Automatic Gait Recognition Based on Statistical Shape Analysis

Liang Wang, Tieniu Tan, *Senior Member, IEEE,* Weiming Hu, and Huazhong Ning

***Abstract*—Gait recognition has recently gained significant attention from computer vision researchers. This interest is strongly motivated by the need for automated person identification systems at a distance in visual surveillance and monitoring applications. This paper aims to propose a simple and efficient automatic gait recognition algorithm using statistical shape analysis. For each image sequence, an improved background subtraction procedure is used to extract moving silhouettes of the walking figure from the background. Temporal changes of the detected silhouettes are then represented as an associated sequence of complex vector configurations in a common coordinate frame, and are further analyzed using the *Procrustes* shape analysis method to obtain mean shape as gait signature. Supervised pattern classification techniques based on the full *Procrustes* distance measure are adopted for recognition. This method does not directly analyze the dynamics of gait, but implicitly uses the action of walking to capture the structural characteristics of gait, especially the shape cues of body biometrics. The algorithm is tested on a database consisting of 240 sequences from 20 different subjects walking at 3 viewing angles in an outdoor environment. Experimental results are included to demonstrate the encouraging performance of the proposed algorithm.**

***Index Terms*—Biometrics, gait recognition, statistical shape analysis, visual surveillance.**

I. Introduction

THE demand for automated person identification systems is significantly growing in many important applications such as visual surveillance, access control, and smart interface. Biometrics is intended to address such a need by making use of the physiological or behavioral characteristics of people [46]. Biometric features used currently include fingerprint, voice, iris, face, signature, etc. Face and fingerprint are the two most widely used in some commercial and law applications.

Recently, vision-based human identification at a distance has been strongly driven by the need for automated person identification systems for visual surveillance and monitoring applications. In 2000, DARPA sponsored the Human ID program [1]. The goal of this program is to develop a full range of multimodal surveillance technologies for successfully detecting, classifying, and identifying humans in order to enhance the protection of facilities from terrorist attacks. Although there is a rich body of work describing vision systems to deal with human detection, tracking, and action recognition [23], [51], computer vision researchers have only recently begun to investigate gait as a biometric feature. Gait recognition aims essentially to discriminate individuals by the way they walk, and it is closely related to vision-based human motion analysis methods, especially to methods that deal with whole-body human movements [10].

Gait has the advantage of being noninvasive, just like automatic face recognition. It is less likely to be obscured than other biometric features [21]. Most first-generation biometrics such as fingerprint and face usually require proximal sensing or physical contact, whereas gait is probably the only biometric feature perceivable at a distance where face or iris information is not available in a resolution high enough for recognition. Hence, from the perspective of surveillance, gait is a particularly attractive modality. Apart from being noninvasive, gait is also hard to conceal (i.e., people need to walk, so gait is usually apparent [21]). Gait is readily captured without a walker's attention, so the walkers seldom hide or disguise their gait deliberately. Although it could be argued that some physical factors such as drunkenness, fatigue, pregnancy, and injuries involving joints may affect an individual's gait motion since they will inevitably bring considerable changes to the normal walking pattern, these factors are similar in principle to factors affecting other biometrics [8].

There is a great deal of related studies on gait, including physical medical studies [33]–[35], psychology [29], [42], and approaches aiming to model human body and to track human motion [21], [23], [39], [51]. In addition, gait classification [37], [38] is the recognition of different types of human locomotion such as running, walking, and jumping. All these related subjects lend ample support to the view that gait has clear potential as a biometric feature for recognition.

The purpose and major contributions of this paper are as follows.

- By applying the statistical shape analysis method to automatic gait recognition, this paper aims to develop a simple and effective method for gait-based human identification.
- The proposed method does not directly analyze the dynamics of gait, but derives a compact statistical description of gait as a continuum from its spatio-temporal motion pattern. So it implicitly captures the body structural (appearances) characteristic of gait.
- An integrated background subtraction procedure is proposed. It combines some effective methods demonstrated in different change detection approaches and obtains

Manuscript received September 27, 2002; revised March 14, 2003. This work was supported in part by the NSFC under Grants 69825105 and 60105002, the Natural Science Foundation of Beijing under Grant 4031004, and the National 863 High-Tech R&D Program of China under Grant 2002AA117010. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Philippe Salembier.

The authors are with the National Laboratory of Pattern Recognition (NLPR), Institute of Automation, Chinese Academy of Sciences, Beijing, China, 100080 (e-mail: lwang@nlpr.ia.ac.cn; tnt@nlpr.ia.ac.cn; wmhu@nlpr.ia.ac.cn; hzning@nlpr.ia.ac.cn).

Digital Object Identifier 10.1109/TIP.2003.815251

smoother gait detection results (which are critical to gait analysis).

- Instead of silhouette images usually used in existing work, here we only analyze binary silhouette boundary shape (i.e., outer-contour). This greatly reduces the subsequent computational cost.
- Performance evaluation is performed on our newly established database. Unlike most previous small databases which only involve a lateral view with respect to the image plane, the database used in this paper includes image sequences taken from three different views.
- A large number of papers in the literature usually reported good recognition results on databases of limited size, but few made informed comparisons among different algorithms. Here, we provide some quantitative comparative experiments to examine the performance of the proposed algorithm.
- The proposed method has several desirable properties: 1) it is very easy to comprehend and implement; 2) as a silhouette-based method, it is insensitive to the color and texture of cloth; 3) the statistical nature of signature extraction makes the method considerably robust to noisy data; 4) it does not require precise segmentation of body parts; 5) it has a relatively lower computational cost; and 6) gait sequences of just over 60 frames with respect to a frame rate of 25 fps are sufficient to obtain a good steady recognition performance.

The remainder of this paper is organized as follows. Section II introduces previous related work. Algorithm overview is given in Section III. Section IV describes human silhouette extraction and representation. Section V discusses gait signature extraction and classification. A large number of experimental results are presented and discussed in Section VI. Section VII concludes this paper.

II. Related Work

Human gait has always been an active research topic in biomechanics, kinesiology, physical medicine for therapy, etc. Medical study from Murray [35] supported the view that if all gait movements were considered then gait was unique. Another psychological research from Johansson [29] showed that people did have a remarkable ability to recognize different types of motion (e.g., gait patterns) even only by watching video sequences of light points affixed to joints of the walker. The earliest attempt to recognize people by gait was probably due to Cutting and Kozlowski [28]. A more recent study by Stevenage [19] again confirmed the possibility of learning gait patterns and recognizing people by their gait.

Interest in automatic gait recognition in the computer vision community only began recently, but considerable efforts have already been made and a large number of methods have been proposed [2]–[14], [20], [22], [24]–[26], [30], [31], [36], [40], [41], [43]–[45], [47]–[50], [52]. These methods can be roughly divided into two major categories, namely model-based methods [5], [6], [12], [14], [20], [24], [25], [30], [31], [36], [48]–[50] (which usually model the human body structure and extract image features to map them into the structural components of models or to derive motion trajectories of body parts) and motion-based methods [2]–[4], [7]–[11], [22], [26], [41], [43], [44], [47], [52] (which generally characterize the whole motion pattern of the human body by a compact representation regardless of the underlying structure).

Model-based approaches aim to explicitly model the human body or motion, and they usually perform model matching in each frame of a walking sequence in order that the parameters such as trajectories, limb lengths, and angular speeds are measured on the model. As a typical example of model-based approaches, Cunado *et al.* [5] considered legs as an interlinked pendulum, and gait signatures were derived from the frequency components of the variations in the inclination of human thigh. These features were analyzed using the phase-weighted Fourier magnitude spectrum to recognize different people. Johnson and Bobick [14], [20] used activity-specific static body parameters for gait recognition without directly analyzing the dynamics of gait patterns. In addition, Yam *et al.* [24] first used running to recognize people as well as walking. Later, they further explored the relationship between walking and running that was expressed as a mapping based on phase modulation [48]. The computational cost of model-based methods is relatively high.

