Human Gait Identification Based on Difference Image

Jinyan Chen, Gang Li
School of Computer Software, Tianjin University, Tianjin, China.
Email: {chenjinyan, ligang}@tju.edu.cn

Abstract—Human gait recognition aims to identify people by their walking style. In this paper a difference image based human gait identification method is proposed. For each human gait images sequence, gauss model based background estimation is used to segment frames of the sequence to obtain the silhouette images with less noise. By comparing the difference of two adjacent silhouettes in the images sequence, we can get a differences images sequence. Every difference image in the difference sequence indicates the body moving feature during ones walking. By projecting every difference image to Y axis or X axis we can get two feature vectors. Project every difference images of the whole walking images in one walking cycle we can get two matrixes. These two matrixes indicate the style of ones walking. Then Two-Dimensional Principal Component Analysis(2DPCA) is used to transform the above matrixes to two vectors while at the same time keep the maximum separability. Finally the similarity of two human gait images is calculated by the Euclidean distance of the two vectors. Experimental results on a gait database of 124 people show that the rank 5 identification accuracy can achieve 92%.

Index Terms—human gait, gait identification, principle component analysis, difference image, visual surveillance

I. INTRODUCTION

A. Background of Gait Based Human Identification

Human gait is the manner of ones walking, which was firstly studied by doctors for medical purpose [1-3]. Doctors analyzed their patients walking style to decide whether they were healthy. Then Roislien et al. [4] found that age, gender and weight might influence human gait. BenAbdelkader et al. [5-7] found that just as other biological features, almost everyone has his own walking features. In fact if the sample is properly chosen human gait could also be used as a biological feature to identify human.

In comparison with the first generation biometric such as fingerprint, face and iris, gait based identification has many advantages [8, 9]: (1) Gait based identification could be done at a distance as long as the object is visible. (2) Human gait is difficult to be disguised: fingerprint and iris can be masked by glove or glasses while walking style is difficult to be disguised. (3) Human gait based identification does not require user’s interaction.

Currently gait recognition approaches can be classified into two categories: model-based identification and model free identification. Model-based gait identification try to segment the silhouette into several components according to the structure of human model, then uses the moving model parameters of the above components to describe the gait features. Cunado et al. [11] took Hough translation to extract arms, legs and body from the binarized silhouette and used Fourier analysis to describe the movement the above parts. Yam et al. [12] adopted double pendulum to describe the thigh and lower leg movements. Lee and Grimson [10] used several ellipses to fit the different parts of the silhouette of the person and estimate the parameters from these ellipses. Yoo et al. [13] divided the body into several parts by image segmentation and extracted features from these segments, then support vector machine was used to identify human. Tafazzoli and Safabakhsh [14] constructed movements model based on anatomical proportions then Fourier transform was used to analysis human walking style.

However, the majority of current approaches are the model free approaches. The model free approaches do not model the structure of human motion, but deal directly with image statistical information. Yu et al. [15] overlapped silhouette to get Gait Energy Image (GEI) and used the GEI to identify human. Cheng et al. [9] used Hidden Markov Models to analyze the relationship between these images. Wei Suyuan et al. [16] and Xue Hongmin et al. [17] used “frieze” patterns to extract features from images sequence. Little and Boyd[18] adopted scale independent features from moments of the dense optical flow to represent the shape of human motion. Wang et al. [19] took Chrono-Gait Image as the gait features. Liu et al.[20] represented ones walking style by outermost contour.

In fact model based human identification focus on analysis the moving features of body parts while model free methods focus on the statistical information of the silhouette.

The raw features extracted by model free identification usually serve the curse or dimensions. So Principal Component Analysis (PCA), Gabor transform [21], Random Forest algorithm[22] and Fourier transform [23, 24] were used to reduce the dimension of these features.

Comparing to other biometric there are still several problems need to be resolved in gait based human identification: In case of human having coat or with a bag the recognition error will increase rapidly; Sometimes it is difficult to extract the silhouette from the background or segment legs, arms and head from silhouette;.
B. About PCA and 2DPCA

Principal Component Analysis (PCA) is an efficient method to extract features from raw data. As a dimension reduction method, PCA has been used in the scenario of computer vision and pattern recognition. One dimensional PCA has been used in gait based human identification. Liang Wang et al. [25] used one dimensional PCA to transform the original silhouette matrix to a low-dimensional eigenspace. Dong Xu et al. [26] used average human gait matrix to express the walking features of human then extract features from the average human gait matrix by one dimensional PCA. But to our knowledge most of the PCA based human gait identification are based on one dimension PCA not two dimensions PCA.

