ORTHOGONAL DIAGONAL PROJECTIONS FOR GAIT RECOGNITION

Daoliang Tan, Kaiqi Huang, Shiqi Yu, Tieniu Tan

National Laboratory of Pattern Recognition
Institute of Automation, Chinese Academy of Sciences, Beijing, China, 100080

ABSTRACT

Gait has received much attention from researchers in the vision field due to its utility in walker identification. One of the key issues in gait recognition is how to extract discriminative shape features from 2D human silhouette images. This paper deals with the problem of gait-based walker recognition using statistical shape features. First, we normalize walkers' silhouettes (to facilitate gait feature comparison) into a square form and use the orthogonal projections in the positive and negative diagonal directions to draw personal signatures contained in gait patterns. Then principal component analysis (PCA) and linear discriminant analysis (LDA) are applied to reduce the dimensionality of original gait features and to improve the topological structure in the feature space. Finally, this paper accomplishes the recognition of unknown gait features based on the nearest neighbor rule, with the discussion of the effect of distance metrics and scales on discriminating performance. Experimental results justify the potential of our method.

Index Terms— Gait, PCA, LDA, shape, metric, scale

1. INTRODUCTION

Both governments and the public have paid great attention to the issue of security over the past years. In particular, the 9.11 incident deepens the consensus of security enhancement. In this context, a variety of biometrics have been investigated in an effort to better protect our society, including gait that refers to the person-specific moving styles. Gait-based human recognition is partially supported by the earlier psychological experiments [3].

In fact, many contributions have been made to this rapidly developing domain. For instance, Niyogi and Adelson [7] took the earliest initiative in recognizing walking people based on gait features. Additionally, Cunado et al. [2] used two inter-linked pendulums to model the motion of human legs and achieved the recognition of walking people using the dynamics of angles in the simplified leg model. Instead of following structure from motion and dynamic cues for walker identification, Bobick and Johnson [1] employed the shape features of four static distances in representative human silhouettes to coarsely describe human gait. Similarly, Wang et al. [9] characterized walking gait via a vector of distances obtained by unwrapping the silhouette around its center of "mass". Moreover, Kale et al. [4] utilized the outermost widths of silhouettes for the description of pedestrians, whereas Liu et al. [6] relied on frieze patterns to delineate the signatures of walking human beings. In particular, Sarkar et al. [5] built an important baseline platform for the easy comparison of daytime gait recognition algorithms, and Tan et al. [8] set up a large night gait dataset in an attempt to facilitate the research of walker recognition at night, considering the real significance of night security. In summary, most of state-of-the-art methods for gait recognition adopt the shape matching-like scheme with the help of appearance features in 2D human silhouettes, since shape features are more important than dynamic ones from the long-term point of view.

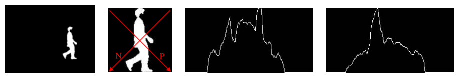
Therefore, it is critical to extract useful gait features when developing gait recognition systems. However, feature extraction involves a trial-and-error process and generally lacks theoretic guidelines. In addition, existing work usually neglects two vital issues associated with gait recognition: distance metrics and scales. This motivates us to address the problem of gait-based walker recognition using shape features with the consideration of metrics and scales. From the perspective of walker recognition from video, this work is necessary. The major contribution of this paper lies in the offering of a promising method to extract gait features.

In the following, we will introduce our method in Section 2. Then Section 3 provides experimental justification for our approach. Finally, Section 4 concludes this paper.

2. APPROACH

In addition to supplying human silhouette data, which can be obtained by background subtraction, current gait databases comply with the assumption that surveillance cameras are static or fixed and there exists only one single pedestrian in the scene of interest. Hence this paper will focus on silhouette postprocesing, feature description, dimension reduction, and classification. The nontrivial step of postprocessing serves as the alignment of human silhouette images for the subsequent matching purpose. Then the objective of feature description

This work is partially supported by the National Basic Research Program of China (2004CB318110) and the National Natural Science Foundation of China (60605014, 60332010, and 60335010).