The majority of current approaches are motion-based. They typically analyze the image sequence by motion or shape that subjects make as they walk to recover gait features for recognition. In an early work, Niyogi and Adelson [4] distinguished walkers through extracting their spatio-temporal gait patterns obtained from the curve-fitted "snake." Using human shapes and their temporal changes, Murase and Sakai [2] presented a template matching method based on eigenspace representation to distinguish different gaits. Huang *et al.* [3] extended the approach of [2] by adding canonical analysis for better discrimination. More recently, Shutler *et al.* [7] extended statistical gait recognition via temporal moments, Hayfron-Acquah *et al.* [8] described an automatic gait recognition method using the generalized symmetry operator, and Foster *et al.* [9] presented an area mask based metric for gait recognition. In addition, Ben-Abdelkader *et al.* [10], [52] proposed an eigengait method using image self-similarity plots, Phillips *et al.* [47] described a baseline algorithm based on spatial-temporal silhouette correlation for the gait identification problem, and He and Debrunner [45] recognized individual gaits via HMMs that used the quantized vector of Hu moments of a moving person's silhouette as the input.

These early results further confirm that gait has a good potential for personal recognition. Compared with other widely used biometric features such as face and fingerprint, gait recognition is still in its infancy. Gait-based human identification is a challenging problem touching on many hard computer vision problems, e.g., matching temporal signatures, automatic figure and background segmentation, modeling and describing human motion and dynamics, etc. Vision-based gait recognition will thus offer us an interesting research topic.

III. Algorithm Overview

Human gait is usually determined by the persons' weight, limb length, habitual posture and so on. It includes both the

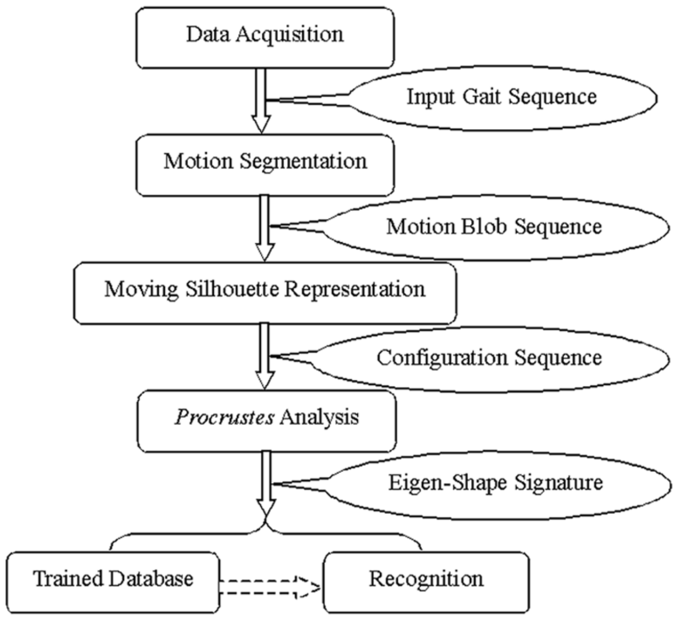


Fig. 1. Overview of the proposed method.

body appearances and the dynamics of human walking motion [41]. In theory, joint angle changes are sufficient for recognition by gait. However, their recovery from a walking video is an unsolved problem for current vision techniques. The particular difficulties of joint angle computation from monocular video sequences are self-occlusions of limbs and joint angle singularities. Empirically, recognizing humans by gait can be achieved by applying the statistical analysis to the temporal patterns of individual subjects, which has been well demonstrated in gait recognition [2], [3], [7], [10], [22], [41], [43], [44], [52]. These techniques remain statistical in essence, describing human motion by a compact representation of motion or structural statistics of a sequence of area distributions rather than the attempt to match the data to a model. Intuitively, recognizing people through gait depends greatly on how the silhouette shape of an individual changes over time. Therefore, we may consider gait to be composed of a set of static poses and their temporal variations can be analyzed to obtain distinguishable signatures. Based upon the above consideration, here we present a model-free automatic gait recognition algorithm using the *Procrustes* shape analysis method.

Fig. 1 gives an overview of the proposed method. For each input sequence, an improved background subtraction procedure is first used to extract the spatial silhouettes of the walking figure from the background. Pose changes of these segmented silhouettes over time are then represented as an associated sequence of complex configurations in a two-dimensional (2-D) shape space and are further analyzed by the *Procrustes* shape analysis method to obtain an eigen-shape as gait signature. The standard pattern classification techniques are adopted for recognition. Like many previous work, this approach also does not directly analyze gait dynamics. It includes the appearance as part of gait recognition features. It is in essence holistic because gait is implicitly characterized by the structural statistics of the spatio-temporal patterns generated by the silhouette of the walking person in image sequences.

IV. Human Silhouette Extraction and Representation

A. Silhouette Extraction

Gait detection is the first step to gait analysis. To extract the walking figure from the background image, change detection based on background subtraction is adopted. Generally speaking, it involves background modeling, the arithmetic subtraction operation and the selection of a suitable threshold.

Background image can be generated by a variety of methods. A potentially more robust approach is to dynamically generate the background image from some portion of image sequence and periodically update it to account for possible changes in the background. Here the Least Median of Squares (*LMedS*) method is used to construct the background image [15]. Let I_{xy}^t represent a sequence including N collected images. The resulting background B_{xy} can be computed by

$$B_{xy} = \min_p \operatorname{med}_t \left(I_{xy}^t - p\right)^2 \tag{1}$$

where p is the background value to be determined for the pixel location (x, y), and t is the frame index ranging within 1–N.

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 as most of moving objects may be missed out since the brightness change is too low to distinguish changing regions from noise. To solve this problem, we use an extraction function to indirectly perform differencing [16]

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

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

For each image, the changing pixels can be detected by comparing the above extraction function against a suitable threshold T decided using the traditional histogram method

$$D_{xy} = \begin{cases} 1, & f(a_{xy}, b_{xy}) \geq T \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

As no change detection algorithm is perfect, there will inevitably be spurious pixels, holes inside moving objects, and other anomalies in the detected sections. Morphological operators such as erosion and dilation are therefore used to further filter spurious pixels and fill small holes inside the extracted silhouettes. Finally, a binary connected component analysis is utilized to extract a single-connectivity moving region. An example of gait detection is shown in Fig. 2.

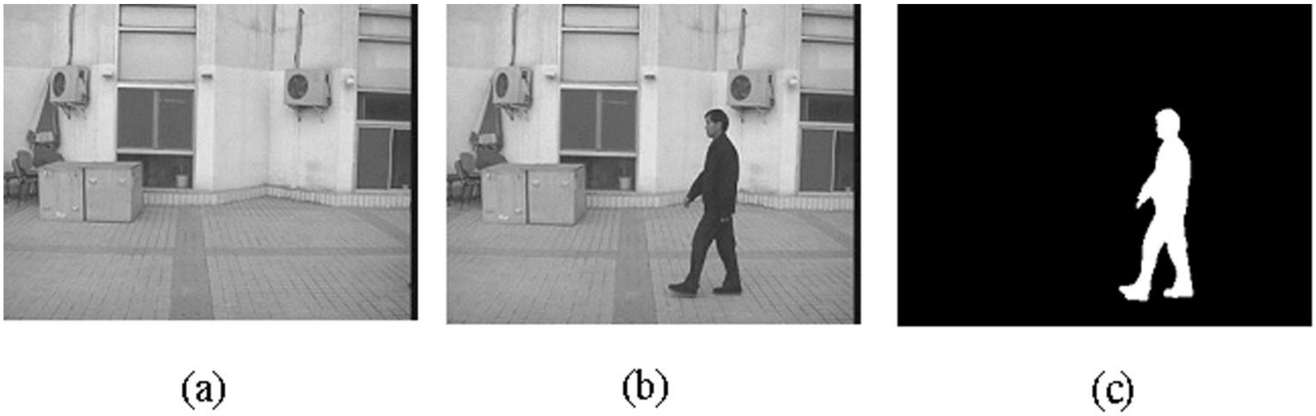


Fig. 2. Example of gait detection. (a) Background image; (b) original image; and (c) extracted silhouette.

