Ekinci M, Aykut M. Human gait recognition based on kernel PCA using projections. JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 22(6): 867~876 Nov. 2007

Human Gait Recognition Based on Kernel PCA Using Projections

Murat Ekinci and Murat Aykut

Computer Vision Lab, Department of Computer Engineering, Karadeniz Technical University, 61080 Trabzon, Turkey

E-mail: ekinci@ktu.edu.tr

Received February 17, 2007; Revised August 6, 2007.

Abstract This paper presents a novel approach for human identification at a distance using gait recognition. Recognition of a person from their gait is a biometric of increasing interest. The proposed work introduces a nonlinear machine learning method, kernel Principal Component Analysis (PCA), to extract gait features from silhouettes for individual recognition. Binarized silhouette of a motion object is first represented by four 1-D signals which are the basic image features called the distance vectors. Fourier transform is performed to achieve translation invariant for the gait patterns accumulated from silhouette sequences which are extracted from different circumstances. Kernel PCA is then used to extract higher order relations among the gait patterns for future recognition. A fusion strategy is finally executed to produce a final decision. The experiments are carried out on the CMU and the USF gait databases and presented based on the different training gait cycles.

Keywords biometrics, gait recognition, gait representation, kernel PCA, pattern recognition

1 Introduction

The image-based individual human identification methods, such as face, fingerprints, palmprints, generally require a cooperative subject, views from certain aspects, and physical contact or close proximity. These methods cannot reliably recognize non-cooperating individuals at a distance in the real world under changing environmental conditions. Gait, which concerns recognizing individuals by the way they walk, is a relatively new biometric without these disadvantages[1~3]. In other words, a unique advantage of gait as a biometric is that it offers potential for recognition at a distance or at low resolution when the human subject occupies too few image pixels for other biometrics to be perceivable.

Various gait recognition techniques have been proposed and can be broadly divided as model-based and model-free approaches. Model based approaches[4,5] aim to derive the movement of the torso and/or the legs. They usually recover explicit features describing gait dynamics, such as stride dimensions and the kinematics of joint angles.

Model-free approaches are mainly silhouette-based approaches. The silhouette approaches[3,6,7] characterize body movement by the statistics of the patterns produced by walking. These patterns capture both the static and dynamic properties of body shape. A hidden Markov models based framework for individual recognition by gait was presented in [6]. In [7], they first extract key frames from a sequence and then the similarity between two sequences is computed using the normalized correlation. The template matching method in [8] was extended to gait recognition by combining transformation based on canonical analysis and used eigenspace transformation for feature selection. In the work in [3], the similarity between the gallery sequence and the probe sequence is directly measured by computing the correlation corresponding time-normalized frame pairs. BenAbdelkader *et al.*[9] presented self similarity and structural stride parameters (stride and cadence) used PCA (Principal Component Analysis) applied to self-similarity plots that are derived by differencing. In [10], eigenspace transformation based on on PCA was first applied to the distance signals derived from a sequence of silhouette images, then classification was performed on gait patterns produced from the distance vectors. Han *et al.*[11] used the Gait Energy Image formed by averaging silhouettes and then deployed PCA and multiple discriminant analysis to learn features for fusion.

In this paper, we present an improved silhouette-based (model-free) approach and kernel PCA to extract the gait features. The main purposes and contributions of this paper are summarized as follows.

• An improved spatio-temporal gait representation, we called gait pattern, is first proposed to characterize human walking properties for individual recognition by gait. The gait pattern is created by using the distance vectors. The distance vectors are differences between

Regular Paper
This work was supported by Karadeniz Technical University under Grant No. KTU-2004.112.009.001.

the bounding box and silhouette, and are extracted by using four projections of silhouette.

• A Kernel Principal Component Analysis (PCA) based on a nonlinear extraction method is then applied. Kernel PCA is a state-of-the-art nonlinear machine learning method. Experimental results achieved by PCA and kernel PCA based methods are comparatively presented.

• Fourier transform is employed to achieve translation invariant for the gait patterns which are especially accumulated from silhouette sequences extracted from the subjects walk in different speed and/or different time. Consequently, Fourier transform based kernel PCA method is developed to achieve higher recognition for individuals in the database for the cases where the training and testing sets do not correspond to the same walking styles.

• A large number of papers in the literature reported their performance without using different training gait cycles. Here, we provide some quantitative comparative experiments to examine the performance of the proposed gait recognition algorithm with different numbers of training gait cycles of each person.

2 Gait Pattern Representation

In this paper, we consider individual recognition by activity-specific human motion, i.e., regular human walking, which is used in most current approaches of individual recognition by gait. We first represent the spatio-temporal information in a single 2D gait template (pattern) by using multi-projections of silhouette. We assume that silhouettes have been extracted from original human walking sequences. A silhouette preprocessing procedure[3,12] is then applied on the extracted silhouette sequences. It includes size normalization (proportionally resizing each silhouette image so that all silhouettes have the same height) and horizontal alignment (centering the upper half silhouette part with respect to its horizontal centroid). In a processed silhouette sequence, the process of period analysis of each gait sequence is performed as follows: once the person (silhouette) has been tracked for a certain number of frames, then we take the projections and find the correlation between consecutive frames, and do normalization by subtracting its mean and dividing by its standard deviation, and then smooth it with a symmetric average filter. In the symmetric filter used, the neighbor values of each center are inspected in symmetric pairs around the center. The average of them is determined as a value smoothed for the center cost[13]. Further we compute its autocorrelation to find peaks indicate the gait frequency (cycle) information. Hence, we estimate the real period as the average distance between each pair of consecutive major peaks[10,13].

