GAIT RECOGNITION BASED ON PROCRUSTES SHAPE ANALYSIS*

Liang Wang, Huazhong Ning, Weiming Hu, Tieniu Tan

National Laboratory of Pattern Recognition
Institute of Automation, Chinese Academy of Sciences, Beijing, P. R. China, 100080
{lwang, hzning, wmhu, tnt}@nlpr.ia.ac.cn

ABSTRACT

Gait recognition has recently attracted increasing attention, especially in vision-based human identification at a distance in visual surveillance. This paper proposes a simple but efficient gait recognition algorithm based on statistical shape analysis. For each gait sequence, a background subtraction procedure is used to segment spatial silhouettes of the walking figures from the background. Static pose changes of these silhouettes over time are represented as a sequence of associated complex configurations in a common coordinate, and are then analyzed using the *Procrustes* shape analysis method to obtain gait signature. The *k*-nearest neighbor classifier and the nearest exemplar classifier based on the full *Procrustes* distance measure are adopted for recognition. Experimental results demonstrate that the proposed algorithm has an encouraging recognition performance.

1. INTRODUCTION

Vision-based human identification at a distance in surveillance has recently gained more attentions, e.g., the HID Program being supported by DARPA [1]. This strong interest is driven by some potential applications such as access control for security-sensitive situations. Its research focus is currently on gait, face or activity-specific recognition. As a newly emergent biometric, gait recognition aims at discriminating individuals by the way they walk. Gait has the advantages of being non-invasive and difficult to conceal, and is also the only perceivable biometric at a distance for personal identification.

Although gait recognition is a very new research area, there have been some attempts in the recent literature [2-14]. They can be simply divided into two main classes [10]. The first class, state-space methods, considers gait to be comprised of a sequence of body poses, and recognizes it through considering temporal variations of observations with respect to those static poses [2,3]. For instance, Murase and Sakai [2] presented a template matching method using eigenspace representation to distinguish different gaits. Huang et al. [3] extended the approach of [2] by adding canonical analysis. The second class, spatio-temporal methods, characterizes the spatio-temporal distribution generated by the motion in its continuum [5-13]. The probably earliest approach to recognizing walking figures was due to Niyogi and Adelson [4]. They distinguished different walkers through extracting their spatio-temporal gait patterns obtained from curve-fitting 'snake'. Little and Boyd [6] used frequency and phase features from optical flow information of walking figures to recognize individuals by their gait. More recently, Shutler et al. [7] extended statistical gait recognition via temporal moments, and BenAbdelkader et al. [10] used image self-similarity plot of a moving person to recognize individual gait. Compared with other widely used biometric features such as face and fingerprint, gait recognition is in its infancy. This is because that the existing approaches are usually carried out under some simplifying assumptions that walking people move frontal-parallel to the fixed camera with a relatively plain background. Further, performance evaluation is also usually performed on a small database due to the lack of a common gait database of reasonable size.

Intuitively, recognizing people through gait depends greatly on how the static silhouette shape of an individual changes over time. To implicitly capture the structural characteristics of gait, this paper presents an automatic gait recognition algorithm based on the *Procrustes* shape analysis approach. For each sequence, a background subtraction procedure is used to extract the spatial silhouettes of walking figures from the background. Static pose changes of these silhouettes over time are then represented as a sequence of complex configurations in 2D space and are analyzed by the *Procrustes* shape analysis method to obtain gait signature. The *k*-nearest neighbor classifier and the nearest exemplar classifier based on the full *Procrustes* distance measure are respectively adopted for gait recognition. Experimental results on a relatively larger NLPR database demonstrate that the proposed algorithm has an encouraging recognition performance.

2. FEATURE EXTRACTION

2.1. Segmentation of gait motion

To extract walking figures from the background, change detection based on background subtraction is adopted. First, the Least Median of Squares (*LMedS*) method is used to reliably construct the background image [15]. The brightness change is usually accomplished by differencing between the background and current image. However, the selection of threshold for binarization is very difficult, especially in the case of low contrast images. To solve this problem, we choose the following extraction function to indirectly perform differencing operation [16].

$$f(a,b)=1-\frac{2\sqrt{(a+1)(b+1)}}{(a+1)+(b+1)}\cdot\frac{2\sqrt{(256-a)(256-b)}}{(256-a)+(256-b)} \qquad (1)$$

In Equation (1), $a(x, y)$ and $b(x, y)$ are the brightness of current image and the background at the pixel position (x, y) respectively, $0 \le a(x,y), b(x,y) \le 255$ and $0 \le f(a,b) < 1$. This function can arrange the change sensitivity according to the brightness of each pixel in the background image.