(a) (b) (c) (d)

Fig. 1. Gait feature extraction. (a) Original silhouette. (b) Normalized silhouette. (c) Positive diagonal projection. (d) Negative diagonal projection.

is to avoid the unnecessary redundancy in the raw silhouette data, and this objective is further solidified by the step of dimension reduction rooted on the linear optimization principles. The following part will present more information about these steps.

2.1. Silhouette Postprocessing

In general, our method falls into the shape matching-like framework. As a result, it is better to normalize raw human silhouettes of different sizes in the hope of removing the inconsistency in recognition results due to the resolution and human position variation, regardless of the loss of personal physical features (e.g., human height). More specifically, we achieve the normalization of silhouettes by means of two steps: 1) calculate the aspect ratio θ of each human silhouette; and 2) resize the raw silhouette with θ preserved and translate the resized silhouette in an $s \times s$ image I_s to make its center of gravity be the middle of I_s in the horizontal direction. Figure 1(a)–(b) illustrate this process.

2.2. Feature Description

The choice of suitable gait features is the key to the success of a gait recognition system; this paper resorts to the use of appearance features to characterize human gait. More precisely, each normalized silhouette I_s is projected in two orthogonal directions: the positive diagonal P and the negative one N shown in Fig. 1(b). Thus, each projection records a vector of the valid number of human pixels in that projective direction. We denote the P projection by x_p and the N one by x_n. Finally, the concatenation $x = (x_p^T, x_n^T)^T \in \mathcal{R}^m$ of x_p and x_n fulfills the description of the gait pattern within I_s. In addition, Figure 1(c)–(d) describe our gait features. It is worth pointing out the difference between Liu's frieze pattern [6] and ours: we focus on the cues in two diagonal directions and collectively employ them rather than the separate horizontal and vertical projections.

2.3. Dimension Reduction

The component features in x certainly have correlation due to the symmetry of human figures; this makes it necessary to remove the correlation-induced redundancy in x. The fact that linear methods to reduce dimensionality generally have lower computational complexity and higher efficiency in comparison with nonlinear techniques prompts us to take advantage of the classical PCA and LDA techniques in an effort to further reduce data redundancy. In brief, PCA can be regarded as a criterion of minimizing information loss or reconstructive errors which can be mathematically expressed as

$$\min_{W \in \mathcal{R}^{m \times n_w}} E \parallel \hat{x} - x \parallel_2^2 \ s.t. \ W^T W = I \qquad (1)$$

where the transpose of W represents a linear transformation mapping x to $u = W^T x$, $\hat{x}$ the reconstructed signal obtained by $\hat{x} = Wu$ (with the columns of W being basis), and n_w is the expected dimension. It is easy to figure out that the columns of the optimal W not only should satisfy the eigen-equation of the covariance matrix S_x of x and but should also correspond to the first n_w largest eigenvalues of S_x. This paper takes the choice for n_w that the ideal n_w equals the minimum number of the eigenvalues of S_x which at least "weigh" a 95% of the total weight of x— the trace of S_x. On the other hand, PCA actually does not makes full use of the interrelationship among training data such as the inter-class and intra-class cues. This is the motivation behind the application of the LDA technique. The basic idea in LDA is to maximize the Euclidean distance between data in different classes (often simplified as the distance between distinct class centers) and minimize the Euclidean distance between the same class data. This aim can be concisely formulated as the optimization problem (2) in matrix terms:

$$\max_{V \in \mathcal{R}^{n_w \times n_v}} \frac{tr(V^T S_b V)}{tr(V^T S_w V)} \qquad (2)$$

where V is the needed transformation and S_b and S_w are the between-class and within-class scatter matrices, respectively. The matrix traces of S_b and S_w gauge the between-class distance and the within-class distance. A series of matrix computation reveals that the column vectors of the computed V constitute the generalized eigenvectors between S_b and S_w. In a similar way to the use of n_w, this paper takes account of a 98% of the generalized "weight"—the sum of the generalized eigenvalues—for the choice of n_v. In a word, we can obtain a relatively compact feature vector $y = V^T W^T x$ through the combination of W and V.