B. Representation of Silhouette Shapes

An important cue in determining underlying motion of a walking figure is the temporal changes in the walker's silhouette shape. To make the proposed method insensitive to changes of color and texture of clothing, we ignore the color of the foreground objects and only use the binary silhouette. Further, for the sake of reducing redundant information, we use spatial edge contours to approximate temporal patterns of gaits.

Once the spatial silhouette of a walking subject is 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

$$x_c = \frac{1}{N_b}\sum_{i=1}^{N_b} x_i, \quad y_c = \frac{1}{N_b}\sum_{i=1}^{N_b} y_i \tag{4}$$

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 the 2-D shape space. We can then unwrap each shape anticlockwise into a set of boundary pixel points sampled along its outer-contour in a common complex coordinate system. That is, each shape can be described as a vector of ordered complex numbers with N_b elements

$$\boldsymbol{z} = [z_1, z_2, \ldots, z_i, \ldots, z_{Nb}]^T \tag{5}$$

where $z_i = x_i + j * y_i$. The silhouette shape representation is illustrated in Fig. 3, where the black dot indicates the shape centroid, and the two axes Re and Im represent the real and imaginary part of a complex number, respectively. Therefore, each gait sequence will be accordingly converted into an associated sequence of such 2-D shape configurations.

We need one method that allows us to compare a set of static pose shapes to gait pattern and is robust to changes of position, scale, and rotation. A mathematically elegant way for aligning point sets in a common coordinate system is *Procrustes* shape analysis [17]. So it is expected that it can be easily adapted to handle spatial patterns of gait motion. In the following section, we will give a brief introduction to the *Procrustes* shape analysis method and show its application in gait signature extraction and classification.

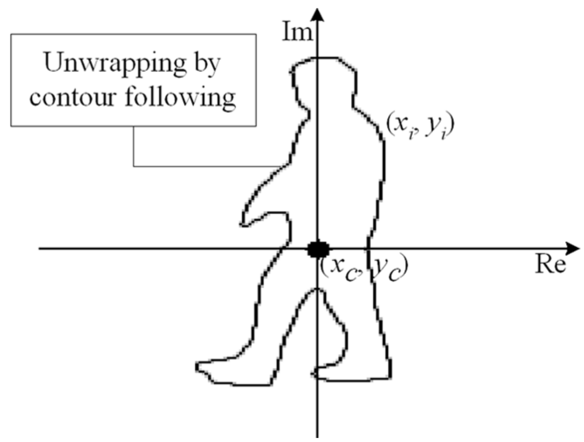


Fig. 3. Illustration of silhouette shape representation.

V. Gait Feature Extraction and Classification

A. Procrustes Shape Analysis

Procrustes shape analysis [17] is a particularly popular method in directional statistics [18], and it is intended to cope with 2-D shapes. A good brief review can be found in [11].

A shape in 2-D space can be described by a vector of k complex numbers, $\boldsymbol{z} = [z_1, z_2, \ldots, z_k]^{\mathrm{T}}$, 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

$$\begin{cases} \boldsymbol{z}_1 = \alpha \mathbf{1}_k + \beta \boldsymbol{z}_2, \alpha, \quad \beta \in C \\ \beta = |\beta| e^{j\angle\beta} \end{cases} \tag{6}$$

where $\alpha\mathbf{1}_k$ translates $\boldsymbol{z}_2$, and $|\beta|$ and $\angle\beta$ scale and rotate $\boldsymbol{z}_2$, respectively, we may consider they represent the same shape [11].

It is very convenient to center shapes by defining the centered configuration $\boldsymbol{u} = [u_1, u_2, \ldots, u_k]^{\mathrm{T}}, u_i = z_i - \bar{z}, \bar{z} = \sum_{i=1}^{k} z_i/k$. The full *Procrustes* distance $d_F(\boldsymbol{u}_1, \boldsymbol{u}_2)$ between two configurations can be defined as

$$d_F(\boldsymbol{u}_1, \boldsymbol{u}_2) = 1 - \frac{|\boldsymbol{u}_1^* \boldsymbol{u}_2|^2}{\|\boldsymbol{u}_1\|^2 \|\boldsymbol{u}_2\|^2} \tag{7}$$

which minimizes

$$\left\| \frac{\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|} - \alpha \mathbf{1}_k - \beta \frac{\boldsymbol{u}_2}{\|\boldsymbol{u}_2\|} \right\|^2. \tag{8}$$

Note that the superscript $*$ represents the complex conjugation transpose 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

$$\min_{\alpha_i,\beta_i} \sum_{i=1}^{n} \|\boldsymbol{u} - \alpha_i 1_{\boldsymbol{k}} - \beta_i \boldsymbol{u}_i\|^2. \tag{9}$$

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

$$\boldsymbol{S_u} = \sum_{i=1}^{n} (\boldsymbol{u}_i \boldsymbol{u}_i^*)/(\boldsymbol{u}_i^* \boldsymbol{u}_i). \tag{10}$$

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].

B. Gait Signature Extraction

Our approach uses these single shape representations from a gait sequence to find their mean shape as gait signatures for recognition. Similar to Eigenface [32], we call this gait signature as "Eigenshape." The following summarizes the major steps in determining the *Procrustes* mean shape for a sequence of shapes from n frames, e.g., a gait pattern.

1) Select a set of k points from the boundary to mean a 2-D shape as a vector configuration $\boldsymbol{z}_i$ in the manner discussed in Section IV-B. We tackle the point correspondence problem through interpolation of boundary pixels so that the point 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 2-D space to move all shapes to a common center to handle translational invariance. So, we can directly set $\boldsymbol{u}_i = \boldsymbol{z}_i, i = 1, 2, \ldots, n$.
3) Compute the matrix $\boldsymbol{S}_u$ using (10). Then, compute the eigenvalues and the associated eigenvectors of $\boldsymbol{S}_u$.
4) Set the *Procrustes* mean shape $\hat{\boldsymbol{u}}$ as the eigenvector that corresponds to the maximum eigenvalue, and this mean shape is used as the gait signature.

C. Similarity Measure and Classifier

To measure similarity between two gait sequences, we make use of 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 V-B.
2) Find the full *Procrustes* distance between the two mean shapes by

$$d(\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2) = 1 - \frac{|\hat{\boldsymbol{u}}_1^* \hat{\boldsymbol{u}}_2|^2}{\|\hat{\boldsymbol{u}}_1\|^2 \|\hat{\boldsymbol{u}}_2\|^2}. \tag{11}$$

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

Gait recognition is a traditional pattern classification problem which can be solved by measuring similarities among gait sequences. We try three different simple classification methods, namely the nearest neighbor classifier (*NN*), the k-nearest-neighbor classifier (k*NN*), and the nearest neighbor classifier with class exemplar (*ENN*).

Let T represent a test sequence and R_i represent the ith reference sequence, we may classify this test sequence into the class c that minimizes the similarity distances between the test sequence and all reference patterns by

$$c = \arg_i \min d(T, R_i) \tag{12}$$

where d is the similarity measure defined in (11).

No doubt, a more sophisticated classifier could be employed, but the interest here is to evaluate the genuine discriminatory ability of the extracted features. We use the leave-one-out cross-validation rule in our experiments in order to obtain an unbiased estimation of recognition accuracy.