2.1 Representation Construction

Gait pattern is produced from the projections of silhouettes which are generated from a sequence of binary silhouette images, $B_t(x, y)$, indexed spatially by pixel location (x, y) and temporally by time t. An example silhouette and the distance vectors corresponding to four projections are shown in Fig.1. The distance vectors (projections) are the differences be-

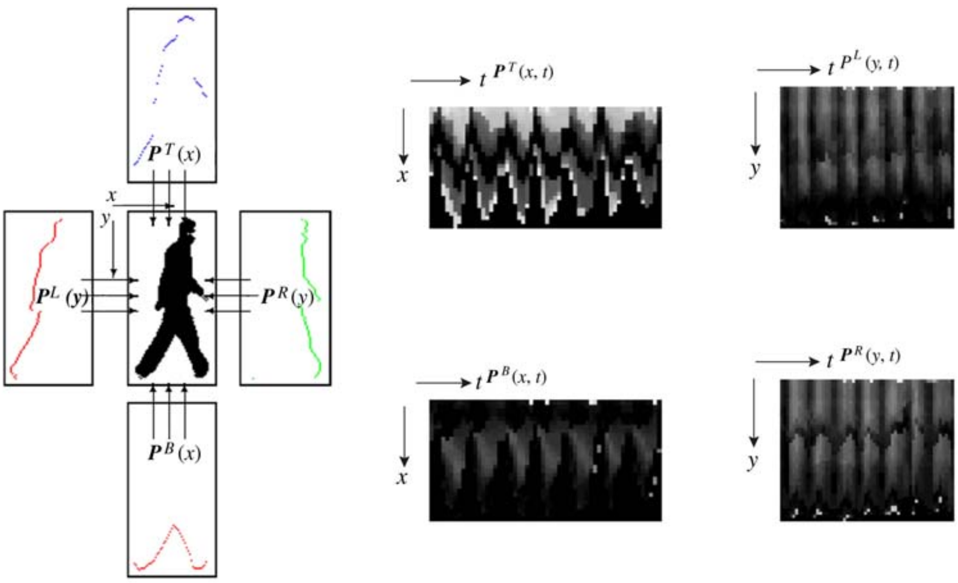

Fig.1. Silhouette representation. Left: Silhouette and four projections; Middle: gait patterns produced from top and bottom projections; Right: gait patterns obtained from left and right projections.

tween the bounding box and the outer contour of silhouette. There are four different image features called the distance vectors: top-, bottom-, left- and right-projections. The size of 1D signals for left- and right-projections is the height of the bounding box. The values in both signals are the number of columns between bounding box and silhouette at each row. The size of the 1D signals for both top- and bottom-distance vectors is the width of the bounding box, and the values of the signals are the number of rows between the box and silhouette at each column.

Thus, each gait pattern can separately be formed as a new 2D image. For instance, gait pattern image for top-projection is formulated as $\boldsymbol{P}^T(x,t) = \sum_y \overline{B_t}(x,y)$ where each column (indexed by time t) is the top-projections (row sum) of silhouette image $B_t(x,y)$, as shown in Fig.1 (Middle-Top). The meaning of $\overline{B_t}(x,y)$ is the complement of silhouette shape, that is empty pixels in the bounding box. Each value of $\boldsymbol{P}^T(x,t)$ is then a count of the number of rows empty pixels between the top side of the bounding box and the outer contours in that columns x of silhouette image $B_t(x,y)$. The result is a 2D pattern, formed by stacking row projections (from top of the bounding box to silhouette) together to form a spatio-temporal pattern. A second pattern which represents the bottom-projection $\boldsymbol{P}^B(x,t) = \sum_{-y} \overline{B_t}(x,y)$ can be constructed by stacking row projections (from bottom to silhouette), as shown in Fig.1 (Middle-Bottom). The third pattern $\boldsymbol{P}^L(y,t) = \sum_x \overline{B_t}(x,y)$ is then constructed by stacking columns projections (from left of the bounding box to silhouette) and the last pattern $\boldsymbol{P}^R(y,t) = \sum_{-x} \overline{B_t}(x,y)$ is also finally constructed by stacking columns projections (from right to silhouette), as shown in Fig.1 (Right), respectively. For simplicity of notation, we write $\sum_y$, $\sum_{-y}$, $\sum_x$, and $\sum_{-x}$ as shorthand for $\sum_{y=Top\text{-}of\text{-}the\text{-}box}^{Contour\text{-}of\text{-}silhouette}$, $\sum_{y=Bottom\text{-}of\text{-}the\text{-}box}^{Contour\text{-}of\text{-}silhouette}$, $\sum_{x=Left\text{-}side\text{-}of\text{-}the\text{-}box}^{Contour\text{-}of\text{-}silhouette}$, and $\sum_{x=Right\text{-}side\text{-}of\text{-}the\text{-}box}^{Contour\text{-}of\text{-}silhouette}$, respectively.

The variation of each component of the distance vectors can be regarded as gait signature of that object. From the temporal distance vector plots, it is clear that the distance vector is roughly periodic and gives the extent of movement of different part of the subject. The brighter a pixel in 2D patterns in Fig.1 (Middle and Right), the larger the value of the distance vector in that position.

