By applying the mirroring method, gait energy images with 10 views from 18º to 180º are generated, resulting in gait energy images with views of 198º, 216º, 234º, 252º, 270º, 288º, 306º, 324º, 342º, and 360º, as depicted in Figure 5.

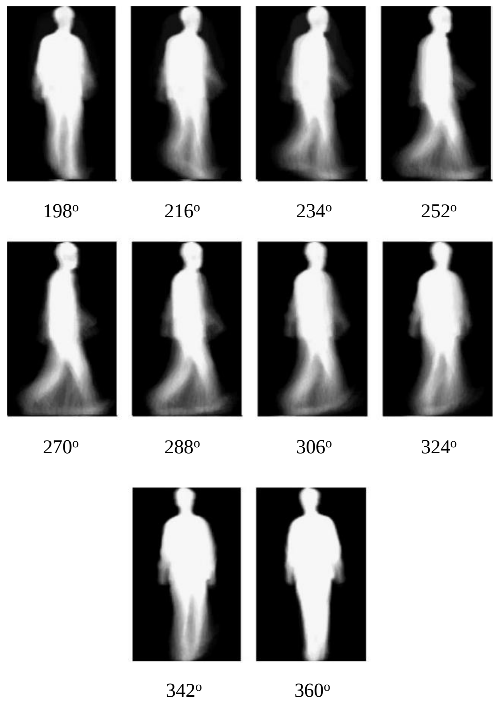


Figure 5. Gait energy maps of 10 views generated by mirroring

Entropy is utilized to effectively describe the texture of the modified gait energy image (MGEI). The paper employs entropy to characterize the features of MGEI using the following expression:

$$E = -\sum_{l=1}^{L} P_l(x, y) \log_2 P_l(x, y) \quad (5)$$

Here, $P_l(x, y)$ represents the probability that the brightness of the pixel point (x, y) in MGEI is l, and L denotes the gray level of MGEI.

The entropy features set with 21 views is denoted as $\{E_1, E_2, \cdots, E_{21}\}$. The dispersion matrix of the entropy features set is calculated as:

$$S = \frac{1}{21}\sum_{i=1}^{21\Sigma}(E_i - \bar{E})^{\bar{T}_i} \tag{6}$$

Here,

$$\bar{E} = \frac{1}{21}\sum_{i=1}^{21} E_i \tag{7}$$

represents the mean of the entropy features.

For an MGEI with an unknown view, the entropy feature E_x is considered. By computing the inner product between E_x and,the eigenvectors corresponding to the first d largest eigenvalues of matrix S (where d = 8 in this paper), a feature image V is obtained, $V = [Y_1, Y_2, \cdots, Y_d]$,

$$Y_k = E_x X_k \tag{8}$$

Here, $X_k, k = 1,2,\cdots,d$, represents the eigenvector corresponding to the d*-th* largest eigenvalue. V is a feature image that represents the unknown view.

Similarly, feature images $V^i = [Y_1^i, Y_2^i, \cdots, Y_d^i], i = 1,2,\cdots,21$, are obtained to represent the 21 different views. The Euclidean distance between the feature image V with 21 different views can be calculated as:

$$D(V^i, V^j) = \sum_{k=1}^{d} \|Y_k^i - Y_k^j\|_2 \tag{9}$$

Based on the nearest neighbor criterion, the view V corresponding to the feature image with the smallest distance is recorded as:

$$\tilde{\imath} = Arg[\min_i D(V^i, V)], i = 1,2,\cdots,21 \tag{10}$$

C. Gait Recognition

After obtaining the views of an unknown gait, the similarity between gaits with the same view can be calculated to reduce the impact of the view on gait recognition performance.

Two-dimensional Principal Component Analysis (2DPCA) is used to extract features from the Modified Gait Energy Image (MGEI) database$\{G_F^{(1)}, G_F^{(2)}, \cdots, G_F^{(n)}\}$, where n is the number of gait samples. The scatter matrix is defined as:

$$S_F = \frac{1}{n}\sum_{i=1}^{n\Sigma}(G_F^{(i)} - \bar{G}_F)^{T_F^{(i)}}{}_F \tag{11}$$

Here,

$$\bar{G}_F = \frac{1}{n}\sum_{i=1}^{n} G_F^{(i)} \tag{12}$$

represents the mean of the gait energy images.

From the scatter matrix S_F, N non-zero eigenvectors are selected to construct a projected image, given by:

$$Y^{(i)} = [\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_N]^T G_F^{(i)}, i = 1,2,\cdots,n \tag{13}$$

The set $\{Y^{(1)}, Y^{(i2)}, \cdots, Y^{(n)}\}$ belongs to C gait classifications.

