

Learning Speed Invariant Gait Template via Thin Plate Spline Kernel Manifold Fitting

Sheng Huang¹
huangsheng@cqu.edu.cn

Ahmed Elgammal²
elgammal@cs.rutgers.edu

Dan Yang¹
dyang@cqu.edu.cn

¹ College of Computer Science
Chongqing University
Chongqing, China

² Department of Computer Science
Rutgers, The State University of New
Jersey,
Piscataway, NJ, USA

Abstract

We present a novel approach for cross-speed gait recognition. In our approach, the cyclic walking action is considered as residing on a manifold which is homeomorphic to a unit circle in the gait space. Thin Plate Spline (TPS) kernel-based Radial Basis Function (RBF) interpolation is used to fit the walking manifold for each gait sequence. The subject related kernel mapping coefficients are learned for representing the gait. According to the property of TPS, the coefficients can be naturally separated as an affine component and a non-affine component. The affine component is the style factor corresponding to the deformation of the homeomorphic manifold caused by the walking action, while the non-affine component is the shape factor, invariant to the walking speed. We denote this non-affine component as *Speed Invariant Gait Template* (SIGT) and use it as cross-speed gait feature. To address the curse of dimensionality issue and speed up the recognition, we use Globality Locality Preserving Projections (GLPP) to reduce the dimensions of SIGTs. Two walking speeds related gait databases are employed for evaluating our proposed method. The experimental results demonstrate the superiority of our method over the state-of-the-art.

1 Introduction

In the recent decade, more and more researchers pay attention to gait recognition, since the gait is the only perceptible biometric trait at a distance, and can tolerate low resolution imagery that other biometrics cannot be perceived [18]. However, compared to other biometrics, gait recognition is sensitive to more factors such as walking speed, clothing, briefcases, and even the walking environment. Among these factors, walking speed is regarded as one of the most common changing factors in real life. Although, several works have been proposed to tackle this problem from different perspectives [10, 16, 21, 24], but this issue still remains unsolved.

Generally speaking, gait recognition approaches can be roughly classified to model-based approaches and model-free approaches (feature/appearance-based approaches) [18]. Model-based approaches try to model the walking action using static or dynamic body parameters [1, 24, 26]. Model-free approaches focus on directly extracting holistic gait features



Figure 1: The visualization of SIGTs (a) The sparsely sampled gait silhouettes of half cycle from OU-ISIR Treadmill Dataset A, (b) The extracted SIGTs from the gait sequences.

from gait sequences. Compared to the model-free approaches, model-based approaches seem to be more popular recently [6, 11, 12, 13, 15, 16, 17, 22, 23] because of its simplicity and robustness.

The solutions for cross-speed gait recognition also follow the previous two paths. A recent representative model-based cross-speed gait recognition method is the speed variation based silhouette transformation model [24]. This method separates the gait silhouettes into dynamic and static features via fitting a human model. Then a bilinear model is applied to the dynamic features for transforming the reference speed to another arbitrary speed. With regard to the model-free cross-speed gait recognition, Liu et al [16] proposed a Hidden Markov Model (HMM) based time-normalization gait feature. Similarity measurement is presented in a Linear Discriminant Analysis (LDA) Space. Similarly, Tanawongsuwan et al [21] proposed a stride normalized gait features for gait recognition via learning the statistic relation between the walking speed and stride. Differential Composition Model (DCM) [10] is a recent cross-speed gait recognition method that utilizes both appearance feature and prior knowledge of body parameters to solve this issue. It uses a descriptor called Higher-order Derivative Shape Configuration (HSC) to describe the gait sequences. Then, it separates the signature into different local parts based on the body parameters, and measure their similarities individually. In general, the normalization-based method is simple but does not perform well in the situation of large speed variation. On the other hand, the methods which need the prior knowledge of body parameters are very complex.

