Towards Scalable View-Invariant Gait Recognition: Multilinear Analysis for Gait

Chan-Su Lee and Ahmed Elgammal

Department of Computer Science
Rutgers University
Piscataway, NJ, USA

Abstract. In this paper we introduce a novel approach for learning view-invariant gait representation that does not require synthesizing particular views or any camera calibration. Given walking sequences captured from multiple views for multiple people, we fit a multilinear generative model using higher-order singular value decomposition which decomposes view factors, body configuration factors, and gait-style factors. Gait-style is a view-invariant, time-invariant, and speed-invariant gait signature that can then be used in recognition. In the recognition phase, a new walking cycle of unknown person in unknown view is automatically aligned to the learned model and then iterative procedure is used to solve for both the gait-style parameter and the view. The proposed framework allows for scalability to add a new person to already learned model even if a single cycle of a single view is available.

1 Introduction

Human gait is a valuable biometric cue that can be used for human identification similar to other biometrics such as faces and fingerprints. Gait has significant advantages compared to other biometrics since it is easily observable in an unintrusive way and is difficult to disguise [4]. Therefore, gait recognition has a great potential for human identification in public spaces for surveillance and for security [4, 10, 11, 21]. A fundamental challenge in gait recognition is to develop robust recognition algorithms that can extract gait features that are invariant to the presence of various conditions which affect people appearance. As a challenging problem in gait recognition, different conditions such as view, clothing, walking surface, and shoe type were presented in the NIST dataset [21]. Many gait recognition algorithms assume constrained conditions to reduce various sources that influence recognition accuracy. Two typical assumptions are fixed view, especially side view, and constant speed.

Generally, appearance-based approaches have been favorable in gait recognition [3, 9, 10, 13, 16–21, 24, 26, 30, 31] because, in typical application scenarios, people might be at a distance from the camera which inhibits accurate fitting of 3D models. Therefore, many gait recognition research focus on extracting view-based invariant gait signature for use in identification. Several attempts have been made to achieve view-invariant gait recognition [2, 8, 12, 23], mainly based on synthesizing side-view images from multiple views. For example, Shakhnarovich [23] used image-based visual hull to render a side view from multiple cameras. Kale [12] also presented the view invariant gait recognition algorithm by synthesizing a side view using perspective projection methods and optical based structures.

T. Kanade, A. Jain, and N.K. Ratha (Eds.): AVBPA 2005, LNCS 3546, pp. 395–405, 2005.

In this paper we introduce a novel approach for learning view-invariant gait signature that does not require synthesizing particular views and doesn't require any camera calibration. Instead, in the learning phase, multiple views are used to extract an invariant gait signature while in recognition phase, any single view can be used to extract the gait signature directly. Given walking sequences captured from multiple views for multiple people, first we learn a nonlinear generative model for each walking cycle which enables re-sampling the cycle into temporally aligned gait cycles. Then we fit a multilinear generative model using higher-order singular value decomposition (HOSVD) [14] that decomposes view factors, body configuration factors, and gait-style factors. Gait-style is a view-invariant, time-invariant, and speed-invariant gait signature that can then be used in recognition. In the recognition phase, a new walking of unknown person in unknown view is aligned to the learned model and then iterative procedure is used to solve for both the gait-style and the view. Related work in using mulitlinear analysis for gait includes [27]

One important feature of the proposed framework is it's scalability to include new people. Given a learned model we can add a new person to the model even if only single view is available for that person, i.e., one cycle gait sequence from one of the multiple possible views is need to include new people to the database. This is a very important feature since in realistic scenarios, it is not always possible to have multiple view sequences of each person to be included in the database. Experimental results using CMU Mobo gait database and NIST-USF database [21] are reported in this paper.

The organization of the paper is as follows: In Section 2, we introduce temporal normalization of gait and cycle detection. Section 3 describes decomposition of gait-style, view, and body configuration parameters, estimation of style, and its application to gait recognition. Experimental results are described in Section 4 prior to the conclusion in Section 5.