To reduce within-class scatter and increase between-class scatter, the paper utilizes the 2D Linear Discriminant Analysis (2DLDA) method to find the optimal projection vector $\boldsymbol{W}$. The Fisher criterion function is defined as:

$$J(\boldsymbol{W}) = \frac{|W^T S_B W|}{|W^T S_W W|} \tag{14}$$

Here, S_W and S_B represent the within-class scatter matrix and the between-class scatter matrix, respectively.

$$S_W = \sum_{i=1}^{C\Sigma}\sum_{Y\in D_i}^{\Sigma}(Y - m_i)(Y - m_i)^T \tag{15}$$

$$S_B = \sum_{i=1}^{C\Sigma} n_i(m_i - m)(m_i - m)^T \tag{16}$$

Here, m_i represents the mean of the samples in class i, and m represents the mean of all the samples, as defined below:

$$m_i = \frac{1}{n_i}\sum_{Y\in D_i} Y \tag{17}$$

$$m = \frac{1}{n}\sum_{i=1}^{C}\sum_{Y\in D_i} Y \tag{18}$$

D_i refers to the sample set belonging to the *ith* category in the database. n_i represents the number of samples in the set D_i.

The 2DLDA method searches for the optimal projection vector $\boldsymbol{W}$ to maximize $J(\boldsymbol{W})$. The optimal projection direction is the normalized eigenvector corresponding to the maximum eigenvalue of $(S_W)^{-1}S_B$, given by:

$$S_B w_i = \lambda_i S_W w_i, i = 1,2,\cdots,C \tag{19}$$

The feature vector $\nu_i, i = 1,2,\cdots,C$, is used to generate the gait feature matrix $Z^{(i)}$, given by:

$$Z^{(i)} = [\nu_1, \nu_2, \cdots, \nu_{C-1}]^T Y^{(i)}, i = 1,2,\cdots,n \tag{20}$$

For each class j, R_j is a set of gait feature matrices with n_j elements, and its centroid is denoted as G_j. The nearest neighbor criterion is employed for classification. The Euclidean distance between the unknown gait feature matrix Z and the centroid of the *j-th* class gait feature matrix set in the database is calculated as $D(Z, G_j)$. The category with the smallest distance is selected as the recognized gait category, given by:

$$\tilde{\jmath} = Arg[\min_j D(Z, G_j)], j = 1,2,\cdots,C \tag{21}$$

IV. SIMULATION AND ANALYSIS

A. Experimental Database and Evaluation Indicators

In this section, the CASIA B gait database provided by the Institute of Automation, Chinese Academy of Sciences ,is used. The database consists of 124 individuals, including 93 males and 31 females. It contains 11 gait views ranging from 0° to 180°, with an 18° difference between adjacent views. Each individual has 10 video sequences for each view, including 6 normal walking sequences, 2 backpack walking sequences, and 2 jacketed walking sequences. The videos have a frame rate of 25fps and a resolution of 320 × 240. In this section, the training and testing process involves using the first two normal walking video sequences from each human target across all views. These sequences are used to train the algorithm and evaluate its performance with other video sequences.

For the evaluation of the gait recognition system, the Average Recognition Rate (ARR) index is commonly used. It is calculated as follows:

$$\text{ARR} = \frac{Number\ of\ targets\ correctly\ identified}{Total\ number\ of\ targets} \times 100\% \quad (22)$$

The ARR provides a measure of the system's performance in correctly identifying the gait of the individuals in the database.

B. View Detection Experiment and Analysis

One of the key differences between the proposed method and existing gait recognition methods is the view detection step, which determines the view of the gait before carrying out gait recognition. Figure 6 presents the results of the view detection experiment conducted on the 11 views in the CASIA B gait database.

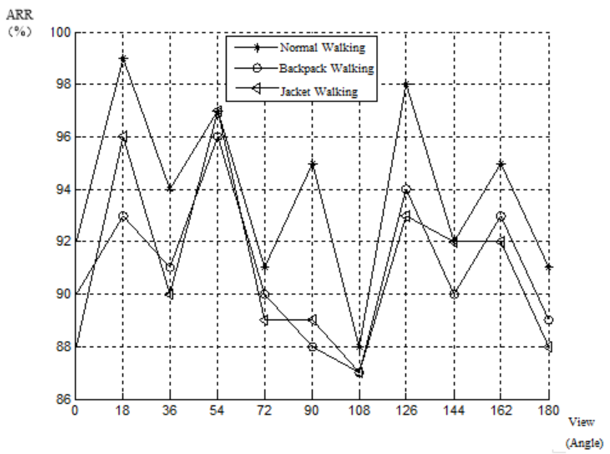


Figure 6. View detection result

The results show that the average recognition rate of the view detection method proposed in this paper is above 87% for different views. The overall average recognition rate for all views is 92.42%. These results indicate that the method presented in this paper performs well in accurately detecting the views of gait sequences.