Then, for each image, the changing pixels can be detected by

comparing the above extraction function against a suitable threshold decided using the histogram method. As no change detection algorithm is perfect, morphological operators are used to further filter spurious pixels. Finally, a binary connected component analysis is utilized to extract a single-connectivity moving region. An example of segmentation is shown in Figure 1.

Figure 1. An example of gait segmentation
(a) Background; (b) Original image; (c) The extracted silhouette.

2.2. Representation of silhouette shapes

An important cue in determining internal motion of a walker is his or her silhouette shape changes over time. For the sake of reducing redundant information, we only use spatial edge contours to approximately represent temporal patterns of gaits. The extraction and representation process of the silhouette's boundary is illustrated in Figure 2. After the spatial silhouette of a walking subject has been extracted, its boundary can be easily obtained using a border following algorithm based on connectivity. Then, we can compute its shape centroid (x_c, y_c) by

$$x_c = \frac{1}{N_b}\sum_{i=1}^{N_b} x_i \ , \ y_c = \frac{1}{N_b}\sum_{i=1}^{N_b} y_i \qquad (2)$$

where N_b is the total number of boundary pixels, and (x_i, y_i) is a pixel on the boundary. Let the centroid be the origin of 2D shape space, and we can re-represent shapes as a set of boundary pixel points anticlockwise in a common complex coordinate. That is, each shape can be described as a vector of ordered complex numbers with N_b boundary elements $\boldsymbol{z}=[z_1, z_2, \ldots, z_i, \ldots z_{Nb}]^T$, where $z_i=x_i+i*y_i$. Therefore, each gait sequence will be accordingly converted into a sequence of such 2D shape configurations.

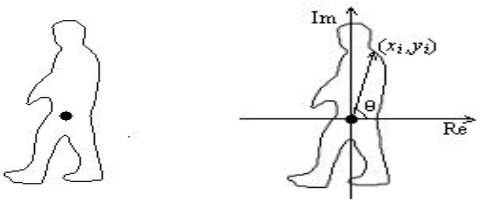

Figure 2. Illustration of shape representation

We need a method that allows us to compare a set of static pose shapes in gait pattern and is robust to position, scale, and slight rotation changes. A mathematically elegant way for aligning point sets in a common coordinate frame is *Procrustes* shape analysis [17]. In the following subsection, we will give a brief introduction to the *Procrustes* shape analysis method and show its application in gait recognition.

2.3. Procrustes shape analysis

Procrustes shape analysis [17] is a particularly popular method in directional statistics [18]. It is intended for coping with two-dimensional shapes and provides a good method to find mean shapes, i.e., a measure of similarity of two shapes. A good review can be found in [11].

A shape in 2D space can be described by a vector of k complex numbers, $\boldsymbol{z}=[z_1, z_2, \ldots, z_k]^T$, called a configuration. For two shapes, $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$, if their configurations are equal through a combination of translation, scaling and rotation, we may consider they are the same shape [11].

$$\begin{cases} \boldsymbol{z}_1 = a\,\boldsymbol{1}_k + b\boldsymbol{z}_2 \,, a\,, b \in C \\ \qquad b = |\,b\,|\,e^{i\angle b} \end{cases} \qquad (3)$$

That is, $a\boldsymbol{1}_k$ translates $\boldsymbol{z}_2$, and $|\,ß\,|$ and $\angle \boldsymbol{b}$ scale and rotate $\boldsymbol{z}_2$. It is very convenient to center shapes through defining the centered configuration $\boldsymbol{u}=[u_1, u_2, \ldots, u_k]^T$, $u_i = z_i - \bar{z}$, $\bar{z} = \sum_{i=1}^{k} z_i / k$.

The full *Procrustes* distance between two configurations can be defined by [11]

$$d_F(\boldsymbol{u}_1, \boldsymbol{u}_2) = 1 - \frac{\left|\boldsymbol{u}_1^T \boldsymbol{u}_2\right|^2}{\|\boldsymbol{u}_1\|^2 \|\boldsymbol{u}_2\|^2} \qquad (4)$$

which minimizes

$$\left\| \frac{\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|} - a\boldsymbol{1}_k - b\frac{\boldsymbol{u}_2}{\|\boldsymbol{u}_2\|} \right\|^2 \qquad (5)$$

Note that the superscript T represents the transpose of the complex conjugate and $0 \le d_F \le 1$. The *Procrustes* distance allows us to compare two shapes independent of position, scale and rotation. Given a set of n shapes, we can find their mean by finding $\boldsymbol{u}$ that minimizes the objective function

$$\min_{a_j, b_j} \sum_{j=1}^{n} \left\| \boldsymbol{u} - a_j \boldsymbol{1}_k - b_j \boldsymbol{u}_j \right\|^2 \qquad (6)$$