In this paper, we aim at extracting speed invariant gait features directly without any body parameters. We use a manifold analysis approach to extract gait features, since recent studies indicate actions reside on manifolds [20, 23], and the manifold of walking action is topologically equivalent to a unit circle [8, 9, 13]. Consequently, the problem of gait feature extraction can be formatted as a problem of gait manifold fitting. For addressing this problem, Thin Plate Spline (TPS) kernel based RBF interpolation [9] is used to fit the gait manifold. The obtained mapping coefficients are considered as a gait feature, since they potentially preserve the characteristics of people motion. Furthermore, there is a unique property of TPS that it separates the mapping coefficients into an affine component and a non-affine component [27] and a natural assumption can be given that the affine component depicts the manifold deformation caused by walking and the non-affine component preserves the static features irrelevant to motion. Hence, the non-affine component is a gait feature robust to the walking speed variation and we call it *Speed Invariant Gait Template* (SIGT) (see figure 1). The dimensions of original SIGTs are still high. It can often results in overfitting and the feature is too sparse for characterizing the topography of essential gait space. Thus, the application of dimensionality reduction is necessary. In this paper, we use a recent method: Globality-Locality Preserving Projections (GLPP) for dimensionality reduction [8]. Unlike Principal Component Analysis (PCA)[25] and Linear Discriminant Analysis (LDA)[19], GLPP is an improved Locality Preserving Projections (LPP) method which can better preserve the data's

geometric structures via considering the geometric structures of both samples and classes.

Two walking speed related gait databases are chosen for evaluating our method. Several recent template-based methods and cross-speed methods are used for comparison. The experimental results demonstrate that our method outperforms them. There are 3-fold contributions of our method:

1. The novelty of our method is that we translate the gait feature extraction problem into a manifold fitting problem and use the RBF interpolation to generatively model the walking action without any body parameters. To the best of our knowledge, our method is the first manifold based cross-speed gait recognition method and template-based gait recognition method.
2. We provide a natural way to separate the dynamic features and static features via thin plate spline kernel based RBF interpolation. The separation between dynamic features and static features is general issue to the other computer vision issues. For example, the facial expression can be separated to expressing motion and static facial characteristics via this model.
3. We choose a new dimensionality reduction method to improve the SIGTs and speed up the recognition. This method can well preserve the geometric structures of data and perform very well cross different gait features.

The rest of paper is organized as follows: we introduce our method in section 2; section 3 describes the dimensionality reduction using GLPP; experiments are presented in section 4; the conclusion is finally summarized in section 5.

2 Speed Invariant Gait Template

2.1 Silhouette Representation

In this paper, we will not discuss the procedure of the gait sequence extraction and period detection. We assume the input sequence is a full cycle gait sequence. Besides, an important point should be noted is that the walking phase should be in the same order for all gait sequences in our case. This is because our method is temporal order cared method. However, manually aligning the gait sequences is very hard, since the length of each gait sequence is different and the frames are sampled discretely. In order to address this problem, we follow the strategy of Lee et al [14]. The work of Lee et al [14] also uses homeomorphic manifold analysis to model gait sequences for a general human identification. However, different to our approach, it focuses on using the manifold model to generatively synthesize the same length gait sequence for each instance, and use a bilinear model to identify them. In our approach, we just use a part of their works for gait sequence alignment and it will be introduced in section 2.3.

Implicit function based representation is chosen to present the gait silhouettes. This is a landmark free representation and typically used in level-set based method. The main advantage of this representation is that it is robust to the noise and silhouette fragmentation. We represent each gait silhouette as an implicit function $y(x)$ at each pixel x that $y(x) = 0$ on contour, $y(x) > 0$ inside the contour, and $y(x) < 0$ outside the contour. The signed distance function is used:

$$y(x) = \begin{cases} d_c(x), & x \text{ inside contour} \\ 0, & x \text{ on contour} \\ -d_c(x), & x \text{ outside contour} \end{cases} \quad (1)$$

where $d_c(x)$ is the distance to the closest point on the contour with a positive sign inside the contour and a negative sign outside the contour. Finally, each represented gait silhouette is a

point $y_i \in R^l, i = 1, \dots, N$ in the gait space where l is the dimension of the gait space, i is the index of silhouette in the gait sequence and N is the length of sequences.

2.2 Manifold Fitting using TPS based Interpolation

Some recent studies indicate the gait manifold is topologically equivalent to a unit circle [9, 10, 11], while some factors like walking style, walking speed and human shape translate and deform the circle depending on the subjects. Thus, these subject related factors can be obtained from gait sequences by learning the nonlinear mapping between the actual gait manifold and the unit circle.