VI. Experimental Results

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

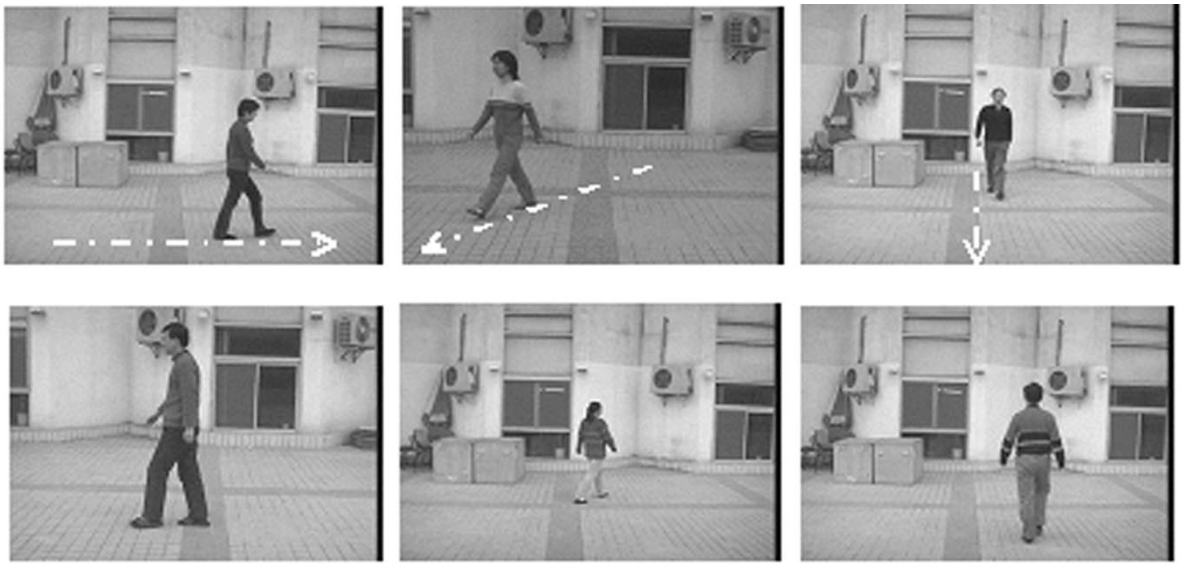
A. Data Acquisition

A new gait database, called the *NLPR* database, has been successfully established for our experiments. A digital camera (Panasonic NV-DX100EN) fixed on a tripod is used to capture gait sequences at a rate of 25 fps on two different days in an outdoor environment. Here we assume that a single subject moves in the field of view without occlusion. All subjects walk on a straight-line path at natural cadences and in three different viewing angles with respect to the image plane, namely frontally (90°), laterally (0°) and obliquely (45°). The images are recovered from video storied in DV tapes to a Microsoft AVI wrapper with an IEEE 1394 interface offline, and finally transcoded using the Sthvcd2000 decoder into 24-bit full-color BMP files with a resolution of 352×240. The resulting *NLPR* gait database includes 20 different subjects and 4 sequences per view per subject. The database thus includes a total of 240 ($20 \times 4 \times 3$) sequences. The length of each collected sequence varies with the pace of the walker, but the average is about 90 frames. Some sample images are shown in Fig. 4, where the white line with arrow represents the walking path.

B. Processing

For each sequence, we perform motion segmentation using the method described in Section IV-A. An example of temporal changes of moving silhouettes in a gait pattern is shown in Fig. 5.

Each sequence is accordingly converted into a sequence of shape representations with the associated configurations in 2-D space (the vector dimensionality is set to 360 here) in the manner described as Section IV-B. Then, we can obtain their associated mean shapes in the manner described in Section V-B. Note that the walking direction is pre-normalized here to avoid the effect on recognition performance, e.g., all sequences with the lateral view are flipped from right to left. Further, we use the class average of mean shapes derived from the same-view sequences of a subject as an exemplar for that class, which aims to avoid selecting a single and random reference sample. Fig. 6 shows plots of mean shapes and their exemplar of four sequences of the same subject and plots of the exemplars of five different



(a) (b) (c)

Fig. 4. Sample images of gait sequences in the *NLPR* database. (a) Lateral view; (b) oblique view; and (c) frontal view.

Fig. 5. Temporal changes of moving silhouette in a gait pattern (Frame 28–Frame 35).

subjects (note that Exemplars 1–5 are corresponding to Subjects 5, 8, 14, 16, and 20, respectively) from which we can see that the intra-subject changes in eigenshapes are very small, while the inter-subject changes are more significant. Such result implies that the mean shapes have considerable discriminating power for identifying individuals.

C. Results

We have tried three classification methods. In the *NN* test, each sequence is classified as belonging to the class of its nearest neighbor. In the k*NN* 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.

First, we evaluate the performance of our approach using classification error in identification mode in which the classifier determines which class a given measurement belongs to. For a small number of examples, we expect to compute an unbiased estimate of the true classification rate using the leave-one-out cross-validation rule since the leave-one-out error rate estimator is known to be an almost unbiased estimator of the true error rate of the classifier. We label the order for the 80 same-view gait sequences subject by subject from 1 to 80. Then we leave one example out, train on the rest, and classify the left-out element according to its *MSD* differences with respect to the rest examples. This process is repeated 80 times, and the recognition rate is obtained as the ratio of the number of correctly classified test samples out of the total 80 for each viewing angle. The correct classification rates (*CCR*) are summarized in Table I. Note that here testing and training are consistent with respect to viewpoints. From Table I, it can be seen that the recognition performance under the frontal walking is better than other two views. This is probably due to the averaging associated with the mean shape analysis owing to less severe shape variations in such gait patterns. It can also be seen that the *ENN* classifier consistently outperforms the other two. For each subject, although his or her gaits at different times are perceived to be almost invariable, there are still slight changes between them. Multiple sample sequences' average may serve to provide a more standard gait pattern for that specific person than a single and random sample se-