3 Human Recognition Using Gait Patterns

In this section, we describe the proposed approach for gait-based human recognition. Binarized silhouettes are first produced by using motion segmentation which is achieved via background modeling using a dynamic background frame estimated and updated in time, for details see [14]. In the training procedure, each gallery (training) silhouette sequence is divided into cycles by gait cycle estimation. Training gait patterns are then computed from each cycle. Once gallery and probe silhouette sequences are obtained from the subjects walking in different speed and/or at different time, there can be translation variant in the gait patterns extracted from that sequences. To achieve a translation invariant approach, the 2D gait patterns are transformed to the frequency domain by applying fast Fourier transform (FFT). Next, we perform kernel PCA based nonlinear feature extraction procedure on the normalized gait patterns transformed into the frequency domain. As a result, training gait transformation matrices and training gait features that form feature databases are obtained. This is independently repeated for each gait pattern produced from the projections (left-, right-, top-, bottom-projections). In the recognition procedure, each probe (testing) silhouette sequence is processed to generate the gait patterns used as testing set. These patterns are then transformed to the feature space by transformation matrices to obtain gait pattern features. Testing features are compared with training features in the database. This is separately performed for the gait pattern features constructed by each projection. Finally a feature fusion strategy is applied to combine gait pattern features at the decision level to improve recognition performance. The system diagram is shown in Fig.2.

3.1 Kernel PCA

Kernel PCA is a technique for nonlinear dimension reduction of data with an underlying nonlinear spatial structure. A key insight behind kernel PCA is to transform the input data into a higher-dimensional feature space[15]. The feature space is constructed such that a nonlinear operation can be applied in the input space by applying a linear operation in the feature space. Consequently, standard PCA can be applied in feature space to perform nonlinear PCA in the input space.

Given k class for training, and each class represents a sequence of the distance vector signals of a person. Multiple sequences of each subject can be added for training, but we use a sequence, which includes one gait cycle. Let $\boldsymbol{P}_{i,j}^w$ be the j-th distance vector signal in the i-th class for w projection to silhouette and M_i the number of such distance vector signals in the i-th class. The total number of training samples is $M_t^w = M_1^w + M_2^w + \cdots + M_k^w$, as the whole training set can be represented by $[\boldsymbol{P}_{1,1}^w, \boldsymbol{P}_{1,2}^w, .., \boldsymbol{P}_{1,M_1}^w, \boldsymbol{P}_{2,1}^w, \ldots, \boldsymbol{P}_{k,M_k}^w]$. For ease of understanding, we denote the

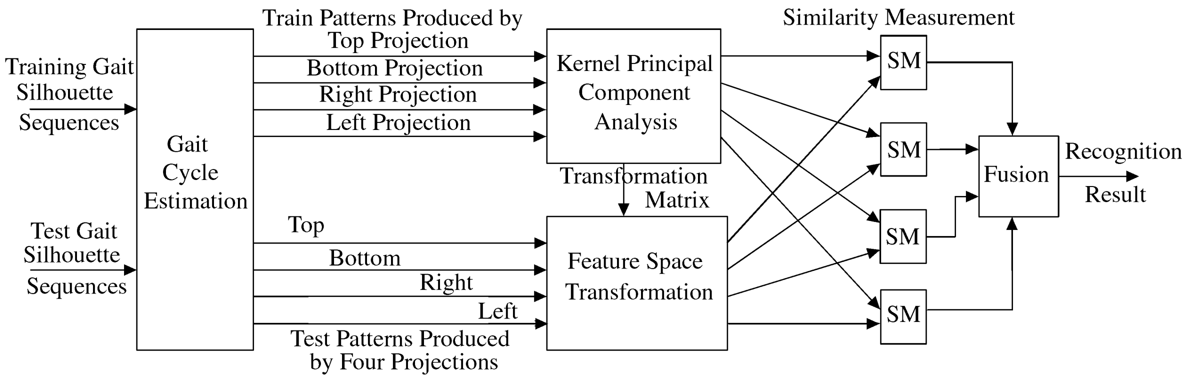

Fig.2. System diagram of human recognition using the proposed approach.

training samples, $\boldsymbol{P}_{i,j}^{w}$, as $\boldsymbol{\chi}_i \in \Re^N$, $i = 1,..,M$, where M is the total number of samples.

Thus, given a set of examples $\boldsymbol{\chi}_i \in \Re^N$, $i = 1,\ldots,M$, which are centered, $\sum_{i=1}^{M} \boldsymbol{\chi}_i = 0$, PCA finds the principal axis by diagonalizing the covariance matrix:

$$\boldsymbol{C} = \frac{1}{M}\sum_{i=1}^{M} \boldsymbol{\chi}_i \boldsymbol{\chi}_j^{\mathrm{T}}. \tag{1}$$

Eigenvalue equation, $\lambda \boldsymbol{v} = \boldsymbol{C}\boldsymbol{v}$ is solved, where $\boldsymbol{v}$ is eigenvector matrix. First few eigenvectors are used as the basic vectors of the lower dimensional subspace. Eigen features are then derived by projecting the examples onto these basic vectors[16].

In kernel PCA, the data, $\boldsymbol{\chi}$, from input space is first mapped to a higher dimensional feature space by using a map $\Phi : \Re^N \to F$, and then performing a linear PCA in F. The covariance matrix in this new space F is,

$$\overline{\boldsymbol{C}} = \frac{1}{M}\sum_{i=1}^{M} \Phi(\boldsymbol{\chi}_i)\Phi(\boldsymbol{\chi}_i)^{\mathrm{T}}. \tag{2}$$