We use TPS kernel based RBF interpolation for learning the mapping between the embedding space of walking action (a unit circle) and the input visual space (the gait space). Let $Y = \{y_i \in R^l, i = 1, \dots, N\}$ be a gait sequence in the gait space and $X = \{x_i \in R^2, i = 1, \dots, N\}$ be the corresponding points in the embedding space. $T = \{t_j \in R^2, j = 1, \dots, N_t\}$ denotes N_t equally spaced centers in the embedding space. We can solve for multiple TPS kernel interpolants $f^k : R^2 \rightarrow R$ where k is the k th element (dimension) of the signed distance function based gait vector (the vectorized gait image) in the gait space and f^k is a RBF interpolant. We minimize a regularized risk criteria to learn the nonlinear mapping from the embedding space to each individual dimension in the visual input space that satisfies $y_i^k = f^k(x_i)$. From the representer theorem [12], such a function admits a representation of the form of linear combination of basis functions around arbitrary points (centers). Thus, to the k th dimension of the input, the form of function $f^k(x)$ is as follows:

$$f^k(x) = p^k(x) + \sum_{i=1}^{N_t} d_i^k \phi(\|x - t_i\|_2) = p^k(x) + \sum_{i=1}^{N_t} d_i^k \|x - t_i\|_2 \log(\|x - t_i\|_2) \quad (2)$$

where function $\phi(u) = u^2 \log(u)$ is a thin plate spline function and $p^k(x) = [1, x^T] \cdot s^k$ denotes the TPS smoothness term as a linear polynomial function with coefficients s^k . The matrix form of interpolation is as follows:

$$f(x) = W \cdot \Phi(x) \quad (3)$$

where W is a $l \times (N_t + 3)$ coefficients matrix with the k th row $[d_1^k, \dots, d_{N_t}^k, s^k]^T$ and $\Phi(x)$ is a vector that $\Phi(x) = [\phi(\|x - t_1\|_2), \dots, \phi(\|x - t_{N_t}\|_2), 1, x^T]$. The matrix W represents the mapping coefficients which are the l nonlinear mappings from the embedding space to gait space. In order to make the problem be well posed and insure the orthogonality, an additional constraint should be added:

$$\sum_{i=1}^{N_t} d_i p_j(x_i) = 0, j = 1, 2, 3 \quad (4)$$

where p_j is the linear basis of the polynomial part $p(x)$.

Thus, the mapping coefficients W can be obtained by directly solving the following linear systems:

$$\begin{pmatrix} C & P_x \\ P_t^T & 0_{3 \times 3} \end{pmatrix} W^T = \begin{pmatrix} C & P_x \\ P_t^T & 0_{3 \times 3} \end{pmatrix} (D \ S)^T = \begin{pmatrix} Y \\ 0_{3 \times l} \end{pmatrix} \quad (5)$$

where C is a $N \times N_t$ matrix with $C_{ij} = \phi(\|x_i - t_j\|_2), i = 1, \dots, N, j = 1, \dots, N_t, P_x$ is a $N \times 3$ matrix with i th row $[1, x_i^T]$, P_t is a $N_t \times 3$ matrix with i th row $[1, t_i^T]$. $D = [d_1, \dots, d_{N_t}]$ is a

$l \times N_t$ coefficient matrix of TPS based interpolants and $S = [s^T]$ is a $l \times 3$ coefficient matrix of the smoothness term. Y is an aligned gait sequence.

Compared to other splines, TPS has a desirable property that it can always be decomposed into a global affine and a local non-affine component and the TPS smoothness term is solely dependent on the non-affine components [17]. Therefore, the TPS based interpolation provides a natural separation between the dynamic features caused by walking action and the static features invariant to the walking action. In our case, the dynamic features are embedded in the matrix D while the static features are embedded in the matrix S . Consequently, the $l \times 3$ matrix S is a speed-invariant gait features. The matrix S naturally provides three l dimensional gait templates. Finally, we yield these three templates as a $3 \times l$ vector. We name this vector *Speed Invariant Gait Templates* (SIGT) and apply it for cross-speed gait recognition.