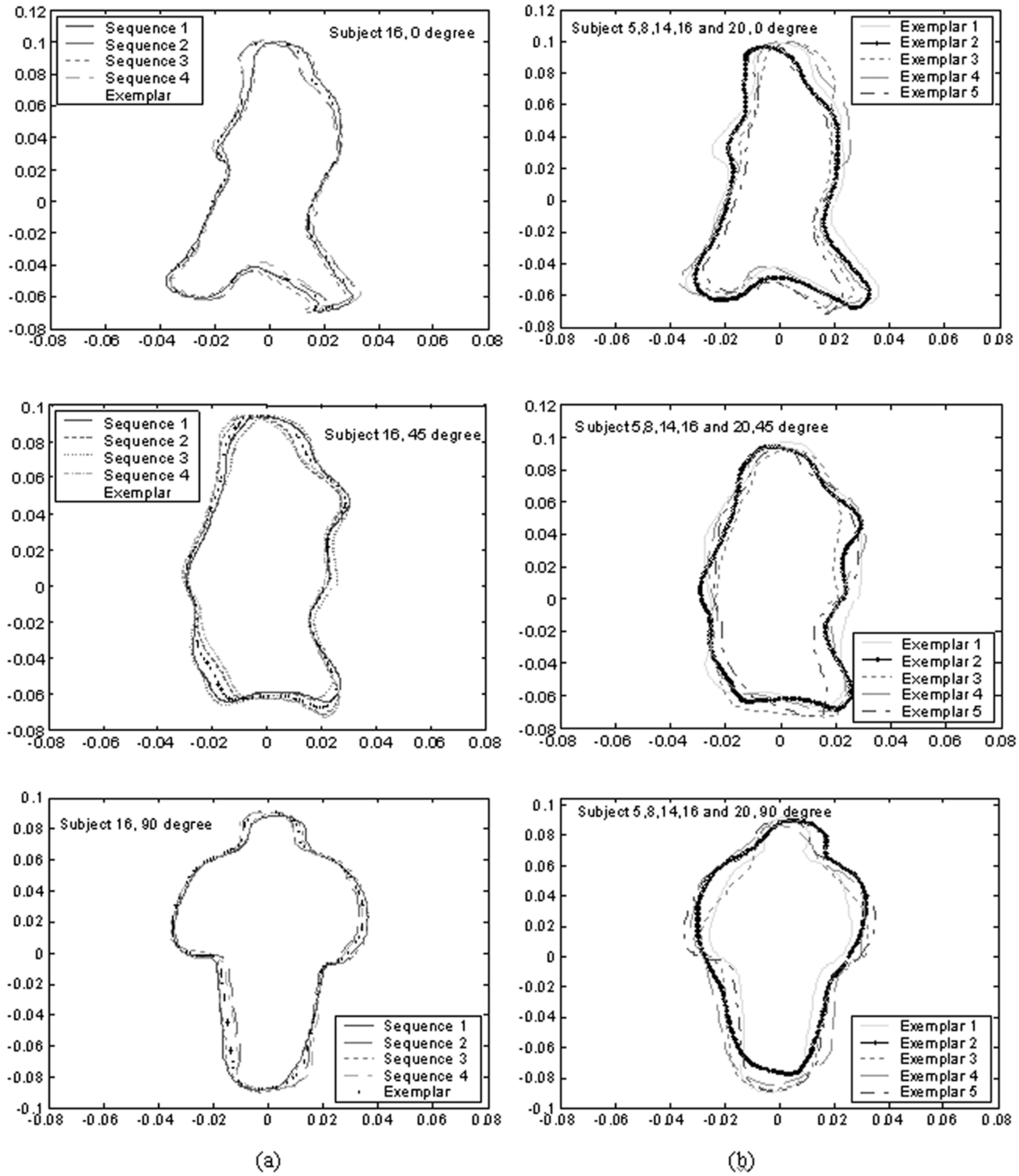


Fig. 6. Plots of mean shapes and the exemplars for different views. (a) Mean shapes and their exemplar of four sequences of the same subject and (b) exemplars of five different subjects.

TABLE I
*CCR*s OF DIFFERENT CLASSIFIERS UNDER DIFFERENT VIEWING ANGLES

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%	87.50%	90.00%

quence. Although as a whole the results are very encouraging, more experiments on a more realistic database still need to be further investigated in order to be more conclusive.

Another useful classification performance measure that is probably more general than classification error is the rank order statistic, which was first introduced by the FERET protocol for the evaluation of face recognition algorithms [27]. It is defined as the cumulative probability $p_{(k)}$ that the real class of a test measurement is among its top k matches. The basic models for evaluating the performance of an algorithm are the closed and open universes. In the closed universe, every probe (unknown measurements) is in the gallery (known measurements). While in an open universe, some probes are not in the gallery. The performance statistics are reported as the cumulative match scores. The rank is plotted along the horizontal axis, and the vertical axis is the percentage of correct matches [27]. Here, we use the closed-universe model and the leave-one-out cross-validation rule with the *NLPR* database to estimate the identification performance of the proposed method. Fig. 7

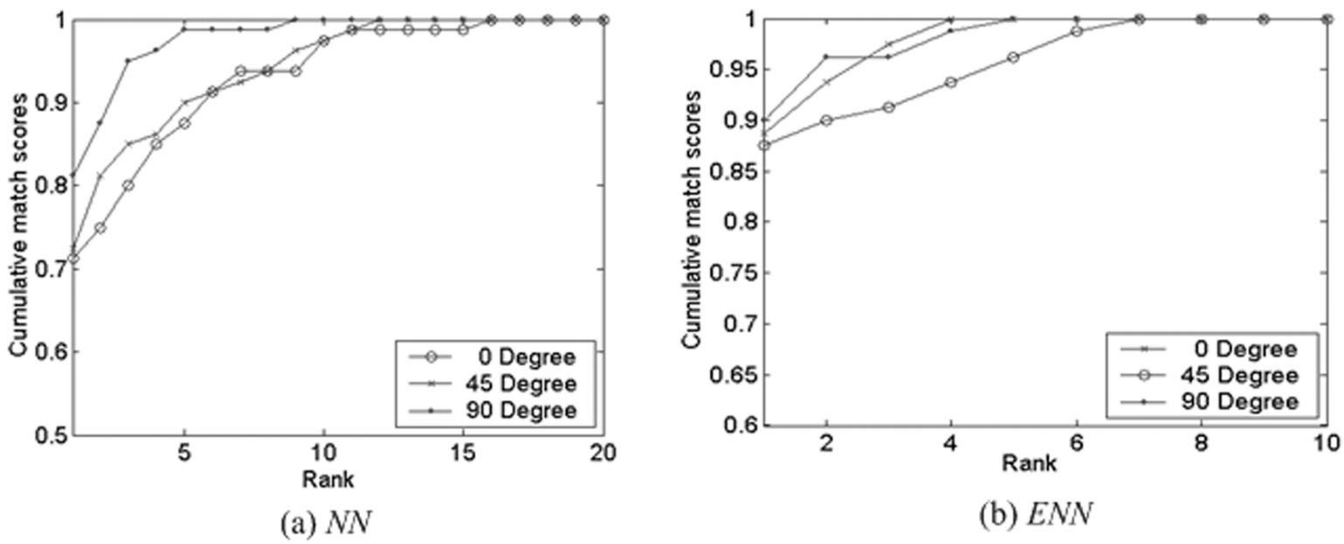


Fig. 7. Identification performance results in terms of the FERET protocol's CMS curve.

shows the cumulative match scores (CMS) for ranks up to 20 in Fig. 7(a) based on *NN* and 10 in Fig. 7(b) based on *ENN* respectively. It is noted that the correct classification rate is equivalent to $p_{(1)}$ (i.e., $\mathrm{Rank} = 1$).

For completeness, we also use the ROC (*Receiver Operating Characteristic*) curve to report results. For verification mode, the pattern classifier is asked to verify whether a new measurement really belongs to certain claimed class. As before, we estimate FAR (*False Acceptance Rate*) and FRR (*False Reject Rate*) via the leave-one-out rule. That is, we train the classifier using all but one left-out sample, and then verify the left-out sample on all 20 classes. Note that in each of these 80 iterations for each viewing angle, there is one genuine attempt and 19 imposters since the left-out sample is known a priori to belong to one of the 20 classes. By varying the decision threshold for the acceptance, we can produce various combination pairs of FAR and FRR. Fig. 8 shows the ROC curve, from which we see that the EERs (*Equal Error Rate*) are about 8%, 12%, and 14% for 0°, 90°, and 45° views respectively.

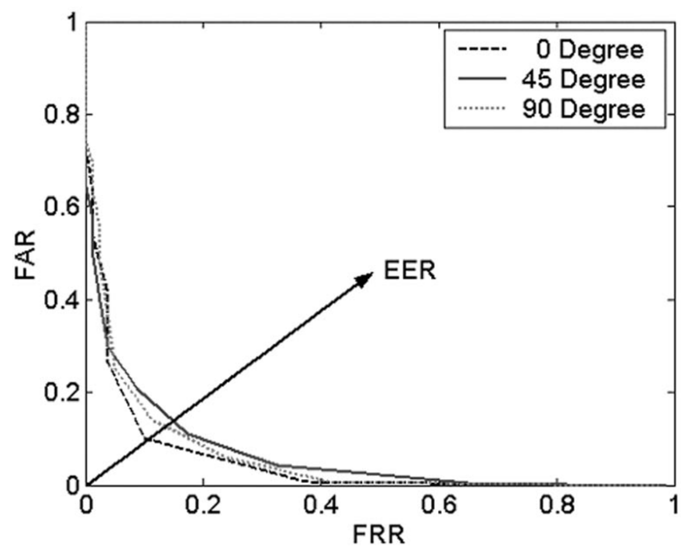


Fig. 8. Verification performance results reported by the ROC curves.

D. Evaluation

The performance of the algorithm is further evaluated with respect to the length of the training sequence and the vector dimensionality of shape representation on the *NLPR* database with the lateral view.