Now the eigenvalue problem becomes $\lambda \boldsymbol{V} = \overline{\boldsymbol{C}}\boldsymbol{V}$. As mentioned previously we do not have to explicitly compute the nonlinear map Φ. The same goal can be achieved by using the kernel function $k(\boldsymbol{\chi}_i, \boldsymbol{\chi}_j) = (\Phi(\boldsymbol{\chi}_i) \cdot \Phi(\boldsymbol{\chi}_j))$, which implicitly computes the dot product of vectors $\boldsymbol{\chi}_i$ and $\boldsymbol{\chi}_j$ in the higher dimensional space[15]. The most often used kernel functions are Gaussian kernel, polynomial kernels, and sigmoid kernels[15]. Gaussian kernel was used for the experimentation in this work, and it is defined as,

$$k(\boldsymbol{\chi}_i, \boldsymbol{\chi}_j) = \exp\Big(-\frac{\|\boldsymbol{\chi}_i - \boldsymbol{\chi}_j\|^2}{2\sigma^2}\Big). \tag{3}$$

Pairwise similarity between input examples are captured in a matrix $\boldsymbol{K}$ which is also called Gram matrix. Each entry $\boldsymbol{K}_{i,j}$ of this matrix is calculated using kernel function $k(\boldsymbol{\chi}_i, \boldsymbol{\chi}_j)$. Eigenvalue equation in terms of Gram matrix written as (see [15]),

$$M\boldsymbol{A}\boldsymbol{\Lambda} = \boldsymbol{K}\boldsymbol{A}, \tag{4}$$

with $\boldsymbol{A} = (\alpha_1, \ldots, \alpha_M)$ and $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_M)$. $\boldsymbol{A}$ is an $M \times M$ orthogonal eigenvector matrix and $\boldsymbol{\Lambda}$ is a diagonal eigenvalue matrix with diagonal elements in decreasing order. Since the eigenvalue equation is solved for $\boldsymbol{A}$'s instead of eigenvectors $\boldsymbol{V}_i$ of kernel PCA, we will have to normalize $\boldsymbol{A}$ to ensure that eigenvalues of kernel PCA have unit norm in the feature space, therefore $\alpha_j = \alpha_j/\sqrt{\lambda_j}$. After normalization the eigenvector matrix, $\boldsymbol{V}$, of kernel PCA is computed as follows,

$$\boldsymbol{V} = \boldsymbol{D}\boldsymbol{A} \tag{5}$$

where $\boldsymbol{D} = [\Phi(\boldsymbol{\chi}_1)\Phi(\boldsymbol{\chi}_2)\ldots\Phi(\boldsymbol{\chi}_M)]$ is the data matrix in feature space. Now let $\boldsymbol{\chi}$ be a test example whose map in the higher dimensional feature space is $\Phi(\boldsymbol{\chi})$. The kernel PCA features for this example are derived as follows:

$$F = \boldsymbol{V}^T\Phi(\boldsymbol{\chi}) = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}, \tag{6}$$

where $\boldsymbol{B} = [\Phi(\boldsymbol{\chi}_1)\Phi(\boldsymbol{x})\Phi(\boldsymbol{\chi}_2)\Phi(\boldsymbol{x})\ldots\Phi(\boldsymbol{\chi}_M)\Phi(\boldsymbol{x})]^{\mathrm{T}}$.

3.2 Similarity Measurement

Weighted Euclidean Distance (WED) measuring has initially been selected for classification[17], and is defined as follows:

$$WED : d_k = \sum_{i=1}^{N} \frac{(\boldsymbol{f}(i) - \boldsymbol{f}_k(i))^2}{(\boldsymbol{s}_k)^2} \tag{7}$$

where $\boldsymbol{f}$ is the feature vector of the unknown gait pattern, $\boldsymbol{f}_k$ and $\boldsymbol{s}_k$ denote the k-th feature vector and its standard deviation, and N is the feature length. In order to increase the recognition performance, a fusion task is developed for the classification results given by each projections.

3.3 Fusion

Two different strategies were developed. In Strategy 1, each projection is separately treated. Then the strategy is to combine the distances of each projection at the end by assigning equal weight. The final similarity using Strategy 1 is calculated as follows:

$$D_i = \sum_{j=1}^{4} w_j \times d_{ji} \tag{8}$$

where D_i is the fused distance similarity value, j is the algorithm's index for projection, w is normalized weight, d_i is single projection distance similarity value, and 4 is the number of projections (left, right, top, bottom). In conclusion, if any 2 of the distance similarity values in the 4 projections give maximum similarities for the same person, then the identification is determined as to be positive. Therefore, fusion strategy 1 has rapidly increased the recognition performance in the experiments.

In the experimental studies, it is seen that some projections can give more robust results than others. For example, while a human moves in the lateral view, with respect to image plane, the back side of the human gives more individual characteristics of gait. The projection corresponding to that side can give more reliable results, and in such a case, it is called the dominant feature. As a result, Strategy 2 has also been developed to further increase recognition performance. In Strategy 2, if the dominant projection, or at least 2 projections of others are positive for an individual, then the final identification decision is positive. The dominant feature in this work is automatically assigned by estimating the direction of motion objects under tracking[12].

4 Experiments and Results

We evaluate the performance of the method on CMU's MoBo database[18], and USF database[3].

4.1 CMU Database

This database has 25 subjects (23 males, 2 females) walking on a treadmill. Each subject is recorded performing four different types of walking: slow walk, fast walk, inclined walk, and slow walk holding ball. There are about 8 cycles in each sequence, and each sequence is recorded at 30 frames per second. It also contains six simultaneous motion sequences of 25 subjects, as shown in Fig.3.

We did mainly two different types of experiments on this database: in type I, all subjects in the training and testing sets walk on the treadmill at the same walking type; in type II, all subjects walk on the treadmill at two different walking types, and it is called that fast walk and slow walk. We did two experiments for each type investigation. They are: I.1) train on fast walk and test on fast walk; I.2) train on slow walk and test on slow walk. Type II: II.1) train on slow walk and test on fast walk; II.2) train on fast walk and test on slow walk.