To find $\boldsymbol{u}$, we compute the following matrix

$$\boldsymbol{S}_u = \sum_{j=1}^{n} (\boldsymbol{u}_j \boldsymbol{u}_j^T) / (\boldsymbol{u}_j^T \boldsymbol{u}_j) \qquad (7)$$

The *Procrustes* mean shape $\hat{\boldsymbol{u}}$ is the dominant eigenvector of $\boldsymbol{S}_u$, i.e., the eigenvector that corresponds to the greatest eigenvalue of $\boldsymbol{S}_u$ [11].

2.4. Gait signature extraction

Our approach uses these single shape representations from a gait sequence to find their mean shape as gait signatures for recognition. The following summarizes the major steps in determining the *Procrustes* mean shape for a sequence of shapes from n frames, e.g., gait pattern.

1. Select a set of k points from the boundary to represent a 2D shape as a vector configuration $\boldsymbol{z}_j$ in the manner discussed in Section 2.2. We tackle the point correspondence problem through inter-interpolation of boundary pixels so that the set is the same for each image;
2. Set the centered configuration. When we represent the silhouette shape, we have used shape centroid as the origin of 2D shape space to move all shapes to a common center to handle translational invariance. So, we can directly set $\boldsymbol{u}_j=\boldsymbol{z}_j$, j=1, 2, …, n;
3. Compute the matrix $\boldsymbol{S}_u$ using Equation 7. Then, compute the eigenvalues and the associated eigenvectors of matrix $\boldsymbol{S}_u$;
4. Set the *Procrustes* mean shape $\hat{\boldsymbol{u}}$ to the eigenvector that corresponds to the maximum eigenvalue, and this mean shape is

used as the gait signature.

3. CLASSIFIER

We try three different classification methods, namely the nearest neighbor classifier (*NN*), the *k*-nearest-neighbor classifier (*kNN*), and the nearest neighbor classifier with class exemplar (*ENN*).

To measure similarity between two gait sequences, we use the *Procrustes* mean shape distance (*MSD*) in the following way.

1. Compute the *Procrustes* mean shape $\hat{\boldsymbol{u}}_1$ and $\hat{\boldsymbol{u}}_2$ of the two gait sequences as discussed in section 2.4.

2. Find the full *Procrustes* distance between the two mean shapes by

$$\boldsymbol{d}_2(\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2) = 1 - \frac{\left|\hat{\boldsymbol{u}}_1^T \hat{\boldsymbol{u}}_2\right|^2}{\left\|\hat{\boldsymbol{u}}_1\right\|^2 \left\|\hat{\boldsymbol{u}}_2\right\|^2} \tag{8}$$

The smaller the above distance measure is, the more similar the two gaits are.

4. EXPERIMENTS

To verify the usefulness of the proposed algorithm, we have performed a number of experiments. Also, we present detailed analysis and discussion on the experimental results.

4.1. Data acquisition

A digital camera (Panasonic NV-DX100EN) fixed on the tripod is used to capture gait sequences in an outdoor environment. We restrict the experimental situation to a single subject moving in the field of view without occlusion. The subject walks to a stationary camera frontally, laterally, and obliquely ($0^o, 45^o, 90^o$ with respect to the image plane) respectively. The resulting *NLPR* gait database includes 20 subjects and four sequences per view per subject. These images with 24-bit full color are captured at a rate of 25 frames per second and the original resolution is 352×240. The database includes a total of 240 sequences. The length of each sequence varies with the time each person takes to traverse the field of view. Some samples are shown in Figure 3.

(a) Lateral view; (b) Oblique view; (c) Frontal view.

Figure 3. Some samples in the NLPR gait database

4.2. Processing

For each sequence, we perform motion segmentation using the method described in section 2.1. An example of temporal changes in a gait pattern is shown in Figure 4.

Each sequence is accordingly converted into a sequence of shape representations with the associated configurations in 2D space in the manner described as section 2.2. Then, we can obtain their associated mean shapes in the manner described in section 2.4. Further, we use the class average of mean shapes derived from four sequences with the same viewing angle of a subject as an exemplar for the class. Figure 5 shows plots of mean shapes of four sequences and their exemplars, from which we can see that these mean shapes have considerable discriminating power.