C. Our Contribution

In this paper, we propose a model free based gait recognition approach which can tolerate imperfect segmentation to some extension. This paper is an extension to our previous work[27]. Comparing to our precious work in this paper we improved our silhouette extraction method and utilized 2DPCA as the features extraction method. Experiment shows these improvement can achieve better result. In comparison with the works of state-of-the-art, the contributions of this paper are:

Difference Image (DI) based silhouette feature representation method: Model based human identification focused on analysis the moving features of body parts while model free methods focused on the statistical information of the silhouette. Model-based method is more precise but is sensitive to segmentation and prone to be affected by noise. Comparing to model-based method, 2DPCA based feature extraction method: Two-Dimensional Principal Component Analysis (2DPCA) is used to extract features from the feature matrix, which can be more efficient and save more storage comparing to widely used One-Dimensional Principal Component Analysis. To our knowledge nobody has used 2DPCA in human gait identification.

The overview of our gait recognition method is shown in Fig. 1. It contains two major parts – the training part and the testing part. In training part, we first extract features from the input training gait sequences. Then 2DPCA is performed to reduce the dimensionality of the extracted features. In testing part, we also extract feature of the test gait sequence firstly. Then the features in the training stage are used to identify the testing gait sequence.

The remainder of this paper is organized as follows: Section 2 describes our feature extraction, identification and verification method in detail. Experimental results are presented in Section 3, and Section 4 gives conclusions of the paper.

II. GAIT FEATURE EXTRACTION AND HUMAN IDENTIFICATION & VERIFICATION

In this section we will introduce the preprocessing of images and silhouette segmentation and method. Then the construction of Difference Image will be described in detail. Finally we will describe how to extract features from the projected feature matrix and use these features to identify and verify human.
A. Silhouettes Sequence Extraction

In this paper Gauss background model is used to estimate the background. We use the method proposed by Wang [25] to subtract background image from the foreground image:

\[ D(a, b) = 1 - \frac{2\sqrt{(a+1)(b+1)}}{(a+1)+(b+1)} \cdot \frac{2\sqrt{(256-a)(256-b)}}{(256-a)+(256-b)} \]  

(1)

\[ D(a, b) \] is the difference between point a(pixel in foreground image) and point b(pixel in background image). Gauss blur is used to smooth the images to eliminate the affect of noise.

For each image I, the distribution of the above extraction function \( f(a(x,y),b(x,y)) \) over x and y can be obtained. Then moving pixels can be extracted by comparing such a distribution against a threshold value decided by the conventional histogram method. That is:

\[
silhouette(x, y) = \begin{cases} 1 & \text{if} (D_{xy}(a,b) > \Delta) \\ 0 & \text{otherwise} \end{cases}
\]  

(2)

By above method we get a sequence of binary silhouettes images from the walking video. Fig. 2 shows the original image and the image segmentation result.

![Figure 2. (a) The original image in gait database, (b) The silhouette image obtained by formula (1) (2).](image)

B. About Difference Image

We can get the difference image by subtracting two silhouettes image. Given two adjacent images \( i \) and \( i+1 \), then the difference image of \( i \) and \( i+1 \) can be described as following:

\[
diff_{img,(x,y)} = \begin{cases} 1 & \text{if} (img_{i+1}(x,y) = 0 \text{ and } img_{i}(x,y) = 1) \\ 0 & \text{otherwise} \end{cases}
\]  

(3)

Fig. 3 shows binarized two frames extracted from a surveillance video and the difference image.

The difference images can be regarded as the body moving mode during ones walking. We use the following method to reduce the dimensions of the difference image, projecting every difference image to Y axis or X axis then we can get two vectors: PX and PY. For frame \( i \) and \( i+1 \), the PX and PY can be defined as following:

\[
PX_i(x) = \sum_y diff_{img,(x,y)}
\]  

(4)

\[
PY_i(y) = \sum_x diff_{img,(x,y)}
\]  

(5)

The difference image matrices can be defined as following:

\[
MX = \begin{bmatrix} PX_1(l) & \cdots & PX_w(w) \\ PX_{T}(l) & \cdots & PX_{T}(w) \end{bmatrix}
\]  

(6)

\[
MY = \begin{bmatrix} PY_1(l) & \cdots & PY_h(h) \\ PY_{T}(l) & \cdots & PY_{T}(h) \end{bmatrix}
\]  

(7)

T is the total frames count in a walking cycle, \( w \) means the width the difference image and \( h \) means the height. We can visualize MX and MY as Fig. 4.