2.3 Gait Sequences Alignment via Synthesized Data

Algorithm 1 Gait Sequence Alignment

Require:

The input unaligned gait sequence, $Y = [y_1, \dots, y_M]$;

The mapping coefficients learned from a manually aligned gait sequence, W_o ;

The equally spaced mapping centers $T = \{t_j = [\cos(2(j-1)\pi/N_t), \sin(2(j-1)\pi/N_t)]^T \in \mathbb{R}^2, j = 1, \dots, N_t\}$

Ensure:

The aligned gait sequence, E ;

1: Calculate the length of input gait sequences, $M = \text{length}(Y)$;

2: Calculate the corresponding equally spaced points in the embedding space of walking action, $X = \{x_i = [\cos(2(i-1)\pi/M), \sin(2(i-1)\pi/M)]^T \in \mathbb{R}^2, i = 1, \dots, M\}$;

3: Calculate the matrices C, P_t, P_x based on X and T ;

4: Synthesize the ordered gait sequence as an exemplar sequence based on C, P_t, P_x and W_o

$$: \begin{pmatrix} G \\ 0_{3 \times l} \end{pmatrix} = \begin{pmatrix} C & P_x \\ P_t^T & 0_{3 \times 3} \end{pmatrix} W_o^T$$

where $G = [g_1, \dots, g_M]$ is the synthesized gait sequence.

5: **for** each $i \in [1, M]$ **do**

6: Shift the input gait sequence Y from frame i to yield a new gait sequence $NY = [y_i, \dots, y_N, y_1, \dots, y_{i-1}]$;

7: Calculate the similarity score between the shifted gait sequence NY and the synthesized gait sequence G and put the similarity score into array $\text{Score}(i) = \|NY - G\|_2$;

8: **end for**

9: Find the index of array j whose element owns the minimum similarity;

10: Align the input gait sequences with that index, $E = [y_j, \dots, y_M, y_1, \dots, y_{j-1}]$;

11: **return** E ;

Our method is a temporal order information cared gait recognition method. Thus, the input gait sequences should be full cycle and their walking phases should be at the same order. So, we must align the input gait sequences into the same order for meeting the input requirement. Since our model is generative, we can follow the Lee's strategy [14] to use the synthesized gait sequence to align the gait sequences and the experimental result from [14] also proves that this strategy can achieve a good alignment. The detail gait sequence alignment procedure is described in Algorithm 1.

3 Dimensionality Reduction using GLPP

Although we convert the whole sequence into a vector yielded by three gait templates, the dimension of this gait feature is still very high. This high dimensions lead to overfitting and slow up the recognition. In order to tackle this problem, we use an Improved LPP method name *Globality-Locality Preserving Projections* (GLPP) [8] to reduce dimensionality. Compared to some conventional methods, such as PCA and LDA, such kind of LPP-based method

takes the geometric structures of data into consideration and this geometric information can benefit the data classification [2, 4, 8]. Furthermore, compared to the conventional LPP, GLPP can better preserve the geometric structures of data, since it not only preserves geometric structures of samples, but also preserves geometric structures of classes.

Given a list of SIGTs, $X = \{x_i \in \mathbb{R}^d, i = 1, \dots, N\}$ where $d = 3l$, we aim at learning the GLPP projections $W = [w_1, \dots, w_m]$ to transform the original SIGTs into a low-dimensional representation $Y = \{y_i \in \mathbb{R}^m, i = 1, \dots, N\}$. The vector $C = [1, \dots, h]$ denotes the category labels and $\bar{x}_j, j \in C$ denotes the mean of the samples belonging to class j . Similarly, $\bar{y}_j, j \in C$ denotes the mean of the projected samples belonging to class j .

GLPP has two objective terms. The globality preserving objective term O_g is used to preserve the geometric structures of classes and the locality preserving objective term O_w is used to preserve the geometric structures of samples. They are defined respectively as follows:

$$O_g = \sum_{i,j \in C} (\bar{y}_i - \bar{y}_j)^2 B_{ij} = \sum_{i,j \in C} (w^T \bar{x}_i - w^T \bar{x}_j)^2 B_{ij} \quad (6)$$

$$O_w = \sum_{c \in C} \sum_{i,j \in c} (y_i - y_j)^2 S_{ij} = \sum_{c \in C} \sum_{i,j \in c} (w^T x_i - w^T x_j)^2 S_{ij} \quad (7)$$

where matrices S and B are the adjacency weight matrices of the objective terms O_w and O_g respectively. According to work of Cai et al [4], there are three ways to construct these matrices. In this paper, we choose the dot-product weighting to construct each adjacency matrix.