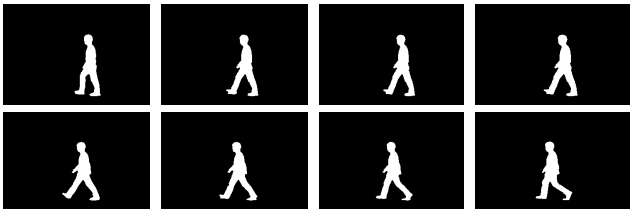

Figure 4. Temporal changes in a gait pattern (Frame 28~35)

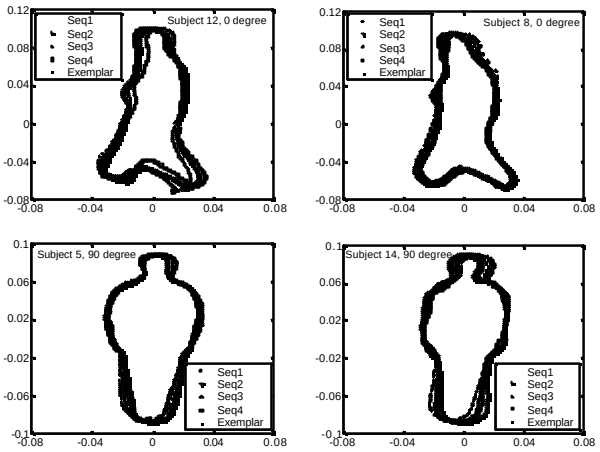

Figure 5. Plots of mean shapes and its exemplar

4.3. Recognition results

Using *MSD*, We try three simple classification methods. In *NN* test, each sequence is classified as belonging to the class of its nearest neighbor. In *kNN* test (*k*=3), we find the three nearest neighbors, and choose the class of the majority, or if no majority, simply the nearest neighbor. The exemplar method (*ENN*) classifies a sequence as the class of its nearest-neighbor exemplar.

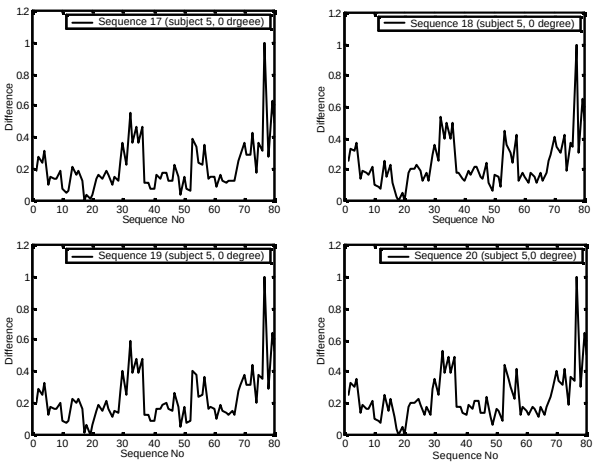

Figure 6. Difference between subject 5 and all others

For a small number of examples, we need compute an unbiased estimate of the true recognition accuracy using a leaving-one-out cross-validation method. First, we label the order from 1 to 80 for the same-type gait sequences subject by subject. Then we leave

one example out, train on the rest, and then classify the omitted element according to its differences with respect to the rest examples based on *MSD*. Figure 6 gives an example to show differences between subject 5 (sequence 17-20) and all others.

The correct classification rates (*CCR*) are summarized in Table 1, from which it can be seen that the recognition performance under frontal walking is the best. This is probably due to the averaging associated with the mean shape analysis owing to less severe shape variations in such gait patterns. Although the results are encouraging, evaluations on a larger database still need to be further investigated in future work in order to be more conclusive.

Table 1. Summary of recognition performance

Classifier	0°	45°	90°
k =1 (*NN*)	71.25	72.50%	81.25%
k =3 (3*NN*)	72.50%	73.75%	80.00%
ENN	88.75%	88.75%	90.00%

4.4. Comparisons

Compared with past work, our method has several advantages: 1) It is easier to comprehend and implement owing to its simpler feature selection. From the segmented silhouette images, past work extracted the features including optical flow [3,6], image self-similarity plots [10], and so on [7,8,13]. However we obtain the feature only from the silhouette boundary, which naturally reduces subsequent computational cost. 2) Our performance evaluation is performed on the presently probably largest database (20 subjects, 240 sequences). Most of past work was realized on a smaller database of usually no more than 12 subjects/48 sequences at the most [2-10,13]. Although our recognition rates are relatively lower than those reported by others based on a small database, the results on a larger database is more convincing. Compared with another recent work [12] (a recognition rate of 73% on a database of 18 subjects/106 sequences), our result is clearly better. 3) Past work was mainly carried out on a small database with only a lateral view [2-13] ([13] used a synthetic virtual lateral view from four viewing angles). Here, we tried three different viewing angles. We examined the recognition sensitivity of the same features to viewingangle and have obtained considerably satisfactory results, which is a more significant progress than previous methods.