![Figure 4. By projecting the difference image along X axis and Y axis we can get the gait feature matrices, then stacking these 1D projections over time we can get the 2D patterns.](image)

C. The Average Difference Image

Equation (6) and (7) gives the original definition of difference image projection. This projection matrix is prone to be affected by noise. To make the features more robust we use the average feature matrix as the final feature matrix.
Figure 5. A “half” walking cycle.

In fact a human gait images sequence can be divided to several walking cycles. Fig. 5 shows a gait cycle. In the first stance the person is at rest and the silhouette size is minimal, then he step out his left leg and the silhouette size will increase with his stepping out, in the sixth stand his legs and arms complete separated. Then the legs and arms closed, at stance 11 the size reach minimum again.

By estimating the size of the silhouette we can divide an images sequence into several walking cycles. In fact this is not a complete gait cycle but a “half” cycle, but as we know that the next “half” cycle is just the same as this one, so in this paper we call a “half” gait cycle as one gait cycle.

By the above method we can divide an images sequence into several walking cycles. Then we can get the average project matrix as follows:

\[
M_X = \sum_{i=1}^{c} M_{X'} = \begin{bmatrix} P_{X'}(l) & \ldots & P_{X'}(w) \\
\vdots & \ddots & \vdots \\
P_{X'}(l) & \ldots & P_{X'}(w) \end{bmatrix},
\]

\[
M_Y = \sum_{i=1}^{c} M_{Y'} = \begin{bmatrix} P_{Y'}(l) & \ldots & P_{Y'}(h) \\
\vdots & \ddots & \vdots \\
P_{Y'}(l) & \ldots & P_{Y'}(h) \end{bmatrix},
\]

Where \( c \) represents the cycles. In this paper the Average Difference Image is abbreviated as DI.

D. Two-Dimensional Principal Component Analysis

In traditional PCA method-based image feature extraction methods, every feature must transform to a vector that is mean the 2D image should be transformed to a vector. Then overlapping these vectors usually lead to a high dimensional vector space. The overcome the problem of high dimensional space Jian Yang[28] proposed a two-dimensional based PCA. Comparing to one dimensional PCA this method seems more efficient. The kernel idea of two-dimensional PCA is to project a \( m \times n \) random matrix onto \( Y \) by the following linear transformation[28]:

\[
Y = AX
\]

Where \( A \) means the mapping matrix and \( Y \) means the target matrix. By formula (10) the original matrix is projected to a \( m \)-dimensional vector \( Y \), which is called the projected feature vector of image \( A \).

To estimate the total discriminatory ability of projection vector \( X \) the total scatter of samples can be introduced. The total scatter can be described by the trace of the covariance matrix of the projected feature vectors:

\[
J(X) = \text{tr}(S_x)
\]

Where \( S_x \) means the covariance matrix of the projected feature vectors, \( \text{tr}(S_x) \) means the trace of \( S_x \). By using formula (11) [28], the trace of \( S_x \) can be described as follows:

\[
\text{tr}(S_x) = X^T \left[ (A - E) (A - E)^T \right] X
\]

We define:

\[
G_i = \frac{1}{M} \sum_{j=1}^{M} (A_j - \bar{A}) (A_j - \bar{A})^T
\]

The criterion described formula (11) can be denoted as following:

\[
J(X) = X^T G_i X
\]

The optimal projection axes \( X_1, \ldots, X_d \) should make \( G_i \) has the maximum trace that is to say \( X_1, \ldots, X_d \) are the orthonormal eigenvectors for \( G_i \) corresponding. That is:

\[
(X_1, \ldots, X_d) = \arg \max \ J(X)
\]

\[
X_i X_j = 0 \quad i \neq j \quad and \quad i, j = 1, \ldots, d
\]

And

\[
Y_k = AX_k \quad k = 1, \ldots, d
\]

E. 2DPCA-Based Reconstruction

The principal components and eigenvectors can be combined to reconstruct the original image. Let \( V=[Y_1, \ldots, Y_d] \) and \( U=[X_1, \ldots, X_d] \) and according to formula (16):

\[
U = AV
\]