C. Comparative Experiment and Analysis of Gait Recognition

To assess the performance of the proposed method, comparative experiments were conducted with other commonly used gait recognition methods, including PCA, LDA, (2D) ^{2}PCA, and W (2D) ^{2}PCA. The average recognition rate index for each method is presented in Tables 1 to 3. The average recognition rates for PCA, LDA, (2D) ^{2}PCA, and W (2D) ^{2}PCA were obtained from literature [15].

From the tables, it can be observed that the proposed method achieves higher average recognition rates than the other four methods across different views and walking states. Particularly, in individual views such as 36°, the proposed method exhibits superior performance. This can be attributed to the incorporation of the view detection strategy in the proposed method. The average recognition rate index differences of the proposed method are significantly smaller than those of the other four methods under different views, with all methods achieving recognition rates of over 70%. This indicates that the proposed method is more robust to changes in view compared to the other four methods. These results demonstrate that the proposed method outperforms the compared methods in terms of gait recognition accuracy, especially in individual views, and exhibits strong robustness to variations in view.

TABLE I. AVERAGE RECOGNITION RATE OF DIFFERENT METHODS IN NORMAL WALKING

View	Methods				
	(2D)²PCA	*W(2D)² PCA*	*PCA*	*LDA*	*Method of this article*
0°	85.3%	85.3%	84.1%	85.1%	87.3%
18°	-	-	-	-	86.7%
36°	42.0%	42.0%	43.2%	41.8%	70.6%
54°	56.6%	56.6%	55.8%	56.2%	73.2%
72°	81.3%	81.3%	79.4%	81.0%	87.9%
90°	80.7%	80.7%	81.6%	80.4%	89.7%
108°	81.3%	81.3%	77.5%	78.1%	83.5%
126°	-	-	-	-	87.5%
144°	67.7%	67.7%	65.6%	67.2%	75.0%
162°	-	-	-	-	79.2%
180°	86.9%	86.9%	85.1%	86.5%	89.1%

TABLE II. AVERAGE RECOGNITION RATE OF DIFFERENT METHODS IN BACKPACK WALKING STATE

View	Methods				
	(2D)²PCA	*W(2D)² PCA*	*PCA*	*LDA*	*Method of this article*
0°	73.4%	75.4%	70.5%	73.1%	79.0%
18°	-	-	-	-	71.4%
36°	52.0%	52.0%	50.7%	51.9%	73.0%
54°	65.7%	65.7%	65.9%	65.2%	73.0%

View	**Methods**				
	$(2D)^2$PCA	***W$(2D)^2$ PCA***	***PCA***	***LDA***	***Method of this article***
72°	78.6%	79.8%	74.9%	78.1%	85.1%
90°	84.6%	84.2%	85.1%	84.1%	87.9%
108°	77.8%	78.1%	73.8%	77.1%	80.2%
126°	-	-	-	-	79.8%
144°	73.4%	72.6%	70.8%	73.0%	78.6%
162°	-	-	-	-	79.4%
180°	81.8%	81.5%	80.3%	81.1%	89.2%

TABLE III. AVERAGE RECOGNITION RATE OF DIFFERENT METHODS IN JACKET WALKING STATE

View	**Methods**				
	$(2D)^2$PCA	***W$(2D)^2$ PCA***	***PCA***	***LDA***	***Method of this article***
0°	80.6%	81.4%	79.1%	80.1%	83.1%
18°	-	-	-	-	73.0%
36°	58.8%	57.7%	58.1%	58.9%	71.0%
54°	71.7%	71.4%	70.6%	71.5%	75.8%
72°	87.1%	89.2%	84.1%	86.9%	89.2%
90°	89.1%	90.3%	88.9%	89.0%	90.7%
108°	86.2%	85.1%	85.1%	86.1%	88.3%
126°	-	-	-	-	73.0%
144°	81.8%	80.2%	80.7%	81.8%	85.1%
162°	-	-	-	-	89.2%
180°	89.5%	89.9%	87.4%	89.1%	91.5%