GLPP aims at preserving the geometric structures of both samples and classes via minimizing both these two objective terms:

$$\begin{aligned} \hat{w} &= \arg \min_w (O_g + \beta \cdot O_w) \\ &= \arg \min_w \left(\sum_{i,j \in C} (w^T \bar{x}_i - w^T \bar{x}_j)^2 B_{ij} + \beta \cdot \sum_{c \in C} \sum_{i,j \in c} (w^T x_i - w^T x_j)^2 S_{ij} \right) \\ &= \arg \min_w (w^T (\bar{X} K \bar{X}^T + \beta \sum_{c \in C} (X_c L_c X_c^T)) w) = \arg \min_w (w^T A w) \end{aligned} \quad (8)$$

where matrix \bar{X} is a matrix containing the mean of each class, and matrix X_c is the matrix of the samples belonging to class c . Matrices K and L_c are the Laplacian matrices of \bar{X} and X_c respectively. $\beta > 1$ is a parameter balances the two objective terms. The matrix A is a positive semi-definite matrix. Therefore, this problem can finally solved as a general eigenvalue problem as follows:

$$\lambda w = A w \quad (9)$$

The optimal solution is the eigenvector corresponding to the first m minimum nonzero eigenvalue λ . Thus we can finally obtain the GLPP projection matrix $W = [w_1, \dots, w_m]$. Then we can transform the SIGTs to a more compact representation via the GLPP projection and measure the similarity for gait recognition.

4 Experiments

4.1 Databases

Two walking speed related databases, CMU Mobo gait database [8] and OU-ISIR Treadmill dataset A [24], are used to evaluate the performances of different gait recognition methods

in this paper. The CMU Mobo gait database [9] has 25 subjects with 6 views. There are four subsets including Slow Walk, Fast Walk, Ball and Incline subsets. We only choose the Slow Walk and Fast Walk subsets for studying the gait recognition performance cross various walking speeds. The walking speed of the Slow Walk subset is 3.3km/h and the walking speed of the Fast Walk subset is 4.5km/h. OU-ISIR Treadmill dataset A [24] is designed for studying the effect of speeds variation on gait recognition. This database contains 34 subjects with speed variation from 2 km/h to 10 km/h at 1 km/h interval. Each subject has two sequences of "gallery" and "probe".

4.2 Compared Methods

Because our method is a template-based method, three recent influential template-based methods, namely, Gait Energy Image (GEI) [9], Gait Flow Image (GFI) [22] and Chrono-Gait Image (CGI) [23], are chosen for comparison. Besides, the experimental results of two cross-speed gait recognition methods, Shape variation-based frieze pattern (SVFP) [15] and Differential Composition Model (DCM) [16], are also referenced for comparison.

4.3 Cross-speed Gait Recognition Results

In these experiments, the size of gait silhouette is 128×88 and 12 equally spaced points around the unit circle are considered as centers. The Euclidean distance is the similarity metric.

With regard to the experiments of Mobo database, only two subsets, Slow Walk subset and Fast Walk subset, are chosen for experiments. Furthermore, we do the experiments for each views individually instead of cross views. This is because our objective mainly focuses on studying the cross-speed gait recognition. We define two group of experiments on Mobo database. In the first group of experiments, the Slow Walk subset is used as a gallery and used to learn the projections while the Fast Walk subset is the probe. In the second group of experiments, the Fast Walk subset is the gallery and the Slow Walk subset is the probe.

Table 2 shows the experimental results of different gait recognition methods on Mobo database. The recognition rates in table 2 are the average recognition rate of the best results under different views. From the observations of table 2, SIGTs, dynamic features (**Deform** in table 2) and whole features (both dynamic and static features, **Whole** in table 2) outperform the compared methods. We also can deduce that the walking-related dynamic features can be also well used for identifying human when the walking speed is not changing a lot. However, there is a disadvantage of the dynamic features and whole features. They have a much higher dimension than the static features (SIGT). So, if the performances of the SIGTs, dynamic features and whole features are similar, we can choose the lower dimensional one.