The experimental results for Type I are summarized in Table 1. This table also shows the performance of the proposed work with the increase in the number of training gait cycles of each person. In cases I.1) and I.2), we conducted seven tests which used 25, 50, 75, 100, 125, 150 and 175 templates corresponding to one, two, three, four, five, six and seven gait cycles from each training sequence for training. The remainder gait cycles were used for authentication, respectively. There are totally 200 gait cycles, 8 cycles for each person. PCA-based method was employed to extract the features from gait patterns, and then the WED based nearest neighbor (NN) classifier was used for the similarity measurement. The fusion was finally performed to achieve the final decision. The comparison of recognition performance is also shown in Table 1. The experimental results reveal that the recognition rate is increased when the more gait cycles are used as training set. We did not need to apply kernel PCA-based feature extraction for Type I experiment, because PCA-based method had achieved the high

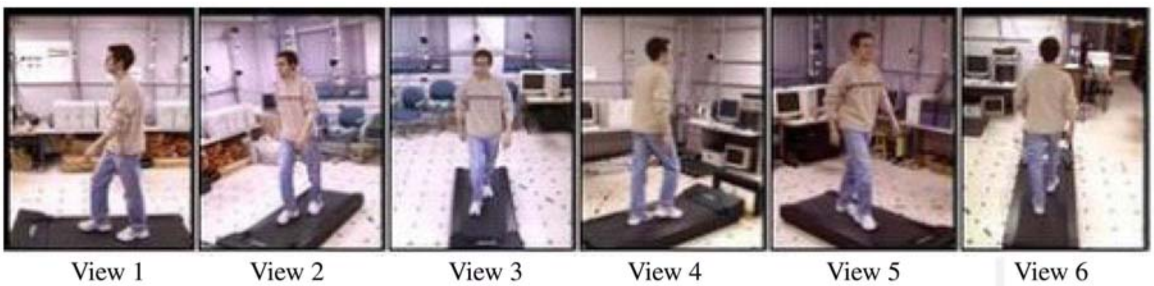

Fig.3. Six viewpoints in CMU database.

recognition rate (100%). The recognition performance was improved by increasing the number of training samples, but the performance was temporarily becoming a bit worse then increased for some view points. This can be due to the fact that some strong noisy in the gait patterns may heavily change the distribution of the data in the feature space. However, the more training samples are considered, the better data distribution in the feature space will significantly be achieved.

The next experiment, called Type II, was also done for the gait sequences extracted from the subjects walk on the treadmill with different speeds. It is called as slow walk and fast walk. For the case of training with fast walk and testing on slow walk, and vice versa, the dip in performance is caused due to the fact that for some individuals as biometrics suggests, there is a considerable change in body dynamics and stride length as a person changes his speed. The results for Type II experiments are also summarized in Table 2. Table 2 shows experimental results obtained by different feature extraction methods presented in this paper. In this table, rank1 performance means the percentage of the correct subjects appearing in the first place of the retrieved rank list and rank5 means the percentage of the correct subjects appearing in any of the first five places of the retrieved rank list. The performance in this table is the recognition rate under these two definitions.

For each person, there are 8 gait cycles at the slow walking and fast walking data sets in each viewpoint. The 8 cycles in one walking type are used as train set, the 8 cycles in other walking type are used as test set. The gait patterns are produced as explained in Subsection 2.1. The features in the gait patterns are extracted by using four different features extraction methods given in Table 2. When it is considered, it is seen that kernel PCA-based feature extraction gives better performance than PCA-based method. There is a quite possible translation variant problem between two gait patterns extracted from the subjects walk with different walking styles and/or at different times. To achieve translation invariant for the proposed method, the gait pattern in the spatial domain is first transformed to the spectral domain by using one dimensional (1-D) FFT. 1-D FFT process is indepen-

Table 1. Classification Performance with Training Number of Gait Cycles for the CMU Dataset for Different Viewpoints

Experiment	Number of Gait Pattern Used		CMU Database View Points				
Type I	Train Set	Test Set	View 1	View 3	View 4	View 5	View 6
I.1	1 gait cycle	7 gait cycles	97.7	99.4	98.8	100	98.8
	2 gait cycles	6 gait cycles	100	100	100	99.3	98.6
	3 gait cycles	5 gait cycles	99.2	99.2	100	100	99.2
	4 gait cycles	4 gait cycles	99	100	100	100	99
	5 gait cycles	3 gait cycles	100	100	100	100	98.6
	6 gait cycles	2 gait cycles	100	100	100	100	100
	7 gait cycles	1 gait cycle	100	100	100	100	100
I.2	1 gait cycle	7 gait cycles	97.7	90.8	97.1	100	98.2
	2 gait cycles	6 gait cycles	98	93.3	98	100	98
	3 gait cycles	5 gait cycles	97.6	94.4	98.4	100	99.2
	4 gait cycles	4 gait cycles	97	97	100	100	99
	5 gait cycles	3 gait cycles	97.3	98.6	100	100	98.6
	6 gait cycles	2 gait cycles	100	100	100	100	100
	7 gait cycles	1 gait cycle	100	100	100	100	100

Table 2. Experiments for Two Different Walking Styles with Different Viewpoints

Train	Method	Rank1 Performance					Rank5 Performance				
Test		View 1	View 3	View4	View5	View6	View1	View3	View4	View5	View6
Slow Fast	PCA	31.5	44	27	29	46	46	64.5	58.5	44	64.5
	KPCA	33	46.5	34.5	35	48	54	68.5	60.5	54	63.5
	FFT+PCA	65	80	63	64.5	67	89	91.5	91	87	87.5
	FFT+KPCA	73	76.5	71.5	64	76	89	92.5	94	89	91.5
Fast Slow	PCA	27	52	28	26	49	50.5	68.5	67.5	47.5	65
	KPCA	39.5	53.5	31.5	24.5	49	62	69	59	51	65
	FFT+PCA	61.5	74.5	62.5	64	73.5	85	88	90.5	85	88
	FFT+KPCA	66.5	79.5	61	67	74	89.5	91.5	89.5	90	88.5

Note: Each walking style includes 8 gait cycles.

