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

Sheng Huang[1]
https://sites.google.com/site/siteshuang/

Ahmed Elgammal[2]
http://www.cs.rutgers.edu/~elgammal

Dan Yang[1]

[1] College of Computer Science, Chongqing University

[2] Department of Computer Science, Rutgers, The State University of New Jersey

Similar to the solutions of general gait recognition approaches, there are two ways to handle the cross-speed gait recognition issues. The first approach is model-based approach which is to model the walking action using static or dynamic body parameters [6]. The second class of cross-speed gait recognition approach is model free approach, also known as feature or appearance-based approach. It focuses on directly extracting holistic gait features from gait sequences [4, 5]. Our proposed method, Speed Invariant Gait Template (SIGT), belongs to the second category.

According to some recent studies, the walking action is considered as residing on a manifold which is topologically equivalent to a unit circle [1]. Thus, we can format the gait feature extraction issue as a gait manifold fitting issue. To address the cross-speed gait recognition issue, Thin Plate Spline (TPS) kernel based RBF interpolation is used to fit the gait manifold, since TPS has a desirable property [7] that it separates the mapping coefficients into an affine component and a non-affine component. And a natural assumption, that the affine component depicts the manifold deformation caused by walking and the non-affine component preserves the static features irrelevant to motion, can be given. Consequently, the non-affine component is a gait feature which is robust to the walking speed variation and we name this new feature *Speed Invariant Gait Template* (SIGT) (see Figure 1).

We show in this paper how to use our proposed TPS-based manifold model to extract SIGTs from the full cyclic gait silhouette sequences. Firstly, The implicit function-based representation is used to represent each gait silhouette, since this representation is robust to the noise and silhouette fragmentation. Each pixel x of a gait silhouette is represented as 0 when it is on the contour, as $d_c(x)$ when it is inside the contour and as $-d_c(x)$ when it is outside the contour, where $d_c(x)$ denotes the distance to the closest point on the contour.

Next, TPS Kernel based RBF interpolation is used to learn the mapping between the manifold embedding space and the represented gait space. Let $Y = \{y_i \in R^l, i = 1, \cdots, N\}$ be a gait sequence in the gait space and $X = \{x_i \in R^2, i = 1, \cdots, N\}$ be the corresponding points in the embedding space. $T = \{t_j \in R^2, i = 1, \cdots, N_t\}$ denotes N_t equally spaced centers in the embedding space. We can solve for multiple TPS kernel interpolants $f^k : R^2 \to R$ where k is the kth element (dimension) of the gait vector (the vectorizated represented gait image) in the gait space and f^k is a RBF interpolant. We minimize a regularized risk criteria to learn nonlinear mapping from the embedding space to each individual dimension in the gait space that satisfies $y_i^k = f^k(x_i)$. From the representer theorem [3], such a function admits a representation of the form of linear combination of basis functions around arbitrary points (centers). Therefore, to the kth dimension of the input, the form of function $f^k(x)$ is as follows:

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

where $\phi(u) = u^2 log(u)$ 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:

$$f(x) = W \cdot \Phi(x) \tag{2}$$

where W is a $l \times (N_t + 3)$ matrix with the kth row $[d_1^k, \cdots, d^k{}_{N_t}, s^{k^T}]$ and $\Phi(x)$ is a vector that $\Phi(x) = [\phi(||x - t_i||_2), \cdots, \phi(||x - t_{N_t}||_2), 1, x^T]$. 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:

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

(a) (b)

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.

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:

$$\begin{pmatrix} C & P_x \\ P_t{}^T & 0_{3\times3} \end{pmatrix} W^T = \begin{pmatrix} C & P_x \\ P_t{}^T & 0_{3\times3} \end{pmatrix} \begin{pmatrix} D & S \end{pmatrix}^T = \begin{pmatrix} Y \\ 0_{3\times l} \end{pmatrix} \tag{4}$$

where C is a $N \times N_t$ matrix with $C_{ij} = \phi(||x_i - t_j||_2), i = 1, \cdots, N, j = 1, \cdots, N_t$, P_x is a $N \times 3$ matrix with ith row $[1, x_i{}^T]$, P_t is a $N_t \times 3$ matrix with ith row $[1, t_i{}^T]$. $D = [d_1, \cdots, d_{N_t}]$ 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. According to the property of TPS, the static feature of gait manifold, which is robust to the speed variation, is embedded in the $l \times 3$ matrix S. Then, the proposed *Speed invariant Gait Template* (SIGT) can be obtained by vectorizing this matrix.

Finally, in order to speed up the recognition and avoid the curse of dimensionality, an recent Improved Locality Preserving Projections (LPP) method named Globality-Locality Preserving Projections (GLPP) [2] is applied to reduce the dimensionality of SIGT. Compared to the state-of-the-arts methods such as LDA, PCA and LPP, GLPP obtains a much better recognition performance via taking geometric structures of both samples and classes into consideration.

The implementations of our method and related algorithms are described in this paper in detail. Our conclusion is that a novel gait template for cross-speed gait recognition is proposed and these gait templates are extracted by fitting the gait manifold via TPS kernel based RBF interpolation. The main contribution of this work is that it provides a natural way to separate the dynamic features and static features and such separation is very general to other computer vision issues.

[1] A. Elgammal and C-S. Lee. Homeomorphic manifold analysis (hma): Generalized separation of style and content on manifolds. *Image and Vision Computing*, 2013.

[2] 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.

[3] 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.

[4] 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.

[5] Z. Liu and S. Sarkar. Improved gait recognition by gait dynamics normalization. *IEEE T-PAMI*, 28(6):863–876, 2006.

[6] A. Tsuji, Y. Makihara, and Y. Yagi. Silhouette transformation based on walking speed for gait identification. In *CVPR*, pages 717–722, 2010.

[7] G. Wahba. *Spline models for observational data*, volume 59. Society for industrial and applied mathematics, 1990.