OU-ISIR database is a more challenging database since the walking speeds of this database are extremely various. For studying the recognition performance with various speeds, we divide the database into three subsets. The configurations of these subsets are described in the table 1.

subset	gallery	probe
subset1	2,3,4 km/h	all walking speeds (2–10 km/h)
subset2	5,6,7 km/h	all walking speeds (2–10 km/h)
subset3	8,9,10 km/h	all walking speeds (2–10 km/h)

Table 1: The configurations of different subsets on OU-ISIR Treadmill dataset A

We can clearly see in table 3 that our method prominently outperforms the recent gait template-based methods in all three subsets. The gains of our method over the second-ranked method are 3.65%, 7.09% and 9.62 % in first, second and third subset respectively. These experimental results demonstrate the superiority of our method for gait recognition particularly when the walking speeds are extremely various. However, compared to the SIGTs, the whole features and the dynamic features are not performing well on this database. This is because the walking speeds are more variant on OU-ISIR database which results in the drastic change of the dynamic information caused by the walking action.

Gallery \ Probe	Rank-1 Recognition Rate(%)			Rank-5 Recognition Rate(%)		
	Slow\fast	fast\slow	Total	Slow\fast	fast\slow	Total
GEI + LDA [16]	84.85	81.73	83.09	89.97	86.63	88.30
GEI + PCA [16]	80.77	75.55	78.27	87.55	81.98	84.77
GEI + LPP	90.39	90.52	90.46	91.84	90.61	91.23
GEI + GLPP	92.32	91.88	92.10	93.22	92.72	92.97
GFI + LDA [16]	83.68	64.97	74.59	86.45	67.94	77.20
GFI + PCA	78.22	57.61	67.92	84.85	69.29	77.07
GFI + LPP	81.95	79.44	80.70	84.09	79.61	81.85
GFI + GLPP	83.89	74.79	79.34	84.99	79.02	82.01
CGI + PCA [23]	75.73	78.60	77.17	81.14	83.42	82.28
CGI + LDA [23]	79.25	81.64	80.45	79.94	82.15	81.05
CGI + LPP	76.97	83.42	80.20	77.18	83.50	80.34
CGI + GLPP	79.60	84.77	82.19	80.43	86.04	83.24
DCM [16] ¹	92.00	88.00	90.00	\	\	\
FSVB [16] ¹	82.00	80.00	81.00	\	\	\
SIGT + PCA	81.88	79.36	80.62	90.11	84.35	87.23
SIGT + LDA	92.67	88.24	90.46	93.29	90.52	91.91
SIGT + LPP	90.53	91.03	90.78	91.77	91.03	91.40
Deform + GLPP	92.32	91.88	92.10	93.22	92.72	92.97
Whole + GLPP	92.88	91.79	92.34	93.98	92.13	93.06
SIGT + GLPP	92.88	91.88	92.38	93.98	92.13	93.06

Table 2: Recognition performance comparison (in percents) using Mobo database. Deform represents the dynamic features extracted by TPS based manifold fitting, Whole represents the whole features (both the dynamic features and SIGTs).

Several experiments are conducted for studying the SIGT’s tolerance of speeds variation. Figure 2 shows these results. In these figures, SIGT, which is the non-affine part of TPS mapping coefficients, is more insensitive to the speed variation in comparison with Whole, which is the whole mapping coefficients, and Deform, which is the affine part of mapping coefficients. Although the recognition performance of SIGT is falling along with the speed interval increasing, but it falls much more smoothly and slowly than Deform and Whole. So, we think this phenomenon justifies that the non-affine part (SIGT) is related to the static feature of gait sequence which is more robust to the speed variation.

Furthermore, according to the experimental results of both Mobo database and OU-ISIR database, GLPP as a dimensionality reduction method almost outperforms all the other dimensionality methods cross all the gait features.

¹These experimental results are directly reference from recent paper of Kusakunniran et al [16].

Methods	The best Rank-1 Recognition Rate (retained dimension)			
	subset1	subset2	subset3	total
GEI + LDA [6]	64.75(33)	69.87(33)	59.90(27)	64.84
GEI + PCA [6]	52.32(313)	54.99(409)	39.61(392)	48.97
GEI + LPP	63.97(79)	69.24(52)	54.63(120)	62.21
GEI + GLPP	66.50(66)	73.95(341)	58.99(120)	66.48
GFI + LDA [10]	30.97(33)	44.17(33)	40.80(33)	38.65
GFI + PCA	28.79(79)	36.80(551)	37.85(35)	34.48
GFI + LPP	33.29(40)	46.14(35)	43.26(103)	40.90
GFI + GLPP	35.39(118)	46.49(69)	40.80(86)	40.89
CGI + PCA [28]	50.42(261)	52.74(69)	41.15(494)	48.10
CGI + LDA [28]	61.03(33)	66.15(33)	45.72(33)	57.53
CGI + LPP	58.57(118)	67.21(69)	45.93(239)	57.24
CGI + GLPP	62.36(40)	68.05(443)	45.79(375)	58.73
SIGT + PCA	42.63(118)	46.14(103)	40.10(52)	42.96
SIGT + LDA	64.53(23)	73.59(31)	53.79(9)	63.97
SIGT + LPP	63.20(79)	71.77(79)	56.32(188)	63.76
Deform + GLPP	47.05(53)	49.44(222)	41.26(424)	45.92
Whole + GLPP	51.06(144)	52.39(103)	41.15(307)	48.20
SIGT+ GLPP	70.15(313)	81.04(104)	68.61(273)	73.27