2 Temporal Normalization by Manifold Embedding and Re-sampling

2.1 Input Representation

The inputs to the training and recognition phases are sequences of human silhouettes detected using background subtraction. We represent each shape instance (silhouette) as an implicit function $y(x)$ at each pixel x such that $y(x) = 0$ on the contour, $y(x) > 0$ inside the contour, and $y(x) < 0$ outside the contour. We use a signed-distance function such that

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

where the $d_c(x)$ is the distance to the closest point on the contour c with a positive sign inside the contour and a negative sign outside the contour. Such representation impose smoothness on the distance between shapes. Given such representation, each input silhouette is represented as a d-dimensional vector, i.e., a point $y \in R^d$ where d is the the dimensionality of the input space. Implicit function representation is typically used in level-set methods.

2.2 Temporal Normalization and Re-sampling

In order to achieve the training and recognition we need to obtain temporally aligned input silhouettes, i.e, obtaining body poses in correspondence during the gait cycle given any input sequence captured at any frame rate with any walking speed. To achieve this task we use any given input sequence to learn a nonlinear generative model that can be used to synthesize silhouettes at any temporal instance within the gait cycle.

The human gait evolves along a one-dimensional manifold embedded in a high dimensional visual space. Only one degree of freedom controls the walking cycle, which corresponds to the constrained body pose as a function of time. Such manifold is nonlinear and can be twisted on the high dimensional space given viewpoint, person shape, and clothing [5, 6]. Therefore, we embed each gait cycle temporally on a unit circle, which is a topologically homeomorphic one-dimensional manifold embedded in a two-dimensional Euclidean space.

In order to obtain synthesized gait poses, we learn a nonlinear mapping function from the manifold embedded on a unit circle and the input silhouettes. Learning nonlinear mapping is necessary since the manifold is embedded nonlinearly and arbitrarily into a unit circle. We use generalized radial basis function (GRBF) [22] to learn this mapping as a collection of interpolation functions. Let N equally spaced centers along a unit circle be $\{t_j \in R^2, j = 1, \cdots, N\}$ and given a set input images $\mathsf{Y} = \{y_i \in R^d, i = 1, \cdots, M\}$ and let their corresponding embedding along the unit circle be $\mathsf{X} = \{x_i \in R^2, i = 1, \cdots, M\}$, we can learn interpolations in the form

$$f^k(x) = p^k(x) + \sum_{i=1}^{N} w_i^k \phi(|x - t_i|), \tag{1}$$

that satisfies the interpolation condition $y_i^k = f^k(x_i)$ where y_i^k is the k-th pixel of input silhouette y_i, $\phi(\cdot)$ is a real valued basic function, w_i^k are real coefficients, $p^k(\cdot)$ is a linear polynomial, and $|\cdot|$ is the norm on R^2. The mapping coefficients can be obtained by solving a linear system [5]. Such mapping can be written in the form of a generative model as

$$f(x) = B \cdot \psi(x), \tag{2}$$

which nonlinearly maps any point x from the two dimensional embedding space into the input space. Therefore, the model can be used to synthesize N intermediate silhouettes at N standard time instances equally spaced along the unit circle. Re-sampling gait from the embedding space enables us to find temporally well aligned gait poses invariant to different walking speed and frame rate using equally spaced N embedding points.

2.3 Gait Cycle Detection in Arbitrary Views

Detection of gait cycles is essential for training and recognition. Typically, for side or frontal views, cycles can be detected using features such as width or height of bounding box, correlation of image sequences, etc, [1, 3]. However, detecting cycles in arbitrary views are difficult. The generative model, described above, facilitates detecting accurate gait cycles in any view given that the model parameter is learned on that particular view.