2.4. Classification

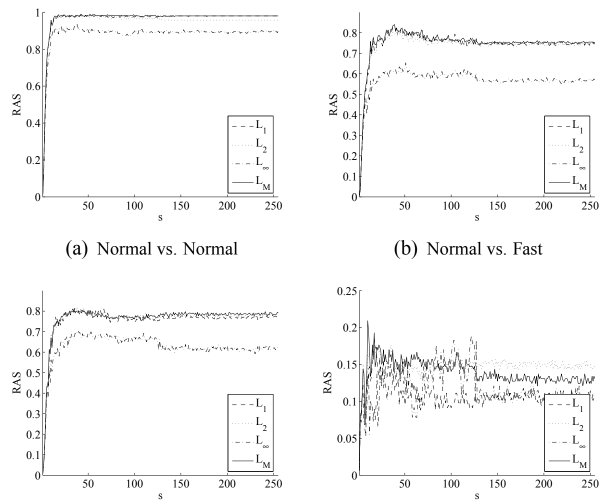
In order to lighten the computational load, this paper simplifies the description of gait features in one video sequence: we employ the average of the features to coarsely represent the signature of the person within this gait sequence. Although W and V are derived on the basis of the Euclidean-2 norm, it is worthwhile to investigate the effect of different distance metrics on recognition accuracy. In this work, we just consider four commonly used metrics: L_1, L_2, L_∞, and the Mahalanobis metric; this paper denotes the Mahalanobis metric by L_M for the notational simplicity. Finally, our method recognizes unknown gait features using the nearest neighbor decision-making rule.

Table 1. Seven Experiments on the USF-NIST Gait Database with the Gallery (G, A, R).

Exp.	Probe[1]	Difference
A	(G, A, L)[71]	View
B	(G, B, R)[41]	Shoe
C	(G, B, L)[41]	Shoe,View
D	(C, A, R)[70]	Surface
E	(C, B, R)[44]	Surface, Shoe
F	(C, A, L)[70]	Surface, View
G	(C, B, L)[44]	Surface, Shoe, View

Table 2. Recognition Accuracy Scores on the USF-NIST Gait Database

	Algo.	A	B	C	D	E	F	G
RAS	[5]	79%	66%	56%	29%	24%	30%	10%
	L_1	97%	78%	66%	23%	17%	17%	19%
	L_2	99%	80%	66%	24%	19%	18%	17%
	L_∞	82%	73%	46%	17%	14%	9%	10%
	L_M	100%	80%	66%	29%	24%	24%	24%



(a) Normal vs. Normal (b) Normal vs. Fast

(c) Normal vs. Slow (d) Normal vs. Bag

Fig. 3. The Accuracy-versus-Scale curves on the CASIA Night Dataset

3. EXPERIMENTS

We experiment with the approach in the last section on two gait databases for the scrutiny of its performance. The two databases are the USF-NIST Gait Database [5] and the CASIA Infrared Night Gait Dataset [8]. The use of these two datasets is because the USF-NIST database provides a baseline platform for an algorithmic comparison and the CASIA dataset has the largest number of subjects in the night gait aspect. In addition, we consider the variation of the normalizing scale s ranging from 1 to 255. The following will describe more details about these experiments.

3.1. USF-NIST Gait Database

We choose the use of the pre-supplied human silhouettes for the May-2001-No-Briefcase data, after considering the computational burden involved in experiments. This data collection consists of 74 people' gait patterns and includes three walking covariate factors: viewpoint, footwear, and ground surface. Meanwhile, Table 1 lists seven experiments designed by Sarkar et al. [5] with regard to this database.