Since \( X_1, \ldots, X_d \) is orthonormal then according to formula (15) the reconstructed image can be expressed as:

\[
\hat{A} = UV^T = \sum_{i=1}^{d} Y_i X_i^T
\]

That is to say the original image can be reconstructed by the \( Y_i \) and \( X_i \). In this paper the features of sample \( i \) is expressed as \( F_i \) that is:

\[
F_i = (Y_i', \ldots, Y_d')
\]

F. Alignment of the Difference Matrix.

From formula (2) (3) we can see that that the columns and rows of the differential matrix may vary with different walker. That is to say that the feature matrix may has different image size. To make every feature
matrix has the same size we use bilinear interpolation to extend every matrix to the standard width and height. In this paper the width of standard image is 13 pixels and the height is 50.

G. Identification and Verification

2DPCA is used to extract the features from X axes and Y axes projected feature matrix. Then for one walking images sequence i, we can get a d dimensions vector F_i. Given two feature vectors F_i and F_j, the distance is defined as following:

\[ D(F_i, F_j) = \sum_{k=0}^{d} (f^i_k - f^j_k)^2 \]  

(20)

Suppose that the training samples are F_1,...,F_M (Where the M is the total number of training samples), and that each of these samples is assigned a given identity class k. Given a test sample T, this sample belongs to n if

\[ d(T, F_n) = \min_{k} d(T, F_k) \]

In the scenario of verification, the similarity between two feature vectors is defined as the negative of distance, that is:

\[ Sim(F_i, F_j) = -D(F_i, F_j) = -\sum_{k=1}^{d} (f^i_k - f^j_k)^2 \]  

(21)

In this paper the similarity between a probe P_j and F_i in the gallery is defined as z-normed similarity [29]:

\[ Sim(P_j, F_i) = \frac{Sim(P_j, F_i) - \text{Mean}_i Sim(P_j, F_i)}{\text{s.d.}, Sim(P_j, F_i)} \]  

(22)

Where s.d. is standard deviation.

FAR (False Acceptance Rate) , FRR (False Rejection Rate) and EER (Equal Error Rate) are used to evaluate the performance of verification[30].

III. EXPERIMENT AND DISCUSSION

A. About the Dataset

In this paper CASIA database (Dataset B) [31] is used to evaluate our proposed method. The CASIA database is provided by the Institute of Automation, Chinese Academy of Sciences (CASIA). Dataset B consists of 124 subjects (93 males and 31 females) captured from 11 view angles (ranging from 0 to 180 degree with view angle interval of 18). For every person there are six normal walking sequences (named normal-01....normal-06) conducted from every view angle. Every walking sequence contains 3-8 gait cycles (about 40-100 frames). The video frame size is 320×240 pixels, and the frame rate is 25 fps. We use all the 124 objects in Dataset B to carry out our experiments. For each person from every view angle, we select 1-3 walking cycles from sequence normal-01 as the training data (gallery), and one walking cycle from sequence normal-02 as the test data (probe).

For every view angle, each time we leave one training images sequence out and use the remainder as the training set. In the scenario of identification we calculate the distance between the probe corresponding to the leave out training images sequence and the 124 classes (including the leave out images sequence). In the scenario of verification, we calculate the similarity between the probe corresponding to the leave out training images sequence and the 124 classes (including the leave out images sequence).

In all the following experiment 2DPCA method was used to get features from the images and the first 5 eigenvalues are used to produce features except with special declaration.

B The Magnitude of Eigenvalues

After projection of the feature image sample onto these axes using (8), we obtained top twenty principal component. The magnitude of G_t’s eigenvalues is plotted in decreasing order in Fig. 7. As showed in Fig. 7 the magnitude of the eigenvalues quickly converges to zero, which is exactly consistent with the results of Fig. 8.
C. The Reconstruction of the Feature Matrix

From the definition of 2DPCA and formula (18) we can reconstruct the original image from the subimages. Fig. 8 shows the result of the reconstruction.

From Fig. 8 we can find that the first 3 subimages contain most energy of the original image while the others contain detail information about the feature matrix. Therefore, it is reasonable to use these component vectors to represent the image for recognition purposes.

D. Comparison between Our Method and GEI

We compare the performance of our Difference Image (DI) base method with that of Gait Energy Image based method (In this paper we use real template for GEI method). Table I shows the rank 1 and rank 5 identification rates comparing with GEI.