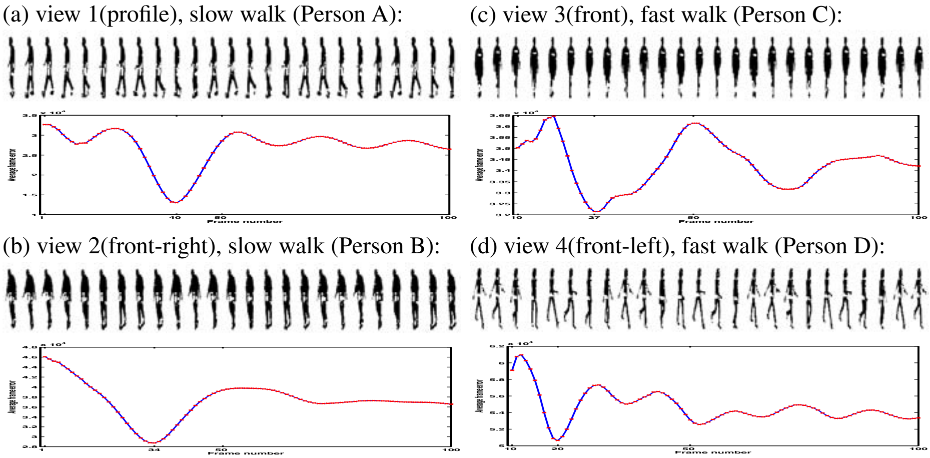

Fig. 1. Cycle detection in different view

Given an input sequence $y_i, i = 1, \cdots, M$ we need to find $k^* \leq M$ such that $y_1, \cdots, y_k$ corresponds to a full cycle. This can be achieved by finding k^* that minimizes the error between an input sequence of length k and model synthesized image sequence, of length k as well, starting from the same point. i.e., we need to find k^* such that

$$k^* = \arg\min_k = \frac{1}{k} \sum_{j=1}^{k} ||f(x_j^k) - y_j||$$

where x_j^k is a point on a unit circle with coordinate $x_j^k = [cos(2\pi \cdot j/k + \delta)\ \ sin(2\pi \cdot j/k + \delta)]$.

To show examples of generative model-based cycle detection, we used CMU Mobo gait data set which shows accurate detection of cycle in different views like side view, front-left view and front views within $1 \sim 2$ frames error. Fig. 1 shows four different view silhouette images (sampled at every 4th frames in the figure). Mean error are shown as a function of k in the range from 10 to 100. Even though we learned generative model for each view from one person, it performs accurately in segmenting cycles at different people.

3 Gait Style and View Decomposition

We model gait image sequences by three components: *gait style:* time-invariant and view-invariant personalized style of the gait which can be used for identification similar to in [16], *gait pose:* time-dependent factor representing body configuration during the gait cycle, and *gait view:* view-dependent factor representing variations of view.

3.1 Multilinear Model for Gait Analysis

Given different people walking sequences from different views, we detect gait cycles using gait cycle detection algorithm in Section 2.3. After cycle detection for every person, each cycle is used to learn the generative model described by equation 2 and re-sampled with the same number of temporally aligned poses. Therefore, the training data consists of N_s gait cycles[1], each captured from N_v different views, and each consists of N_p silhouette images representing aligned body poses. Each silhouette image is represented as a d dimensional vector using the representation described in section 2. The whole collection of aligned cycles for all different people and views is arranged into order four tensor (4-way array) $\boldsymbol{\mathcal{D}}$ with dimensionality $N_s \times N_v \times N_p \times d$.

The data tensor $\boldsymbol{\mathcal{D}}$ can be decomposed to parameterize orthogonal style, view, and pose factors using higher-order singular value decomposition (HOSVD). Higher-order singular value decomposition (HOSVD) is a generalization of SVD for multilinear model analysis by [14, 27, 28]. Multilinear model is a generalization of linear model (one-factor models) and bilinear model (two-factor models) [25] into higher-order tensor decomposition (multi-factor models). The data tensor $\boldsymbol{\mathcal{D}}$ is decomposed to establish forth-order tensor using HOSVD which yields the decomposition

$$\boldsymbol{\mathcal{D}} = \boldsymbol{\mathcal{Z}} \times_{\mathbf{1}} \boldsymbol{S} \times_{\mathbf{2}} \boldsymbol{V} \times_{\mathbf{3}} \boldsymbol{P} \times_{\mathbf{4}} \boldsymbol{M}, \tag{3}$$

where $\boldsymbol{S}$, $\boldsymbol{V}$, $\boldsymbol{P}$, and $\boldsymbol{M}$ are orthogonal matrices with dimensionality $N_s \times N_s$, $N_v \times N_v$, $N_p \times N_p$, $d \times d$ corresponding to style, view, pose, and image orthogonal bases respectively. $\boldsymbol{\mathcal{Z}}$ is a core tensor with the same dimensionality as the data tensor $\boldsymbol{\mathcal{D}}$ which represents the interaction of the gait style, view, pose, and image pixel subspaces[2].