ACKNOWLEDGMENT

In conclusion, a gait recognition method based on modified gait energy image and view angle detection was proposed to mitigate the influence of view changes, clothing variations, and carrying objects on gait recognition performance. The proposed method first modifies the gait energy image to reduce the impact of changing clothing and carrying objects on the gait energy representation. Additionally, the view angle of the gait sequence is detected using entropy features and nearest neighbor classification criteria. This allows for accurate identification of the view, which in turn improves gait recognition performance by reducing the impact of view changes. Gait features are then extracted using a combination of 2DPCA and 2DLDA techniques, and gait recognition is performed based on nearest neighbor distance. The experimental results on the CASIA B gait database demonstrate that the proposed method achieves higher average recognition rates compared to other methods. Furthermore, it exhibits robustness to changes in view, clothing, and carrying objects. However, it is worth noting that the addition of view detection introduces some inefficiency to the method. Further improvements are needed to enhance the efficiency of the proposed method while maintaining its high recognition performance and robustness to various factors.

REFERENCES

[1] Yang Qi, Xue Dingyu, “Gait Recognition Based on Sparse Representation and Segmented Frame Difference Energy Image,” J. Information and Control. China.vol.1, pp.27-32,2013.

[2] K. Annbuselvi, N. Santhi, S. Sivakumar, “Role of Feature Extraction Techniques : PCA and LDA for Appearance Based Gait Recognition,” J. International Journal of Computer Sciences and Engineering. Vol.6 , pp. 38-44, 2018.

[3] Liao R., Li Z., Bhattacharyya S.S., York G, “PoseMapGait: A model-based gait recognition method with pose estimation maps and graph convolutional networks” J. Neurocomputing. Vol.501 , pp. 514-528, 2022.

[4] Boulgouris N V, Hatzinakos D, Plataniotis K N, “Gait recognition: a challenging signal processing technology for biometric identification,” J. Signal Processing Magazine IEEE, vol.22,pp.78-90,2005.

[5] LI Lin-jie,GU Guang-hua,WANG Cheng-ru,ZHANG Chun-feng, “Features Fusion Gait Recognition by Combining Energy Image and Canonical Correlation Analysis,” J. Journal of Chinese Computer Systems. China.vol.11,pp.2558-2561,2014.

[6] Xu D, Huang Y, Zeng Z, et al, “Human Gait Recognition Using Patch Distribution Feature and Locality-Constrained Group Sparse Representation,” IEEE Transl on Image Processing, vol.21,pp.316 – 326,December 2011.

[7] Boulgouris N V, Chi Z X, “Gait Recognition Using Radon Transform and Linear Discriminant Analysis,” J. IEEE Trans Image Process, vol.16,pp.731-740, 2007.

[8] Wang Kejun, Yan Tao, Lü Zhuowen, “Feature level fusion method based on the coupled metric learning and its application in gait recognition,” J. Journal of Southeast University (Natural Science Edition),China.vol. S1,pp.7-11, 2013.

[9] Yang Q, Xue D, “Gait Recognition Based on Sparse Representation and Segmented Frame Difference Energy Image,” J. Information & Control, vol.42pp.27-32,2013.

[10] Huang D Y, Lin T W, Hu W C, et al, “ Gait recognition based on Gabor wavelets and modified gait energy image for human identification,” J. Journal of Electronic Imaging, 2013,vol.22,pp.6931-6946.

[11] Yang Q, Xue D, “Gait Recognition Based on Sparse Representation and Segmented Frame Difference Energy Image,” J.Information & Control, 2013,vol.42,pp.27-32.

[12] BEN Xianye, ZHANG Peng, PAN Tingting, WANG Kejun, “ Performance analysis of matrix gait recognition under linear interpolation framework,” J. CAAI Transactions on Intelligent Systems, 2013, vol.5,pp.415-425.

[13] MA Q Y l, “ Moment Gait Energy Image Based Human Recognition at a Distance,” J. Acta Electronica Sinica, 2012, vol. 35,pp.2078-2082.

[14] Lin C L, Wang K J, Li Y, “ Behavior classification algorithm based on enhanced gait energy image and two-dimensional locality preserving projection,” J. Journal of Computer Applications, 2011,vol.31,pp.721-720.

[15] LIUZhi-yong, FENGGuo-can,CHENWei-fu, “ Gait Recognition Based on Local Binary Pattern and Discriminant Common Vector,” J. Computer Science, 2013, vol.9, pp.262-265.

[16] Li Ping, “ Gait Recognition Method Based on n Haar Wavelet and Fused HMMs ,” J. Computer Applications and Software, 2013,vol.3, pp.244-246.

[17] Jia J, Ruan Q, “ Gait recognition basded on Contourlet Transform and Collaborative Representation,” C.ICSP, IEEE, pp.952-957, 2014