4.5. Discussions

To provide a general approach to human identification in unconstrained environments, much remains to be done. 1) Although our results are encouraging, we are limited in our ability to extrapolate them. Our sample size is still small and no steps are taken to ensure a random sample, though it is far lager than past databases. We are planning to establish a larger database. 2) Clothes will bring considerable effects on the shapes of moving people. Our method is based on shape, so it is inevitably affected. In fact, except [5], past work also resulted from this influence owing to their direct use of motion segmentation. Creating more reference sequences with different clothes is probably useful to solve this problem. 3) As shown in our experiments, there is no reason to expect that extracted features are invariant to viewing angle. A useful experiment would be to determine the sensitivity of the features to viewing angles so as to enable a multi-camera tracking system to select an optimal view for recognition. 4) Both static and dynamic information derived from gait plays an important role in gait recognition. To extract dynamic information such as the oscillatory trajectories of joints or limbs may be more useful for recognition. Therefore, 3D human body modeling and tracking might prove to be of benefit.

5. CONCLUSIONS

This paper describes a new gait recognition method based on statistical shape analysis. An improved background subtraction technique is used to segment silhouettes from the background. Shape changes of these silhouettes over time are represented as the associated configurations in the common coordinate frame, and are then analyzed using the *Procrustes* shape analysis method. Experimental results demonstrate that the proposed algorithm has an encouraging performance.

REFERENCES

[1] Website: http://www.darpa.mil/ito/research/hid/index.html.

[2] H. Murase and R. Sakai, "Moving Object Recognition in Eigenspace Representation: Gait Analysis and Lip Reading", *Pattern Recognition Letters*, vol. 17, pp 155-162, 1996.

[3] P. Huang, C. Harris and M. S. Nixon, "Human Gait Recognition in Canonical Space Using Temporal Templates", *Vision Image and Signal Processing*, 146(2), pp. 93-100, 1999.

[4] S. Niyogi and E. Adelson, "Analyzing and Recognizing Walking Figures in XYT", *CVPR*'94, pp. 469-474, 1994.

[5] D. Cunado, M. Nixon and J. Carter, "Using Gait as a Biometric, via Phase-weighted Magnitude Spectra", *AVBPA*'97, pp. 95-102.

[6] J. Little and J. Boyd, "Recognizing People by Their Gait: The Shape of Motion", *Videre: Journal of Computer Vision Research*, 1(2), pp. 2-32, 1998.

[7] J. Shutler, M. Nixon and C. Harris, "Statistical Gait Recognition via Temporal Moments", *IEEE Southwest Symposium on Image Analysis and Interpretation*, pp. 291-295, 2000.

[8] J. Hayfron-Acquah, M. Nixon, and J. Carter, "Automatic Gait Recognition by Symmetry Analysis", *AVBPA*'01, pp. 272-277.

[9] J. Foster, M. Nixon, and A.Prugel-Bennett, "New Area Based Metrics for Gait Recognition", *AVBPA*'01, pp. 312-317, 2001.

[10] C. BenAbdelkader et al., "EigenGait: Motion-based Recognition of People using Image Self-Similarity", *AVBPA*'01, pp. 312-317, 2001.

[11] J. Boyd, "Video Phase-Locked Loops in Gait Recognition", in *Proc. of International Conferences on Computer Vision*, 2001.

[12] R. Tanawongsuwan and A. Bobick, "Gait Recognition from Time-normalized joint-angle Trajectories in the Walking Plane", *CVPR*, 2001.

[13] G. Shakhnarovich, L. Lee and T. Darrell, "Intergrated Face and Gait Recognition From Multiple Views", *CVPR*, 2001.

[14] A. Bobick and A. Johnson, "Gait Recognition Using Static, Activity-Specific Parameters", *CVPR*, 2001.

[15] Y. Yang and M. Levine, "The Background Primal Sketch: An Approach for Tracking Moving Objects". *Machine Vision and Applications*, vol. 5, pp. 17-34, 1992.

[16] Y. Kuno et al, "Automated Detection of Human for Visual Surveillance System", *ICPR*'96, pp. 865-869, 1996.

[17] J. Kent, "New Directions in Shape Analysis", *The art of statistical science: a tribute to g. s. Watson*, pp: 115-127, Wiley, Chichester, 1992.

[18] K. Mardia and P. Jupp, "Directional Statistics", Wiley, Chichester, 2000.

* This work is supported in part by NSFC (Grant No. 69825105 and 60105002).