(a) Style clusters: (b) View vectors:

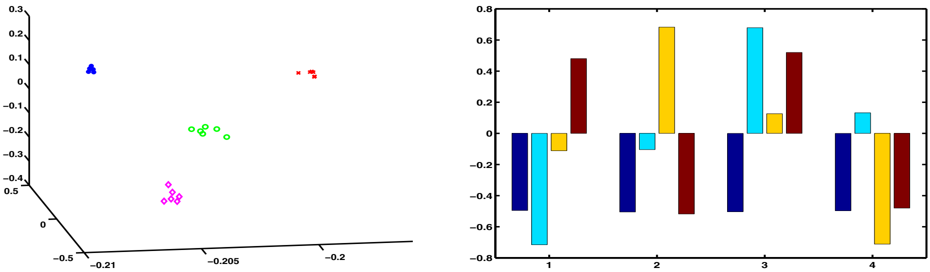

Fig. 2. Tensor analysis: 4 people with 6 cycles each from 4 different views. (a) First three style parameters for 6 gait cycles of each person. Each person's style shows good clustering within the person and good separation between different persons. (b)Four different view vectors, which are orthogonal to each others.

The orthogonal $N_s \times N_s$ matrix $\boldsymbol{S}$ spans the space of gait style parameters. In the style basis matrix $\boldsymbol{S} = [\boldsymbol{s}^1 \boldsymbol{s}^2 \cdots \boldsymbol{s}^s]^T$, each vector $\boldsymbol{s}^i$ represents a style parameter of person i as an N_s dimensional vector. This parameterization of the gait style independent of the view and body configuration is the basic feature we use in the recognition. Fig. 2 shows an example of the decomposition of gait style. We use 4 people from

[1] Each person can be represented by multiple cycles in the training data. So N_s represents the total number of cycles for all people.

[2] Reduced dimensional approximation can be achieved using higher-order orthogonal iteration method [15][29]

CMU-Mobogait data set with 6 cycles each from 4 different views to fit the model. As apparent in the figure, gait style parameters estimated from the different cycles of each person are clustered together in the style space.

Equation 3 can be rewritten as a generative model to synthesize gait cycles given any style vector $\boldsymbol{s}$ and view vector $\boldsymbol{v}$. This can be achieved by defining a new core tensor $\boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{Z}} \times_3 \boldsymbol{\mathcal{P}} \times_4 \boldsymbol{\mathcal{M}}$. Therefore, gait cycle images can be synthesized as $\boldsymbol{D}^{\boldsymbol{sv}}$ where

$$\boldsymbol{D}^{\boldsymbol{sv}} = \boldsymbol{\mathcal{B}} \times_1 \boldsymbol{s} \times_2 \boldsymbol{v} \tag{4}$$

3.2 Gait Style Estimation from Unknown View and Style

Given images $y_1, \cdots, y_k$ representing a full gait cycle from unknown view with k frames, estimation of gait style is required for person identification. First, the sequence is used to learn a generative model in the form of Equation 2 and then the model is used to re-sample p gait images, $\boldsymbol{i}_1\boldsymbol{i}_2 \cdots \boldsymbol{i}_p$, which are aligned with gait poses used in multilinear analysis. By stacking the gait images into a matrix $\boldsymbol{D} = [\boldsymbol{i}_1\boldsymbol{i}_2 \cdots \boldsymbol{i}_p]$, the estimation of style and view can be formulated as solving for $\boldsymbol{s}$ and $\boldsymbol{v}$ that minimize error

$$E(\boldsymbol{s}, \boldsymbol{v}) = ||\boldsymbol{D} - \boldsymbol{\mathcal{B}} \times_1 \boldsymbol{s} \times_2 \boldsymbol{v}||, \tag{5}$$