Figure 2 depicts the curve of recognition accuracy scores (RAS) that vary with the normalization scale s. It can be seen from Fig. 2 that the accuracy of recognition at extremely small scales is low largely due to the limited, confusion-prone features, but the nonzero scores at small scales reflect that our gait description has certain discriminability; in general, the L_M metric has the best performance thanks to its second-order homoscedasticity and the L_∞ metric produces the lowest accuracy owing to its winner-take-all metric; the performance of L_1 and L_2 lies in between that of L_M and L_∞. In addition, the accuracy response to scale s in Fig. 2(a)–(c) tends to be stabilized because of the consistency in the appearance of training and testing silhouettes; the fluctuation of the accuracy in Fig. 2(d)–(g) results from the inconsistent segmentation noise in probe silhouettes. Furthermore, Table 2 presents the RAS values corresponding to the baseline algorithm [5] and $s = 32$. The results illustrate that our approach is promising.

[1]The value in the bracket indicates the number of subjects in the test.

3.2. CASIA Night Gait Dataset

The CASIA dataset keeps a record of night gait from 153 individuals and takes into account four cases: walk normally, walk fastly, walk slowly, and walk normally but with a bag. More precisely, each subject has ten sequences of gait samples: four sequences for the normal walking case and two sequences for each of the remaining cases. As far as this dataset is concerned, we use the collection of the first two normal-walking sequences of each individual as the training data and the remaining sequences as the testing data. Hence we perform four kinds of experiments: normal-versus-normal, normal-versus-fast, normal-versus-slow, and normal-versus-bag; this paper repeats the recognition experiment for each walking case two times, since there have two testing sequences for each case, and adopts the average of the two times's results to report the performance. Figure 3 displays the accuracy-versus-scale curves in the four experiments. Apart from a similar conclusion to that drawn from Fig. 2, we can also obtain from Fig. 3 that the L_2 metric can bring relatively more stable performance. The accuracy variation in Fig. 3(d) should be attributed to the bag-induced appearance noise. Furthermore, Figure 4 shows cumulative match scores (CMS) in the case of $s = 32$ for ranks up to 20 and justifies once again