Table 3. Comparison of Several Algorithms on MoBo Dataset

Train Test	Slow Slow		Fast Fast		Slow Fast		Fast Slow	
	View 1	View 3	View 1	View 3	View 1	View 3	View 1	View 3
Proposed Method	100	100	100	100	73	76.5	66.5	79.5
BenAbdelkader *et al.*[9]	100	96	100	100	54	43	32	33
UMD[6,19,20]	72	-	70	-	32	-	58	-
UMD[4]	72	-	76	-	12	-	12	-
CMU[7]	100	-	-	-	76	-	-	-
Baseline[3]	92	-	-	-	72	-	-	-
MIT[21]	100	-	-	-	64	-	-	-

Fig.4. Some sample images in the database described in [3, 22].

dently performed in horizontal or vertical directions for the gait patterns produced from both the left and right-projections or for the gait patterns produced from both the top- and bottom-projections, respectively. Then PCA- and kernel PCA-based feature extraction methods are employed to achieve higher recognition rates, as illustrated in Table 2. Consequently, highest recognition rates for most view points were achieved by using FFT+KPCA based feature extraction method.

Table 3 compares the recognition performance of different published approaches on MoBo database. Several papers have published results on this dataset, hence, it is a good experiment dataset to benchmark the performance of the proposed algorithm. Table 3 lists the reported identification rates for seven algorithms on eight commonly reported experiments. The first row lists the performance of the proposed method. For seven experiments the performance of the proposed algorithm has always highest score. The numbers given in Table 3 are read from graphs and tables in the cited papers. The number of the subjects in the training set and test set is 25. In the experiments for training on fast walk and testing on slow walk, or vice versa, 200 gait patterns (25 persons $\times$ 8 gait cycles) in each dataset were used to represent the performance of the proposed method. It can be seen from Table 3 that the right person in the top first matches 100% of times for the cases where training and testing sets correspond to the same walking styles. The proposed algorithm has achieved significantly better results than the other approaches on experiments which training and testing samples were extracted from different walking styles.

Table 4. Classification Performance for the USF Dataset, Version 1.7

Experi-ments	PCA		Kernel PCA	
	Strategy 1	Strategy 2	Strategy 1	Strategy 2
C,A,L (71)	78.8%	85.9%	84.5%	90.1%
C,A,R (71)	85.9%	88.7%	85.9%	87.3%
C,B,L (43)	74.4%	86.04%	81.3%	90.6%
C,B,R (43)	83.7%	93.02%	79.06%	88.3%
G,A,L (68)	86.7%	92.6%	88.2%	92.6%
G,A,R (68)	79.4%	82.3%	80.8%	85.2%
G,B,L (44)	90.9%	93.1%	93.1%	95.4%
G,B,R (44)	77.2%	86.3%	86.3%	90.9%

Note: The number of subjects in each subset is given in parenthesis.

4.2 USF Database

The USF database[3] is finally considered. This database consists of persons walking in elliptical paths in front of the camera. Some samples are shown in Fig.4. For each person, there are up to five covariates: viewpoints (left/right), two different shoe types, surface types (grass/concrete), carrying conditions (with/without a briefcase), and time and clothing. Eight experiments are designed for individual recognition as shown in Table 4. Sarkar *et al.*[3] propose a baseline approach to extract human silhouette and recognize an individual in this database.

The experiments in this section begin with these extracted binary silhouette data. These data are noisy, e.g., missing of body parts, small holes inside the obje-