where $\boldsymbol{D}$ is $d \times N_p$ matrix. If the view vector $\boldsymbol{v}$ is known, we can obtain closed form solution for $\boldsymbol{s}$. This can be done by evaluating the product $\boldsymbol{\mathcal{H}} = \boldsymbol{\mathcal{B}} \times \boldsymbol{v}$ and unfolding the tensor $\boldsymbol{\mathcal{H}}$ into a matrix by style-mode, i.e., $\boldsymbol{H}_{(1)} = unfolding(\boldsymbol{\mathcal{H}}, 1)$. *Matrix unfolding* operation is explained in the appendix of this paper. The dimensions of $\boldsymbol{H}_{(1)}$ is $N_s \times (N_v \times N_p \times d)$. Solution for $\boldsymbol{s}$ can be obtained in closed form by solving the linear system $\boldsymbol{D} = \boldsymbol{H}_{(1)}{}^T \boldsymbol{s}$. Therefore estimation of $\boldsymbol{s}$ can be obtained by

$$\boldsymbol{s} = \left(\boldsymbol{H}_{(1)}{}^T\right)^+ \boldsymbol{D} \tag{6}$$

where $+$ is matrix pseudo-inverse operation using singular value decomposition (SVD). Similarly, we can analytically solve for $\boldsymbol{v}$ if the style vector $\boldsymbol{s}$ is known by forming a tensor $\boldsymbol{\mathcal{G}} = \boldsymbol{\mathcal{B}} \times_1 \boldsymbol{v}$ and forming its view-mode unfolding $\boldsymbol{G}_{(2)}$. Therefore, we can obtain the view as

$$\boldsymbol{v} = \left(\boldsymbol{G}_{(2)}{}^T\right)^+ \boldsymbol{D} \tag{7}$$

Iterative estimation of $\boldsymbol{s}$ and $\boldsymbol{v}$ using (6) and (7) leads to local minima for the error in (5). We can start initial style estimation by mean style $\boldsymbol{s} = (\sum_{i=1}^{N_s} s^i)/N_s$.

Given the estimated gait style vector $\boldsymbol{s}$, and different people's gait style vectors learned in the training, the recognition is a typical pattern classification problem. For the experimental results shown in this paper we used two simple classification approaches: Nearest Neighbor and Nearest class mean which shows very good recognition results. However, more sophisticated classification methods can be used to achieve even better results. The proposed framework can easily scale to include new people. Given a new person, theoretically, only one cycle from a single view is required to be able to solve for the person style parameter which can then be added to the trained database. In Section 4 we show experimental evaluation of the scalability and generalization of the model to learn style parameters from a single view and to recognize at different views.

(a) L_2 distance to each style vector in iteration: (b) L_2 distance to each view vector in iteration:

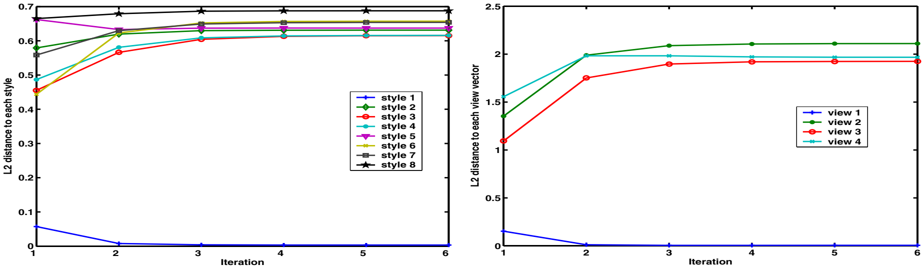

Fig. 3. Measurement of distance to style and view

Fig. 3 shows an example of the iterative estimation of view and gait style parameters. In this experiment we used 8 people with 4 different views from the Mobogait dataset to learn the model. The figure shows the change in the Euclidean distance to each mean style vector and mean view vector with the iterations. In this figure, a side view cycle for the first person was used for testing. It shows convergence to the correct style and view from the first iteration.

4 Experimental Results

We demonstrate the performance of the proposed algorithms on two databases: one is CMU mobo database and the other is USF-NIST gait database. In the preprocessing step, we applied median filter to remove noisy holes and spots. Bounding boxes which cover each person silhouettes were found and normalized to fixed size. Each silhouette shape is represented by a signed-distance function as described in section 2.