1) Influence of the Dimensionality of Shape Representation: The influence of the dimensionality of shape representation (i.e., the number of points sampled along the boundary contour) is examined by changing the sampling interval. Fig. 9 shows the general trend of correct classification rate vs the dimensionality of shape representation, from which we can see that the *CCR* starts to level off at 36 points. That is, 36 points may be sufficient to represent a shape in 2-D space as far as gait recognition is concerned. Clearly, the reduced dimensionality results in a concomitant decrease in computational cost.

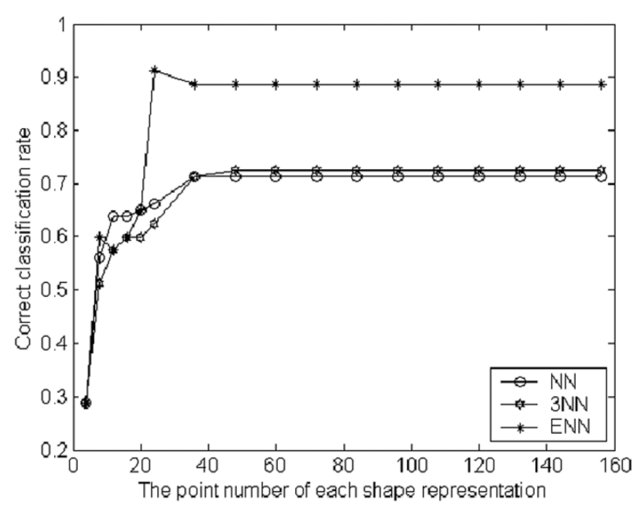


Fig. 9. Performance evaluations with respect to the shape dimensionality.

2) Influence of the Training Sequence Length: To evaluate the effects of the length of training samples, we conducted five tests which respectively make use of the first 15, 30, 45, 60, and 75 frames corresponding approximately to one, two, three, four and five walking cycles from each gait sequence captured at a

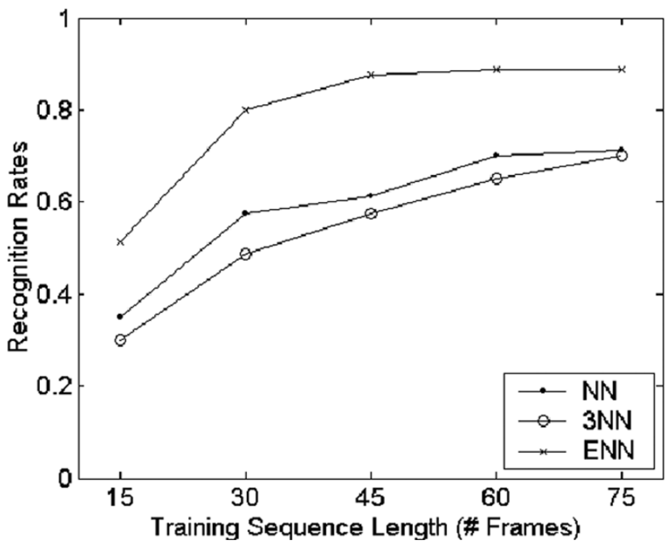


Fig. 10. Performance evaluations with respect to the training sequence length.

rate of 25 fps. (Note that $1 \text{ stride period} = 2$ cycles.) An average cycle is typically 15 frames in terms of 25 fps according to the study of biomechanics though it seems to have a little difference on cadences of different people. The comparisons of recognition performances are shown in Fig. 10. The results reveal that the best performance is achieved by using just over four walking cycles of training samples from each subject (i.e., 60 frames). Furthermore, the recognition performance is improved by increasing the number of training samples. The results thus appear to confirm recognition sensitivity to the sequence length and imply that in a more extended analysis, care must be taken to include sufficient samples in the training data.

E. Comparisons

Identification of people by gait is a challenging problem and has attracted growing interest in the computer vision community. However, there is no baseline algorithm or standard database for measuring and determining what factors affect performance. The unavailability of an accredited common database (e.g., something like the FERET database in face recognition) of a reasonable size and evaluation methodology has been a limitation in the development of gait recognition algorithms. A large number of papers in the literature reported good recognition results usually on a small database, but few of them made informed comparisons among various algorithms. To examine the performance of the proposed algorithm, here we provide some basic comparative experiments.

The first comparative experiment is to test our method on the early *SOTON* gait database [8]. This database collected six subjects and four sequences of each subject. Walkers are required to move frontal-parallel to the image plane. The gray images were captured by a fixed camera with a stationary indoor background at a rate of 25 fps, and the original resolution is 384×288. The length of each sequence is about 60 frames except that the sequences of two subjects have only 30 frames. Fig. 11 gives several samples in the *SOTON* gait database. Nixon and his research group have made one of the first attempts on gait recognition and have developed many algorithms [3], [5], [7]–[9], [24], [30], [43], [48], most of which evaluate performance on the whole or a subset of the *SOTON* database. Hence we evaluate the proposed algorithm on such a database so as to make a direct quantitative comparison with some of their recent methods. Table II shows the comparison results of several different approaches, where we directly select the best recognition accuracy reported in [7], [8], and [9] without re-implementing them. From Table II, we can see that the recognition performance of our method is superior to others. Also, we plan to test the proposed method on their new larger dataset if available.

Another comparative experiment is to compare the performance of the proposed algorithm with those of five recent methods which are from Maryland [10], [52], CMU [26], MIT [41] and USF [47] respectively, and to some extent reflect the best work of these research groups in gait recognition. BenAbdelkader *et al.* [10] used image self-similarity plots as the original measurements to recognize gait based on the idea that the image self-similarity plot of a moving person is a projection of its planar dynamics. Reference [52] is a slight extension of [10]. Based on body shape and gait, Collins *et al.* [26] established a template matching method based on body silhouettes in key frames for human identification. Lee *et al.* [41] described a moment-based representation of gait appearance features for the purpose of person identification and classification. Phillips *et al.* [47] proposed a baseline algorithm for human identification using spatio-temporal correlation of silhouette images. Here, we re-implement these methods using the same silhouette data from the *NLPR* database with a lateral viewing angle. The results are summarized in Table III. From Table III, we can see that our method compares favorably with others, with performance very similar to [41]. Gait feature vector of [41] is composed of parameters of moment features in image regions containing the walking person aggregated over time. Intuitively, the mean features describe the average-looking ellipses for each of the regions of the body; taken together, the 7 ellipses describe the average shape of the body, which is in essence similar to the idea of our method. We see that our method outperforms the methods described in [10], [52], [26], and [47]. From experiments it is also found that the computational cost of [26] and [47] was relatively higher than that of [10], [52], [41] and our method.

The above only provides preliminary comparative results and may not be generalized to say that a certain algorithm is always better than others. Algorithm performance is dependent on the gallery and probe sets. Some similar-size [26], [41], [52] or larger [47] databases have concurrently emerged, so further evaluations and comparisons on a larger and more realistic database are needed in future work.

F. Discussions and Future Work

To provide a more general approach to human identification in unconstrained environments, much work 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. We are planning to establish a larger database which will

Fig. 11. Some samples in the *SOTON* gait database

TABLE II
COMPARISON OF SEVERAL DIFFERENT APPROACHES ON THE *SOTON* DATABASE

Methods	Data sets	Correct Classification Rate (%)
Shutler 2000 [7]	4 Subjects, 4 Sequences per subject	87.50 (*NN*), 93.75 (3*NN*)
Hayfron-Acquah 2001 [8]	4 Subjects, 4 Sequences per subject	100.0 (*NN*), 100.0 (3*NN*)
Foster 2001 [9]	6 Subjects, 4 Sequences per subject	83.00 (*NN*)
Our method	6 Subjects, 4 Sequences per subject	95.83 (*NN*), 87.50 (3*NN*), 100.0 (*ENN*)

TABLE III
COMPARISON OF SEVERAL RECENT ALGORITHMS ON THE *NLPR* DATABASE (0°)

Methods	Top 1	Top 5	Top 10
BenAbdelkader 2001 [10]	72.50%	88.75%	96.25%
BenAbdelkader 2002 [52]	82.50%	93.75%	100.0%
Collins 2002 [26]	71.25%	78.75%	87.50%
Lee 2002 [41]	87.50%	98.75%	100.0%
Phillips 2002 [47]	78.75%	91.25%	98.75%
Our method	88.75%	96.25%	100.0%

scan multiple days, multiple views, clothing variations, etc.