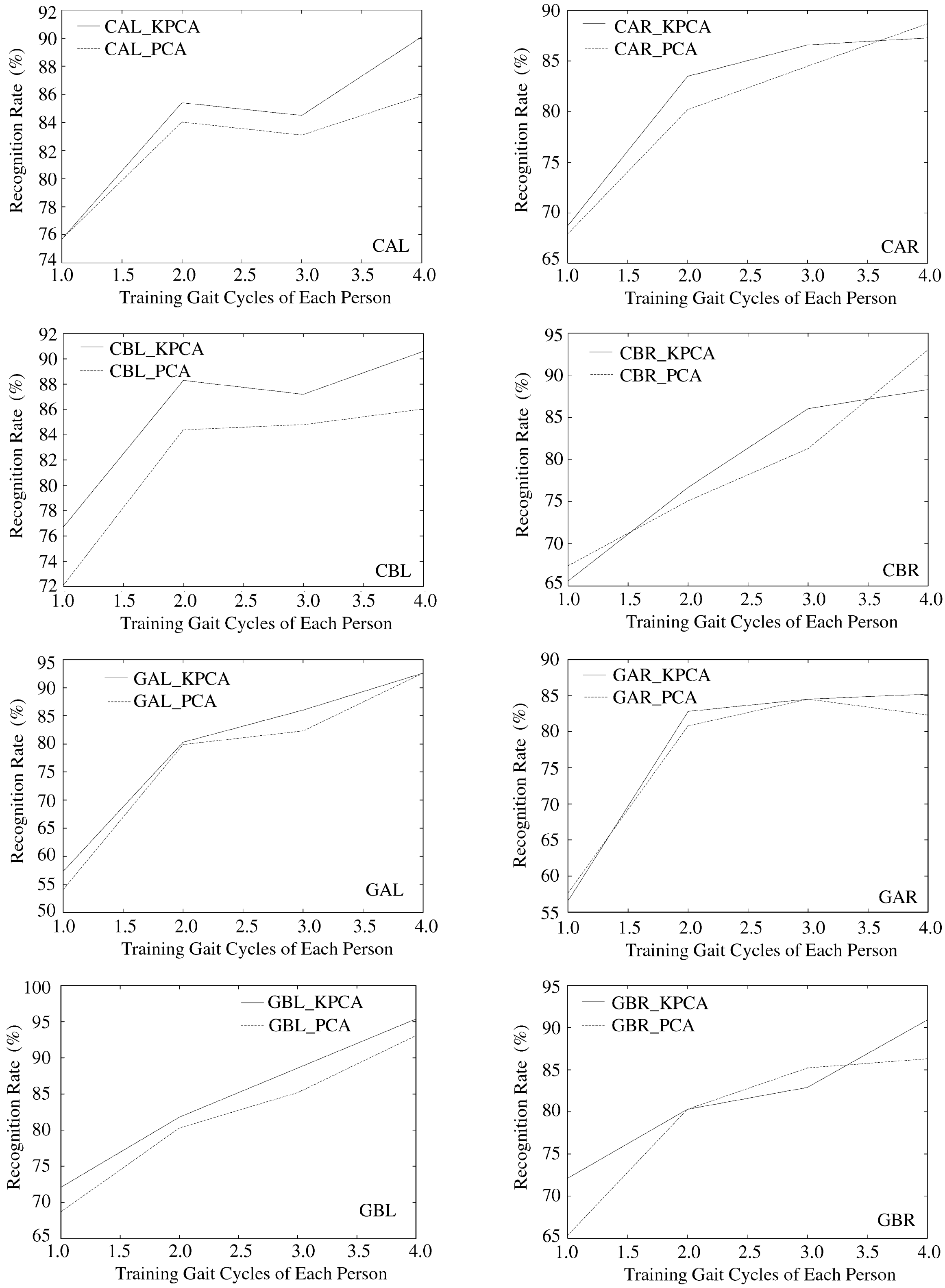

Fig.5. Illustration of the recognition performance variation with different training gait cycles of each person.

Table 5. Comparison of Recognition Performance Using Different Approaches on USF Silhouette Sequence Version 1.7

Experiments	Proposed Method	Baseline[22]	NLPR[10]	UMD-Indirect[6]	UMD-Direct[6]	GEI[11]
G,A,L (68)	92.6	79	70.42	91	99	100
G,B,R (44)	90.9	66	58.54	76	89	90
G,B,L (44)	95.4	56	51.22	65	78	85
C,A,R (71)	87.3	29	34.33	25	35	47
C,B,R (43)	88.3	24	21.43	29	29	57
C,A,L (69)	90.1	30	27.27	24	18	32
C,B,L (43)	90.6	10	14.29	15	24	31

cts, severe shadow around feet, and missing and adding some parts around the border of silhouettes due to background characteristics. In Table 4, G and C indicate grass and concrete surfaces, A and B indicate shoe types, and L and R indicate left and right cameras, respectively. The number of subjects in each subset is also given in brackets. Each one also includes 4∼5 gait cycle sequences. The experimental results on the standard USF HumanID Gait database version 1.7 are summarized in Table 4. In this table, the performance of PCA- and kernel PCA-based feature extraction methods are comparatively illustrated.

The matching is also conducted independently based on weighted Euclidean distance classifier. The decision results based on the fusion strategies, explained in Subsection 3.3, are additionally given in Table 4. Fusion 1 and Fusion 2 indicate that the results are produced by using Strategy I and Strategy II, respectively.

To analyze the relationship between the performance of the proposed method and number of training gait cycles of each person, four kinds of experimental types were designed: one (two, three, or four) training gait cycle(s) of each person was randomly selected for training, and the other gait cycles were used for authentication, respectively. These experimental results are given in Fig.5. Kernel PCA- and PCA-based features extraction methods are comparatively illustrated, as well. In Fig.5, y-axis indicates recognition rate (%), and x-axis indicates the number of training gait cycles of each person. When the plotted results in Fig.5 are considered, it can be seen that kernel PCA-based feature extraction approach achieves better performance than PCA-based approach. From the results we can report that the accuracy can be greatly improved with the growth of the training gait cycles. For instance, when the proposed algorithm is trained using 1 gait cycle in the experiment GBL, an accuracy of 72.1% is achieved. When 4 gait cycles are used for training, a higher accuracy of 95.4% can be gotten. It is evident that training gait cycle number can play an important role in the matching process. More training gait cycles lead to a high recognition rate.

Table 5 finally compares the recognition performance of different published approaches on the USF silhouette version 1.7. The performance of the proposed algorithm is better than other approaches in GBR, GBL, CAR, CBR, CAL, and CBL, and slightly worse in GAL.

5 Conclusions

In this paper, we first propose to improve the spatio-temporal gait representation, which is multi-projections of silhouettes developed by our previous work[13], for individual recognition by gait. As the other contributions and novelties in this paper, 1) kernel PCA based features extraction approach for gait recognition was then presented, 2) FFT-based preprocessing was also proposed to achieve translation invariant for the gait patterns which are produced from silhouette sequences extracted from the subjects walk in different walking styles, and 3) the experimental results were finally submitted to examine the performance of the proposed algorithm with different training gait cycles. The proposed approach achieves highly competitive performance with respect to the published major gait recognition approaches.