Experiment 1: Recognition of Gait in Different Speeds and Views: In this experiment we used CMU Mobo database, which has slow and fast walking sequences on a treadmill with six different views [7], to test gait recognition in different speeds and views. We chose a subset of 18 subjects which provided silhouettes for all different views and allowed finding proper bounding box for the subjects. Four different views (profile view, front-right view, front view, front-left view) were selected for multilinear gait analysis. Three cycles of slow walk for each person are used to learn the multilinear model parameters. In summary, the training data contains 18 people, 3 cycles each, from 4 views. Each person style is represented by the mean of the three style vectors obtained from three training cycles.

For evaluation we used three different slow-walk cycles and three fast-walk cycles for each of the 18 people with 4 views each. Overall there are 216 slow-walk evaluation cycles and 216 fast-walk evaluation cycles. For each evaluation cycle we estimate the view and the style of parmeteters of gait as described in Section 3.2. Finally, people are identified by finding closest style class mean. Table 1 shows the experiment result. For the slow-walk we achieve 100 % correct recognitions for all the views. For the fast-walk, we achieve around 90 % accuracy in average. The results shows fairly consistent recognition for all the different views. In both cases we achieve 100% view

Table 1. Gait recognition in different view and speed (CMU Data)

View class	slow walking sequences	fast walking sequences	Collins[3] (fast walking)
1(profile)	100%	88.9%	76%
2(front-right)	100%	88.9%	N/A
3(front)	100%	92.6%	100%
4(front-left)	100%	88.9%	N/A
Average	100%	90.0%	88%

identification. Even though we perform recognition for each cycle without knowing the view label, our results show better identification than template matching of key frames by Collins [3], shown in the forth column, which is tested for profile and front view separately using whole sequences.

Experiment 2: Generalization and Scalability Across Different Views: We evaluate the scalability of the proposed framework, i.e., Given a learned model, can we extend it to recognize a new person from different view points given that only one gait cycle from a single view is available for that person for training?

To evaluate this, we performed a new experiment by learning the model with a subset of subjects. Among 18 subjects, we learned the model using only eight subjects' slow walk sequences from 4 views. For the rest 10 subjects, only a single cycle data of slow walk from one view was given. We used this single view cycle to estimate gait style parameters. All the estimated style parameters are used as a database for recognition. The recognition is then evaluated using a test set consisting of 3 different slow-walk cycles and 3 fast-walk cycles from 4 views for all the 18 people.

Table 2 shows recognition results. We repeated the experiment by varying the view used in training for the 10 people with each single view cycle. Results show general identification capability to unknown views using style learned from a specific view. This clearly shows that the gait-style parameter is invariant to different view point. The identification performance varies across different views and the view used for training shows better performance on trained view class than others. Others, which do not learned style at all for the views,still, shows potentials for gait recognition. The performanace can be improved by using multiple cycles in the style estimation for given views.

Table 2. Gait recognition across different views(CMU Data)

View class	V1:slow	V2:slow	V3:slow	V4: slow	V1:fast	V2:fast	V3:fast	V4:fast
V1(profile)	96.3%	72.2%	53.7%	75.9%	53.7%	55.6%	40.7%	55.6%
V2(front-right)	72.2%	88.9%	59.3%	66.7%	53.7%	64.8%	48.2%	63.0%
V3(front)	51.9%	66.7%	90.9%	57.4%	50.0%	59.3%	92.6%	53.7%
V4(front-left)	59.3%	75.9%	70.7%	98.1%	46.3%	46.3%	55.6%	63.0%
Average(all)	69.9%	75.9%	68.7%	87.5%	50.9%	56.5%	59.3%	58.8%

Experiment 3: Recognition of Gait with Continuous Variation of Views (USF dataset): In this experiment we use NIST-USF Gait database [21] to evaluate perfor-

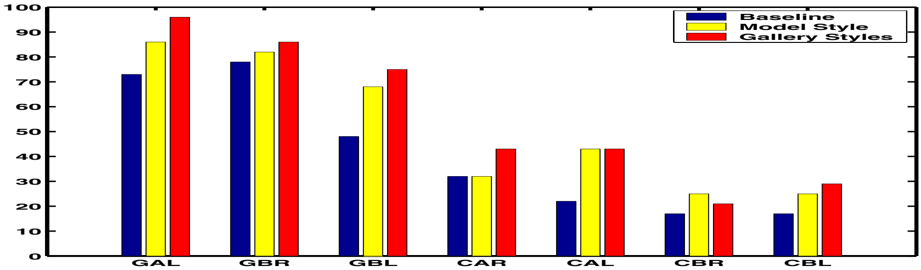

Fig. 4. Recognition result

Table 3. Comparison of Recognition with Baseline (USF Data)