2) Clothes in different seasons will bring considerable effects on the shapes of moving people (e.g., loose vs. close-fitting). Our proposed method is based on the appearance shape just like others [2]–[4], [6], [7], [9], [10], [26], [36], [41], [45], [47], [52], so it is inevitably affected. In fact, except [5] and [12], previous work also resulted from the influences of shape owing to their direct use of motion segmentation information accordingly. Creating multiple reference sequences with different clothes is probably useful to solve this problem. Although we obtained very clean segmentation results on our data set, it should be mentioned that results might degrade significantly with video captured by a poor quality camera. Therefore, more robust motion detection algorithm needs to be developed for reliably handling low quality images.
3) The main drawback of the current method is that it is view-dependent, which is analogous to the state of the art of past algorithms except [20], [25], and [40]. There is no reason to expect that extracted features are invariant to viewing angle. As shown in our experiments, the same feature extraction with different viewing angles has different recognition ability. So a useful experiment would be to determine the sensitivity of the features to viewing angles. The results would enable a multicamera tracking system to select an optimal view for the purpose of recognition. Additionally, an obvious way to generalize the views is to store training sequences taken from multiple viewpoints, and classify both the subject and the viewpoint [26].
4) Both static and dynamic information derived from gait plays an important role in gait recognition. Our work has focused more on model-free recognition based on the static shapes. It may be more useful for recognition to extract dynamic information such as the oscillatory trajectories of joints or limbs as determined by model parameters though it is very hard to be well solved in computer vision. Therefore, 3D human body modeling and tracking might prove to be of benefit [49], [50]. The combination of static and dynamic information must be a promising direction, and this work is under development at our group.

Also, some efforts will be taken to select better similarity measure, design more powerful classifier, extract the extended gait features, etc.

VII. CONCLUSION

With the increasing demands of visual surveillance systems, human identification at a distance has recently gained more interest. Gait is a potential behavioral feature, and many allied studies have demonstrated that it can be used as a useful biometric feature for recognition. The development of computer vision techniques has also assured that vision-based automatic gait analysis can be gradually achieved.

This paper has described a novel 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 then represented as the associated configurations in the common coordinate system, and are analyzed using the *Procrustes* shape analysis method to obtain eigenshape signatures representing implicitly the structural shape cue of the walking figure's appearance. The standard pattern classification technique is utilized for recognition. Experimental results have demonstrated the effectiveness and advantages of the proposed algorithm.

ACKNOWLEDGMENT

The authors would like to thank Dr. M. Nixon from the University of Southampton, U.K., for providing us with the *SOTON* database. Our thanks also go to those who have contributed to the establishment of the *NLPR* database, and to the anonymous referees for their thorough review and constructive comments.