Table 3: Recognition performance comparison (in percents) using OU-ISIR Treadmill dataset A

5 Conclusion

We presented a novel gait templates for cross-speed gait recognition. These gait templates can be extracted by a TPS Kernel based RBF interpolation. Furthermore, a very recent method, GLPP is applied to reduce the dimensions of gait templates. The experimental results demonstrate that the proposed method outperforms the compared methods, particularly when the walking speeds are extremely changing. In this paper, our proposed method is a very basic version. There are still many meaningful tasks that can be done. For example, the extracted dynamic features can be fused with SIGT to present a more general gait recognition. We also can manipulate the SIGT like other template-based methods to synthesize more templates to predict the gait recognition under unknown situations. Furthermore, this model is not also just suitable to the gait recognition, the dynamic feature and static feature separation can be also applied in other computer vision tasks, such as expression analysis and gesture recognition.

Acknowledgement

This work has been partially funded by NSF CAREER award number IIS- 0546372, Fundamental Research Funds for the Central Universities (Grant No. CDJXS11181162) and the National Natural Science Foundation of China (Grant No. 91118005). The authors would like to thank the reviewers for their useful comments.

References

- [1] I. Bouchrika and M.S. Nixon. Model-based feature extraction for gait analysis and recognition. In *MIRAGE*, pages 150–160. Springer, 2007.
- [2] D. Cai, X. He, J. Han, and T.S. Huang. Graph regularized nonnegative matrix factorization for data representation. *IEEE T-PAMI*, 33(8):1548–1560, 2011.

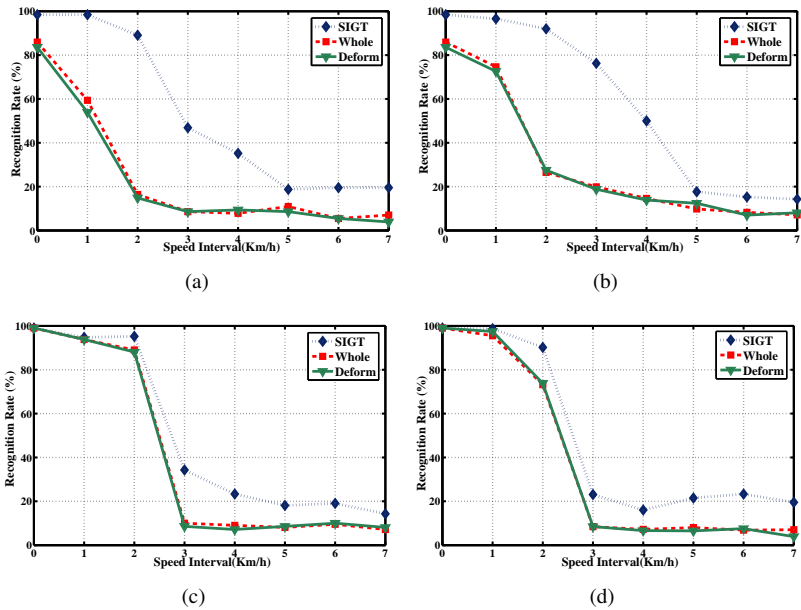


Figure 2: Speed variation tolerances of SIGT, Whole and Deform using (a) 3 km/h probe and 3-10 km/h galleries, (b) 3 km/h gallery and 3-10 km/h probes, (c) 10 km/h probe and 3-10 km/h galleries, and (d) 10 km/h gallery and 3-10 km/h probes.