Difference	Probe Set	Baseline	Nearest Mean	Nearest Neighbor	Kale [13]
View	GAL	73%	86%	96 %	89 %
Shoe	GBR	78%	82%	86 %	88 %
Shoe, view	GBL	48%	68%	75 %	68 %
Surface	CAR	32%	32%	43 %	35 %
Surface, shoe	CBR	22%	43%	43 %	28 %
surface, view	CAL	17%	25%	21 %	15 %
Surface, shoe, view	CBL	17%	25%	29 %	21 %

mance of gait recognition with continuous variation of the view due to the elliptical course that people used in capturing the database. We arbitrary select 28 people for a preliminary evaluation. We choose GAR, which is the gait sequence in grass surface, shoes type A, and right camera sequences, as a gallery set and tested by seven probe sets with variants in view, shoe and surface. Seven cycles were detected from the gallery sets and the probe sets. Three representative cycles of different views were selected from each sequence of gallery sets to learn the model.

For recognition we evaluated two classifiers for each estimated gait-style parameter for each test cycle: nearest style class mean (Model Style) and nearest neighbor style (Gallery styles). In both cases, we used majority vote from different test cycles to determine final person id. Results are shown in Table 3 and Fig. 4. Table 3 also shows recognition results reported in baseline evaluation [21] and recognition results reported using HMM by Kale *et al* in [13].

5 Conclusion

We presented a new framework to gait recognition which first uses a nonlinear generative model to re-sample gait sequences and then uses multilinear analysis to decompose view-invariant time-invariant gait parameters for identification. We showed promising human identification results in different views and speeds in CMU dataset. In USF dataset, which has continues view point variations within each probe set, also shows improvement in identification using the proposed view invariant iterative style estimation framework. We used very simple classification algorithms for identification from the estimated gait style parameters. Recognition can be further improved by employing more sophisticated classification algorithms such as support vector machine (SVM) using style vectors. In the future we plan to report gait recognition for larger data sets.

Appendix

Matrix unfolding operation: Given an r-order tensor $\mathcal{A}$ with dimensions $N_1 \times N_2 \times \cdots \times N_r$, the mode-$n$ matrix unfolding, denoted by $unfolding(\boldsymbol{\mathcal{A}}, n)$, is flattening $\mathcal{A}$ into a matrix whose column vectors are the mode-n vectors [14, 27]. Therefore, the dimension of the unfolded matrix $\boldsymbol{A_{(n)}}$ is $N_n \times (N_1 \times N_2 \times \cdots N_{n-1} \times N_{n+1} \times \cdots N_r)$.