REFERENCES

[1] [Online]. Available: http://www.darpa.mil/iao/HID.html
[2] H. Murase and R. Sakai, "Moving object recognition in eigenspace representation: Gait analysis and lip reading," *Pattern Recognit. Lett.*, vol. 17, pp. 155–162, 1996.
[3] P. Huang, C. Harris, and M. S. Nixon, "Human gait recognition in canonical space using temporal templates," *Vis. Imag. Signal Process.*, vol. 146, no. 2, pp. 93–100, 1999.
[4] S. A. Niyogi and E. H. Adelson, "Analyzing and recognizing walking figures in XYT," in *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, 1994, pp. 469–474.
[5] D. Cunado, M. S. Nixon, and J. N. Carter, "Using gait as a biometric, via phase-weighted magnitude spectra," in *Proc. 1st Int. Conf. Audio- and Video-Based Biometric Person Authentication*, 1997, pp. 95–102.
[6] J. Little and J. Boyd, "Recognizing people by their gait: The shape of motion," *J. Comput. Vis. Res.*, vol. 1, no. 2, pp. 2–32, 1998.
[7] J. Shutler, M. Nixon, and C. Harris, "Statistical gait recognition via temporal moments," in *Proc. 4th IEEE Southwest Symp. Image Analysis and Interpretation*, 2000, pp. 291–295.
[8] J. B. Hayfron-Acquah, M. S. Nixon, and J. N. Carter, "Automatic gait recognition by symmetry analysis," in *Proc. 3rd Int. Conf. Audio- and Video-Based Biometric Person Authentication*, 2001, pp. 272–277.
[9] J. Foster, M. Nixon, and A. Prugel-Bennett, "New area based metrics for gait recognition," in *Proc. 3rd Int. Conf. Audio- and Video-Based Biometric Person Authentication*, 2001, pp. 312–317.
[10] C. BenAbdelkader, R. Culter, H. Nanda, and L. Davis, "EigenGait: Motion-based recognition of people using image self-similarity," in *Proc. 3rd Int. Conf. Audio- and Video-Based Biometric Person Authentication*, 2001, pp. 284–294.
[11] J. Boyd, "Video phase-locked loops in gait recognition," in *Proc. Int. Conf. Computer Vision*, 2001.
[12] R. Tanawongsuwan and A. Bobick, "Gait recognition from time-normalized joint-angle trajectories in the walking plane," in *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, 2001.
[13] G. Shakhnarovich, L. Lee, and T. Darrell, "Integrated face and gait recognition from multiple views," in *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, 2001.
[14] A. Bobick and A. Johnson, "Gait recognition using static, activity-specific parameters," in *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, 2001.
[15] Y. Yang and M. Levine, "The background primal sketch: An approach for tracking moving objects," *Mach. Vis. Applicat.*, vol. 5, pp. 17–34, 1992.
[16] Y. Kuno, T. Watanabe, Y. Shimosakoda, and S. Nakagawa, "Automated detection of human for visual surveillance system," in *Proc. Int. Conf. Pattern Recognition*, 1996, pp. 865–869.
[17] J. Kent, *New Directions in Shape Analysis, in the Art of Statistical Science: A Tribute to G. S. Watson*. New York: Wiley, 1992, pp. 115–127.
[18] K. Mardia and P. Jupp, *Directional Statistics*. New York: Wiley, 2000.
[19] S. V. Stevenage, M. S. Nixon, and K. Vince, "Visual analysis of gait as a cue to identity," *Appl. Cogn. Psychol.*, vol. 13, pp. 513–526, 1999.
[20] A. Johnson and A. Bobick, "A multi-view method for gait recognition using static body parameters," in *Proc. 3rd Int. Conf. Audio- and Video-Based Biometric Person Authentication*, 2001, pp. 301–311.
[21] M. S. Nixon, J. N. Carter, D. Cunado, P. S. Huang, and S. V. Stevenage, "Automatic gait recognition," in *BIOMETRICS Personal Identification in Networked Society*, A. K. Jain, Ed. Norwell, MA: Kluwer, 1999.
[22] L. Wang, W. M. Hu, and T. N. Tan, "A new attempt to gait-based human identification," in *Proc. Int. Conf. Pattern Recognition*, vol. I, 2002, pp. 115–118.
[23] D. M. Gavrila, "The visual analysis of human movement: A survey," *Comput. Vis. Image Understand.*, vol. 73, no. 1, pp. 82–98, 1999.
[24] C. Y. Yam, M. S. Nixon, and J. N. Carter, "Gait recognition by walking and running: A model-based approach," in *Proc. 5th Asian Conf. Computer Vision*, Melbourne, Australia, 2002, pp. 1–6.
[25] C. BenAbdelkader, R. Culter, and L. Davis, "Stride and cadence as a biometric in automatic person identification and verification," in *Proc. Int. Conf. Automatic Face and Gesture Recognition*, Washington, DC, 2002, pp. 372–376.
[26] R. Collins, R. Gross, and J. Shi, "Silhouette-based human identification from body shape and gait," in *Proc. Int. Conf. Automatic Face and Gesture Recognition*, Washington, DC, 2002, pp. 366–371.
[27] J. Phillips, H. Moon, S. Rizvi, and P. Rause, "The FERET evaluation methodology for face recognition algorithms," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, pp. 1090–1104, Oct. 2000.
[28] J. E. Cutting and L. Kozlowski, "Recognizing friends by their walk," *Bull. Psychonomic Soc.*, vol. 9, no. 5, pp. 353–356, 1977.
[29] G. Johansson, "Visual perception of biological motion and a model for its analysis," *Percept. Psychophys.*, vol. 14, no. 2, pp. 201–211, 1973.
[30] D. Cunado, M. S. Nixon, and J. N. Carter, "Automatic gait recognition via model-based evidence gathering," in *Proc. Workshop on Automatic Identification Advanced Technologies*, 1999, pp. 27–30.
[31] S. A. Niyogi and E. H. Adelson, "Analyzing gait with spatiotemporal surfaces," in *Proc. IEEE Workshop on Motion of Non-Rigid and Articulated Objects*, 1994, pp. 64–69.
[32] M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cogn. Neurosci.*, vol. 3, pp. 71–86, 1991.
[33] M. Köhle, D. Merkl, and J. Kastner, "Clinical gait analysis by neural networks: Issues and experiences," in *Proc. IEEE Symp. Computer-Based Medical Systems*, Maribor, Slovenia, 1997, pp. 138–143.
[34] D. Winter, *The Biomechanics and Motor Control of Human Gait: Normal, Elderly, and Pathological*. Waterloo, ON, Canada: Univ. of Waterloo Press, 1991.
[35] M. Murray, "Gait as a total pattern of movement," *Amer. J. Phys. Med.*, vol. 46, no. 1, pp. 290–332, 1967.
[36] J. Boyd and J. Little, "Phase in model-free perception of gait," in *Proc. IEEE Workshop Human Motion*, 2000.
[37] D. Meyer, J. Psl, and H. Niemann, "Gait classification with HMM's for trajectories of body parts extracted by mixture densities," in *Proc. Brit. Machine Vision Conf.*, 1998, pp. 459–468.
[38] A. Bissacco, A. Chiuso, Y. Ma, and S. Soatto, "Recognition of human gaits," in *Proc. Conf. Computer Vision and Pattern Recognition*, 2001.
[39] J.-C. Cheng and J. Moura, "Capture and representation of human walking in live video sequences," *IEEE Trans. Multimedia*, vol. 1, no. 2, pp. 144–156, 1999.
[40] S. Carlsson, "Recognizing walking people," in *Proc. Eur. Conf. Computer Vision*, 2000.
[41] L. Lee and W. Grimson, "Gait analysis for recognition and classification," in *Proc. Int. Conf. Automatic Face and Gesture Recognition*, Washington, DC, 2002, pp. 155–162.
[42] L. Kozlowski and J. Cutting, "Recognizing the sex of a walker from a dynamic point-light display," *Percept. Psychophys.*, vol. 21, pp. 575–580, 1977.
[43] P. S. Huang, C. J. Harris, and M. S. Nixon, "Comparing different template features for recognizing people by their gait," in *Proc. Brit. Machine Vision Conf.*, 1998.
[44] L. Wang, H. Ning, W. Hu, and T. Tan, "Automatic gait recognition based on *Procrustes* shape analysis," in *Proc. Int. Conf. Image Processing*, New York, 2002.
[45] Q. He and C. Debrunner, "Individual recognition from periodic activity using hidden Markov models," in *Proc. IEEE Workshop on Human Motion*, 2000.
[46] A. Jain, R. Bolle, and S. Pankanti, *Biometrics: Personal Identification in Networked Society*. Norwell, MA: Kluwer, 1999.
[47] P. J. Phillips, S. Sarkar, I. Robledo, P. Grother, and K. Bowyer, "Baseline results for the challenge problem of human ID using gait analysis," in *Proc. Int. Conf. Automatic Face and Gesture Recognition*, Washington, DC, 2002, pp. 137–142.
[48] C. Y. Yam, M. S. Nixon, and J. N. Carter, "On the relationship of human walking and running: Automatic person identification by gait," in *Proc. Int. Conf. Pattern Recognition*, 2002.
[49] S. Dockstader, K. Bergkessel, and A. Tekalp, "Feature extraction for the analysis of gait and human motion," in *Proc. Int. Conf. Pattern Recognition*, 2002, pp. 5–8.
[50] B. Bhanu and J. Han, "Individual recognition by kinematic-based gait analysis," in *Proc. Int. Conf. Pattern Recognition*, 2002, pp. 343–346.
[51] L. Wang, W. Hu, and T. N. Tan, "Recent developments in human motion analysis," *Pattern Recognit.*, vol. 36, no. 3, pp. 585–601, 2003.
[52] C. BenAbdelkader, R. Cutler, and L. Davis, "Motion-based recognition of people in eigengait space," in *Proc. Int. Conf. Automatic Face and Gesture Recognition*, Washington, DC, 2002.

Liang Wang received the B.Sc. degree in electrical engineering and the M.Sc. degree in video processing and multimedia communication from Anhui University, Hefei, China, in 1997 and 2000, respectively. He is currently pursuing the Ph.D. degree in pattern recognition and intelligent system in the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China.

He has published more than ten papers in major international journals and conferences. His current research interests include computer vision, pattern recognition, digital image processing and analysis, multimedia, etc.

Tieniu Tan (M'92–SM'97) received the B.Sc. degree in electronic engineering from Xi'an Jiaotong University, China, in 1984 and the M.Sc., D.I.C., and Ph.D. degrees in electronic engineering from Imperial College of Science, Technology, and Medicine, London, U.K., in 1986, 1986, and 1989, respectively.

He joined the Computational Vision Group at the Department of Computer Science, The University of Reading, Reading, U.K., in October 1989, where he worked as Research Fellow, Senior Research Fellow, and Lecturer. In January 1998, he returned to China to join the National Laboratory of Pattern Recognition, the Institute of Automation of the Chinese Academy of Sciences, Beijing, China. He is currently Professor and Director of the National Laboratory of Pattern Recognition as well as President of the Institute of Automation. He has published widely on image processing, computer vision and pattern recognition. His current research interests include speech and image processing, machine and computer vision, pattern recognition, multimedia, and robotics. He serves as referee for many major national and international journals and conferences. He is an Associate Editor for *Pattern Recognition* and the Asia Editor of *Image and Vision Computing*.

Dr. Tan is an Associate Editor of the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE. He is an elected member of the Executive Committee of the British Machine Vision Association and Society for Pattern Recognition (1996–1997) and is a founding co-chair of the IEEE International Workshop on Visual Surveillance.

Weiming Hu received the Ph.D. degree from the Department of Computer Science and Engineering, Zhejiang University, China.

He was a Postdoctoral Research Fellow at the Institute of Computer Science and Technology, Founder Research and Design Center, Peking University, China, from April 1998 to March 2000. In April 2000, he joined the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China, as an Associate Professor. He has published more than 40 papers on major national journals and international conferences. His current research interests include visual surveillance and monitoring of dynamic scenes, recognition and filtering of Internet objectionable images, neural network, computer vision, etc.

Huazhong Ning received the B.Sc. degree in 2000 in computer science from the Department of Special Class for Gifted Young of University of Science and Technology of China. He is currently a master candidate in the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China.

He has published several papers on major national journals and international conferences. His main research interests include computer vision, video computing, tracking, human computer interaction, image processing, pattern recognition, graphics, etc.