- [3] A. Elgammal and C-S. Lee. Inferring 3d body pose from silhouettes using activity manifold learning. In *CVPR*, volume 2, pages II-681, 2004.
- [4] A. Elgammal and C-S. Lee. Homeomorphic manifold analysis (hma): Generalized separation of style and content on manifolds. *Image and Vision Computing*, 2013.
- [5] R. Gross and J. Shi. The cmu motion of body (mobo) database. *Technical Report CMU-RI-TR-01-18*, 2001.
- [6] J. Han and B. Bhanu. Individual recognition using gait energy image. *IEEE T-PAMI*, 28(2):316-322, 2006.
- [7] X. He and P. Niyogi. Locality preserving projections. In *NIPS*, 2003.
- [8] S. Huang, D. Yang, F. Yang, Y. Ge, X. Zhang, J. Zhou, and J. Lu. Face recognition via globality locality preserving projections. *Neurocomputing*, (under review), 2013.
- [9] G.S. Kimeldorf and G. Wahba. A correspondence between bayesian estimation on stochastic processes and smoothing by splines. *The Annals of Mathematical Statistics*, 41(2):495-502, 1970.
- [10] W. Kusakunniran, Q. Wu, J. Zhang, and H. Li. Gait recognition across various walking speeds using higher order shape configuration based on a differential composition model. *IEEE T-MSCB*, pages 1654-1668, 2012.
- [11] T.H.W. Lam, R.S.T. Lee, and D. Zhang. Human gait recognition by the fusion of motion and static spatio-temporal templates. *Pattern Recognition*, 40(9):2563-2573, 2007.

- [12] T.H.W. Lam, K.H. Cheung, and J.N.K. Liu. Gait flow image: A silhouette-based gait representation for human identification. *Pattern Recognition*, 44(4):973–987, 2011.
- [13] C-S. Lee and A. Elgammal. Gait style and gait content: bilinear models for gait recognition using gait re-sampling. In *FG*, pages 147–152, 2004.
- [14] C-S. Lee and A. Elgammal. Dynamic shape style analysis: Bilinear and multilinear human identification with temporal normalization. *IJPRAI*, 24(07):1133–1157, 2010.
- [15] S. Lee, Y. Liu, and R. Collins. Shape variation-based frieze pattern for robust gait recognition. In *CVPR*, pages 1–8, 2007.
- [16] Z. Liu and S. Sarkar. Improved gait recognition by gait dynamics normalization. *IEEE T-PAMI*, 28(6):863–876, 2006.
- [17] J. Lu and E. Zhang. Gait recognition for human identification based on ica and fuzzy svm through multiple views fusion. *Pattern Recognition Letters*, 28(16):2401–2411, 2007.
- [18] M.S. Nixon and J.N. Carter. Automatic recognition by gait. *Proceedings of the IEEE*, 94(11):2013–2024, 2006.
- [19] N.B. Peter, P. Hespanha, and J.K. David. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE T-PAMI*, pages 711–720, 1997.
- [20] H.S. Seung and D.D. Lee. The manifold ways of perception. *Science*, 290(5500): 2268–2269, 2000.
- [21] R. Tanawongsuwan and A. Bobick. Modelling the effects of walking speed on appearance-based gait recognition. In *CVPR*, volume 2, pages II–783, 2004.
- [22] D. Tao, X. Li, X. Wu, and S.J. Maybank. General tensor discriminant analysis and gabor features for gait recognition. *IEEE T-PAMI*, 29(10):1700–1715, 2007.
- [23] J.B. Tenenbaum. Mapping a manifold of perceptual observations. *NIPS*, pages 682–688, 1998.
- [24] A. Tsuji, Y. Makihara, and Y. Yagi. Silhouette transformation based on walking speed for gait identification. In *CVPR*, pages 717–722, 2010.
- [25] M. Turk and A. Pentland. Eigenfaces for recognition. *J. Cognitive Neuroscience*, 3(1): 71–86, January 1991.
- [26] D.K. Wagg and M.S. Nixon. On automated model-based extraction and analysis of gait. In *FG*, pages 11–16, 2004.
- [27] G. Wahba. *Spline models for observational data*, volume 59. Society for industrial and applied mathematics, 1990.
- [28] C. Wang, J. Zhang, L. Wang, J. Pu, and X. Yuan. Human identification using temporal information preserving gait template. *IEEE T-PAMI*, 34:2164–2176, 2012.