References

1. C. BenAdbelkader, R. Cutler, and L. Davis. Stride and cadence as a biometric in automatic person identification and verification. In *Proc. FGR*, pages 357–363, 2002.
2. B. Bhanu and J. Han. Individual recognition by kinematic-based gait analysis. In *Proc. ICPR*, volume 3, pages 343– 346, 2002.
3. R. Collins, R. Gross, and J. Shi. Silhouette-based human identification from body shape and gait. In *Proc. FGR*, pages 351–366, 2002.
4. D. Cunado, M. S. Nixon, and J. Carter. Automatic extraction and description of human gait models for recognition purposes. *Computer Vision and Image Understanding*, 90:1–41, 2003.
5. A. Elgammal. Nonlinear generative models for dynamic shape and dynamic appearance. In *Proc. Int. Workshop GMBV*, 2004.
6. A. Elgammal and C.-S. Lee. Separating style and content on a nonlinear manifold. In *Proc. CVPR*, pages 478–485, 2004.
7. R. Gross and J. Shi. The cmu motion of body (mobo) database. Technical Report CMU-RI-TR-01-18, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, June 2001.
8. J. Han and B. Bhanu. Statistical feature fusion for gait-based human recognition. In *Proc. CVPR*, pages 842–847, 2004.
9. Q. He and C. Debrunner. Individual recognition from periodic activity using hidden markov models. In *In IEEE Workshop on Human Motion*, pages 47–52, 2000.
10. P. Huang, C. Haris, and M. Nixon. Recogising humans by gait via parametric canonical space. *Artificial Intelligence in Engineering*, 13:359–366, 1999.
11. A. Y. Johnson and A. F. Bobick. A multi-view method for gait recognition using static body parameters. In *Proc. AVBPA*, pages 301–311, June 2001.
12. A. Kale, A. K. R. Chowdhury, and R. Chellappa. Towards a view invariant gait recognition algorithm. In *Proc. on Advanced Video and Signal Based Surveillance*, pages 143–150, 2003.
13. A. Kale, A. Sundaresan, A. N. Rajagopalan, N. P. Cuntoor, A. K. Roy-Chowdhury, V. Kruger, and R. Chellappa. Identification of human using gait. *IEEE Trans. Image Processing*, 13(9):1163–1173, 2004.
14. L. D. Lathauwer, B. de Moor, and J. Vandewalle. A multilinear singular value decomposiiton. *SIAM Journal On Matrix Analysis and Applications*, 21(4):1253–1278, 2000.
15. L. D. Lathauwer, B. de Moor, and J. Vandewalle. On the best rank-1 and rank-(r1, r2, ..., rn) approximation of higher-order tensors. *SIAM Journal On Matrix Analysis and Applications*, 21(4):1324–1342, 2000.
16. C.-S. Lee and A. Elgammal. Gait style and gait content: Bilinear model for gait recogntion using gait re-sampling. In *Proc. FGR*, pages 147–152, 2004.

17. L. Lee, G. Dalley, and K. Tieu. Learning pedestrian models for silhouette refinement. In *Proc. ICCV*, pages 663–670, 2003.
18. Y. Liu, R. Collins, and Y. Tsin. Gait sequence analysis using frienze patterns. In *Proc. ECCV*, pages 657–671, 2002.
19. Z. Liu and S. Sarkar. Simplest representation yet for gait recognition: Averaged silhouette. In *Proc. ICPR*, pages 211–214, 2004.
20. H. Murase and R. Sakai. Moving object recognition in eigenspace representation: gait analysisand lip reading. *Pattern Recognition Letters*, 17:155–162, 1996.
21. P. J. Phillips, S. Sarkar, I. Robledo, P. Grother, and K. Bowyer. Baseline results for the challenge problem of human id using gait analysis. In *Proc. FGR*, pages 137–142, 2002.
22. T. Poggio and F. Girosi. Networks for approximation and learning. *Proceedings of the IEEE*, 78(9):1481–1497, 1990.
23. G. Shakhnarovich, L. Lee, and T. Darrell. Integrated face and gait recognition from multiple views. In *Proc. CVPR*, pages 439–446, 2001.
24. R. Tanawongsuwan and A. Bobick. Modelling the effects of walking speed on appearance-based gait recognition. In *Proc. CVPR*, pages 783–790, 2004.
25. J. B. Tenenbaum and W. T. Freeman. Separating style and content with biliear models. *Neural Computation*, 12:1247–1283, 2000.
26. D. Tolliver and R. T. Collins. Gait shape estimation for identification. In *Proc. AVBPA*, pages 734–742, 2003.
27. M. A. O. Vasilescu. Human motion signatures: Analysis, synthesis, recogntion. In *Proc. ICPR*, volume 3, pages 456–460, 2002.
28. M. A. O. Vasilescu and D. Terzopoulos. Multilinear analysis of image ensembles: Tensorfaces. In *Proc. ECCV*, pages 447–460, 2002.
29. M. A. O. Vasilescu and D. Terzopoulos. Multilinear subspace analysis of image ensembles. In *Proc. CVPR*, 2003.
30. L. Wang, T. Tan, H. Ning, and W. Hu. Silhouette analysis-based gait recognition for human identification. *IEEE Trans. PAMI*, 25(12):1505–1518, 2003.
31. G. Zhao, R. Chen, G. Liu, and H. Li. Amplitude spectrum-based gait recognition. In *Proc. FGR*, pages 23–28, 2004.