Acknowledgements The authors would like to thank the anonymous reviewers for their valuable comments, Dr. R. T. Collins from Carnegie Mellon University, USA, for providing us with the CMU database, Dr. S. Sarkar from the University of South Florida, USA, for providing us with the USF database, and Eyup Gedikli for helping to obtain gait patterns from the gait databases.

References

[1] Nixon M S, Carter J N. Automatic recognition by gait. *Proceeding of the IEEE*, 2006, 94(11): 2013∼2023.

[2] Veres G V, Gordon L, Carter J N, Nixon M S. What image information is important in silhouette-based gait recognition? In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Washington DC, USA, June 27∼July 2, 2004, 2: 776∼782.

[3] Sarkar S *et al.* The HumanID gait challenge problem: Data sets, performance, and analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(2): 162∼177.

[4] BenAbdelkader C, Cutler R, Davis L. Motion-based recognition of people in eigengait space. In *Proc. International Conference on Automatic Face and Gesture Recognition*, Washington DC, USA, May 20∼21, 2002, pp.254∼259.

[5] Bazin A I, Nixon M S. Gait verification using probabilistic methods. In *Proc. IEEE Workshop on Applications of Computer Vision*, Breckenridge, CO, USA, January 5∼7, 2005, pp.60∼65.

[6] Kale A *et al.* Identification of humans using gait. *IEEE Transactions on Image Processing*, 2004, 13(9): 1163∼1173.

[7] Collins R, Gross R, Shi J. Silhouette-based human identification from body shape and gait. In *Proc. International Conference on Automatic Face and Gesture Recognition*, Washington DC, USA, May 20∼21, 2002, pp.351∼356.

[8] Huang P, Harris C, Nixon M. Human gait recognition in canonical space using temporal templates. *Vision, Image and Signal Processing, IEE Proceedings*, April 1999, 146(2): 93∼100.

[9] BenAbdelkader C, Cutler R G, Davis L S. Gait recognition using image self-similarity. *EURASIP Journal of Applied Signal Processing*, 2004, 4: 1∼14.

[10] Wang L, Tan T, Ning H, Hu W. Silhouette analysis-based gait recognition for human identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(12): 1505∼1518.

[11] Han J, Bhanu B. Individual recognition using gait image energy. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 2006, 28(2): 316∼322.

[12] Ekinci M, Gedikli E. Silhouette-based human motion detection and analysis for real-time automated video surveillance. *Turkish Journal of Electrical Engineering and Computer Sciences,* 2005, 13(2): 199∼230.

[13] Ekinci M. Gait recognition using multiple projections. In *Proceeding of IEEE International Conference on Automatic Face and Gesture Recognition*, Southampton, U.K., April 10∼12, 2006, pp.517∼522.

[14] Ekinci M, Gedikli E. Background estimation based people detection and tracking for video surveillance. *Lecture Notes in Computer Science,* 2003, 2869: 421∼429.

[15] Scholkopf B, Somala A. Learning with Kernel: Support Vector Machine, Regularization, Optimization and Beyond. MIT Press, 2002.

[16] Ali S, Shah M. A supervised learning framework for generic object detection. In *IEEE, Proc. International Conference on Computer Vision,* Beijing, China, Oct. 17∼20, 2005, 2: 1347∼1354.

[17] Zhu Y, Tan T. Biometric personal identification based on handwriting. *Pattern Recognition*, 2000, 2: 797∼800.

[18] Gross R, Shi J. The CMU motion of body (MOBO) database. Technical Report CMU-RI-TR-01-18, Robotics Institute, Carnegie Mellon University, June 2001.

[19] Kale A, Cuntoor N, Chellappa R. A framework for activity specific human identification. In *Proc. International Conference on Acoustics, Speech, and Signal Processing,* Orlando, Florida, USA, May 13∼17, 2002, 4: 3660∼3663.

[20] Kale A, Rajagopalan A, Cuntoor N, Krugger V. Gait-based recognition of humans using continuous HMMMs. In *Proc. IEEE International Conference on Automatic Face and Gesture Recognition,* Washington DC, USA, May 20∼21, 2002, pp.321∼326.

[21] Lee L, Grimson W. Gait analysis for recognition and classification. In *Proc. IEEE, International Conference on Automatic Face and Gesture Recognition,* Washington DC, USA, May 20–21, 2002, pp.155∼162.

[22] Phillips P *et al.* Baseline results for challenge problem of Human ID using gait analysis. In *Proc. IEEE International Conference on Automatic Face and Gesture Recognition*, Washington DC, USA, May 20–21, 2002, pp.130∼135.

Murat Ekinci received the B.Sc. and M.Sc. degrees in electronic engineering from Karadeniz Technical University, Trabzon, Turkey, in 1990 and 1993, respectively. He received his Ph.D. degree in computer science from the University of Bristol, U.K. in 1997. Since 1997, he has been with the Computer Engineering Department at Karadeniz Technical University, Trabzon, where he is currently an assistant professor. His research interests include computer-vision based gait biometrics, pattern recognition, video surveillance, machine learning, human motion analysis, image processing and computer vision. He is a member of the IEEE and IEEE Computer Society.

Murat Aykut received the B.Sc. degree (with high honors) in computer engineering in 2005 from Karadeniz Technical University, Turkey where he is pursuing the M.Sc. degree. He is currently a research assistant in the Department of Computer Engineering, Karadeniz Technical University. His research interests include biometrics, machine learning, numerical analysis and computer vision.