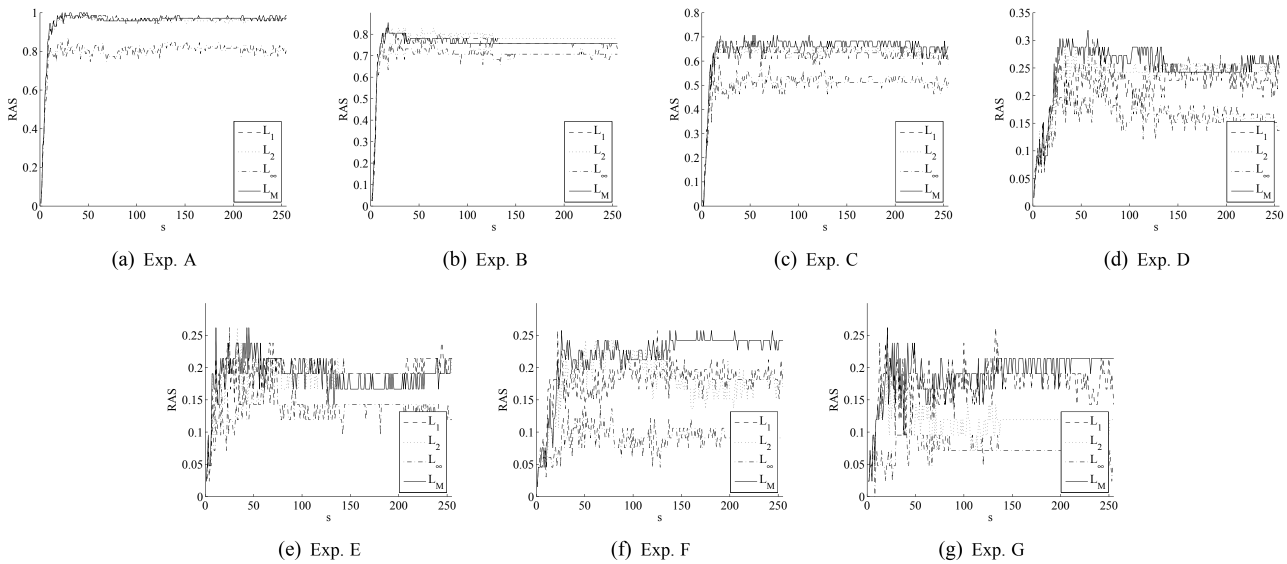


Fig. 2. The Accuracy-versus-Scale curves on the USF-NIST Gait Database

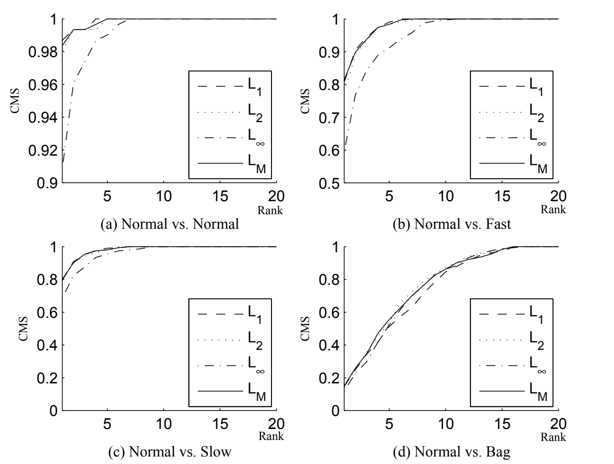


Fig. 4. The cumulative matching scores on the CASIA Night Gait Dataset

that the performance of our approach is promising, particularly when having no severe appearance changes.

4. CONCLUSION

This paper has addressed the problem of gait recognition based on appearance features in human silhouettes, with considering the issues of distance metrics and scales. Experimental results indicate that the Mahalanobis distance can produce the best recognition performance, that the increase in scales does not always brings the corresponding rise of recognition accuracy, and that the approach proposed in this work presents encouraging performance. Our major contribution lies in the offering of a promising method to extract gait features.

5. REFERENCES

[1] A. Bobick and A. Johnson, "Gait recognition using static, activity-specific parameters," in *Proc. CVPR*, December 2001, vol. 1, pp. 423–430.

[2] D. Cunado, M. Nixon, and J. Carter, "Automatic extraction and description of human gait model for recognition purposes," *CVIU*, vol. 90, no. 1, pp. 1–41, 2003.

[3] J. E. Cutting and L. T. Kozlowski, "Recognizing friends by their walk: Gait perception without familiarity cues," *Bulletin of the Psychonomic Society*, vol. 9, no. 5, pp. 353–356, 1977.

[4] A. Kale et al., "Identification of humans using gait," *IEEE TIP*, vol. 13, no. 9, pp. 1163–1173, 2004.

[5] S. Sarkar et al., "The human gait challenge problem: data sets, performance and analysis," *PAMI*, vol. 27, no. 2, pp. 162–177, 2005.

[6] Y. Liu, R. Collins, and Y. Tsin, "Gait sequence analysis using frieze patterns," in *Proc. ECCV*, May 2002, vol. 2, pp. 657–671.

[7] S. Niyogi and E. Adelson, "Analyzing and recognizing walking figures in xyt," in *Proc. CVPR*, June 1994, pp. 469–474.

[8] D. Tan, K. Huang, S. Yu, and T. Tan, "Efficient night gait recognition based on template matching," in *Proc. ICPR*, August 2006, vol. 3, pp. 1000–1003.

[9] L. Wang, T. Tan, H. Ning, and W. Hu, "Silhouette analysis-based gait recognition for human identification," *PAMI*, vol. 25, no. 12, pp. 1505–1518, 2003.