<table>
<thead>
<tr>
<th>View angle</th>
<th>rank 1 performance</th>
<th>rank 5 performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>DI</td>
<td>GEI</td>
<td>DI</td>
</tr>
<tr>
<td>0</td>
<td>67%</td>
<td>68%</td>
</tr>
<tr>
<td>18</td>
<td>50%</td>
<td>37%</td>
</tr>
<tr>
<td>36</td>
<td>31%</td>
<td>22%</td>
</tr>
<tr>
<td>54</td>
<td>30%</td>
<td>26%</td>
</tr>
<tr>
<td>72</td>
<td>60%</td>
<td>44%</td>
</tr>
<tr>
<td>90</td>
<td>80%</td>
<td>77%</td>
</tr>
<tr>
<td>108</td>
<td>70%</td>
<td>62%</td>
</tr>
<tr>
<td>126</td>
<td>40%</td>
<td>36%</td>
</tr>
<tr>
<td>144</td>
<td>39%</td>
<td>34%</td>
</tr>
<tr>
<td>162</td>
<td>42%</td>
<td>35%</td>
</tr>
<tr>
<td>180</td>
<td>86%</td>
<td>84%</td>
</tr>
</tbody>
</table>

From Table I we can see that our method has better performance under view angle 0, 90, 108 and 180. That is may because the image token from above the angle is not prone to be affected by deformation.

To compare the performance of verification, we also evaluate the FAR (False Acceptance Rate) and FRR (False Rejection Rate) for AGDI and GEI. The ROC (Receiver Operating Characteristic) curves under view angle 0, 90, 180 are shown in Fig. 9(a)-(c).

The Comparison Between PCA and 2DPCA.

We also design an experiment to compare the performance of 2DPCA and PCA, which were applied to the feature matrix extracted by this paper. The data set

![Figure 8. The reconstruction of feature matrix (projected to Y axes). All images are enlarged for the sake of clarity.](image)

![Figure 9. a-c The comparison of ROC curves of AGDI and GEI with view angle 0°(a), 90°(b), 180°(c) respectively](image)
view angel is 90.

As illustrated in Fig. 10, the performance of 2DPCA, achieving maximum at about 5 dimensions, is much better than PCA which will achieve the best performance at about 15 dimensions. This experiment shows that 2DPCA is more efficient in dimension reduction.

![Figure 10: The Rank 5 recognition Accuracy comparison of 2DPCA and PCA.](image)

The Affection of Images Cycle Count

From formula (8) and (9) we can see that the feature matrix is the average value of differential images. It should be expected that the use of more images as sample would contribute to a more precise result. To demonstrate this effect, a test was conducted by selectively choosing 1, 2, 3 gait cycles image from 90 degree in sequence normal-02 as test dataset probe. The precision of identification and verification is showed in Fig. 11.

From Fig. 11 we can find that the accuracy of identification and verification will increase with the increase of walking cycles.

IV. CONCLUSION AND FUTURE WORK

In this paper, a difference image and two-dimensional principal component analysis based gait identification and verification method is proposed. The difference image is obtained by comparing two adjacent silhouettes. By projecting difference image to X-axis and Y-axis we get feature matrix. Bilinear interpolation is used to make every feature matrix has the same with and height then 2DPCA is used to get the features. Experimental result shows that 2DPCA has better performance comparing to traditional PCA. Minimum distance classification algorithm is used to identify the human gait. Experiment based on CASIA database result shows our method has an encourage identification and verification performance. In the future there are still some problems to be resolved: Our method is still sensitive to silhouettes segmentation. The matrix alignment method in this paper may lose some useful information.

ACKNOWLEDGMENTS

Dataset used in this paper is provided by Institute of Automation, Chinese Academy of Sciences. This work is partially supported by Ph.D. Programs Foundation of Ministry of Education of China (No. 20100032120011).

REFERENCES


Jinyan Chen born in 1974. Received his Ph.D’s degree at School of Computer Science and Technology in 2008, Tianjin University. He is currently a lecturer in Tianjin University. His research is in the areas of human gait identification, human activity and action recognition.

Gang Li born in 1976. Received his Ph.D’s degree at School of Computer Science and Technology in 2012, Tianjin University. He is currently a lecturer in Tianjin University. His research is in the areas of image processing and software engineering.