Cross-View Gait Recognition Using View-Dependent Discriminative Analysis

Al Mansur, Yasushi Makihara, Daigo Muramatsu and Yasushi Yagi
Osaka University
{mansur,makihara,muramatsu,yagi}@am.sanken.osaka-u.ac.jp

Abstract

Gait is a unique and promising behavioral biometrics which allows to authenticate a person even at a distance from the camera. Since a matching pair of gait features are often drawn from different views due to differences in camera position/attitude and walking directions in the real world, it is important to cope with cross-view gait recognition. In this paper, we propose a discriminative approach to cross-view gait recognition using view-dependent projection matrices, unlike the existing discriminant approaches which utilize only a single common projection matrix for different views. We demonstrated the effectiveness of the proposed method through cross-view gait recognition experiments with two publicly available gait datasets. In addition, since the success of the discriminant analysis relies on the training sample size, we show the effect of transfer learning across two gait datasets as well as provide the rigorous sensitivity analysis of the proposed method against the number of training subjects ranging from 10 to approximately 1,000 subjects.

1. Introduction

Biometrics [17] are person authentication methods by individual physiological or behavioral features. Up to now, many of the modalities such as DNA, fingerprint, palm print, iris, face, signature, keystroke, and gait, have been proposed. Among this large body of biometric modalities, gait [38] is particularly regarded as a promising technology since a gait feature of a target person can be captured even at a distance from a camera without subject's awareness or cooperation [2] under unconstrained situations. Gait recognition is therefore expected to be applied to wide-area surveillance, criminal investigation, and forensics [6, 14]. In fact, automatic gait recognition has been admitted as an evidence of a burglar in UK courts [6].

Approaches to gait recognition are mainly categorized into two groups: model-based approaches [1, 5, 9] and appearance-based approaches [12, 26, 40, 46]. Although the model-based approaches can extract gait motion (or pose sequence) and static part (or body shape) separately, they often suffer from model fitting error and relatively high computational cost. Appearance-based approaches are therefore currently dominant in the gait recognition community and achieve better performances than the model-based approaches in general.

However, appearance-based approaches are largely affected by various factors (e.g., views, clothing, carriages, walking speeds, shoes, surfaces, and physical and psychological conditions) [7, 40]. View covariate is one of the significant issues which gait recognition in the wild often faces, because CCTVs in the public or private space are set up at various positions and attitudes and also people changes their walking direction depending on the destinations. The view difference naturally induces considerable changes of the appearance-based gait features, and hence it makes gait recognition much harder, as reported in [3, 50].

A straightforward approach to cope with such view variations in gait recognition is to collect multi-view gallery (database) for the target person, and to match a probe (query) with the gallery whose view is the closest to that of the probe, which results in a simple nearest neighbor classifier. However, view inconsistency may still remain to a greater or lesser extent between the probe and the nearest gallery as long as the gallery is captured from a set of discrete multiple views. A more sophisticated approach to overcome this difficulty is to reconstruct 3D gait volumes from multi-view synchronous gait silhouettes by visual intersection method or space curving and project it to the same view as that of the probe, which facilitates matching the gait features under exactly the same view [28, 41]. In any cases, it is relatively difficult to acquire multi-view synchronous gait silhouette sequences for uncooperative subjects such as a perpetrator or a suspect in forensic scenarios, and the 3D gait volume-based approaches are therefore inappropriate for the purpose of surveillance and criminal investigation.

On the other hand, discriminative approaches aim at extracting discriminative view-invariant latent space via various discriminant analyses such as principal component analysis (PCA), linear discriminant analysis (LDA) [10], uncorrelated discriminant simplex analysis [32], view invariant discriminative projection [13], sparse Grassmannian locality preserving discriminant analysis [8] and ranking-based support vector machine (SVM) [35]. In addition, Liu et al. [30] extend the discriminative framework to multi-view set matching. All of these approaches define a single common matrix to project gait features from different views into a common discriminative space. A major advantage of these methods is that they do not require view estimation process before applying the projection matrix, while the generative approaches [23, 33] require the view information in advance. On the other hand, using the single matrix leads to discarding meaningful information which is exploited by the generative approaches (e.g., a gait feature value at a specific position from one view is transferred into that at another specific position from another view).

In order to represent such a feature transfer information across views as well as to enhance the discrimination capability at the same time, discriminative analysis using domain-dependent projection matrices has been proposed and reported superior performances in the other research fields. Such approaches include discriminative canonical correlation [21], partial least squares (PLS) [42], generalized multi-view linear discriminant analysis (GMLDA) [43], multi-view discriminant analysis (MvDA) [20], and heterogeneous transfer discriminant-analysis for canonical correlation (HTDCC). As applications, discriminative canonical correlation, PLS, GMLDA and MvDA have been used for cross-view (cross-pose) face recognition and HTDCC has been used for cross-view action recognition.

We therefore propose to use such a multi-domain discriminative analysis approach to cross-view gait recognition where domain-dependent projection matrices are computed and used for projection. We choose MvDA [20] because it provides an efficient generalized eigenvalue based solution and is free from parameter tuning. The contributions of this paper are summarized as follows.

Cross-view gait recognition using domain-dependent discriminative projection matrices

Whereas the existing approaches to cross-view gait recognition exploit only either a feature transfer information in the generative approach or discriminant analysis with a single common projection matrix, we bring domain-dependent discriminative projection matrices into cross-view gait recognition problem. We show that the proposed approach achieves state-of-the-art performance against two publicly available multi-view gait databases: CASIA Gait database B (CASIA-B) [49] with large view variations but with limited number of subjects and OU-ISIR Gait Database, the large population data sets (OU-LP) [15] with large number of subjects but with limited view variations.

Transferring domain-dependent discriminative projection matrices across different gait databases

Since success of the machine learning-based approaches including the proposed method, depend on the variety of the training samples, we employ a transfer learning approach to compensate limited subject variations in the CASIA-B with the large subject variations in the OU-LP.

Sensitivity analysis on the number of training subjects

In addition to transfer learning experiment, we provide a rigorous sensitivity analysis results on the number of subjects ranging from 10 to approximately 1,000 subjects. This reveal the generalization capability of the proposed approach and provides useful insight to future venue of cross-view gait recognition studies.

2. Related work

Approaches to cross-view gait recognition can be summarized into two groups: geometric approaches and machine learning (or example-based) approaches.

Geometric approaches often assume that a person (a 3D object) is well approximated by a planar object on a sagittal plane[1] and/or that the weak perspective geometry is well satisfied with respect to a target person. Jean et al. [18] and Goffredo et al. [11] extract foot and head trajectories in the 2D image plane and project them into the sagittal plane, which are equivalent to trajectories in side view. Kale et al. [19] generates gait silhouettes themselves from canonical view (side view) in a similar way. Kusakunniran et al. [25] employ domain transformations to transform a sequence of gait silhouettes obtained from a certain view to a common canonical view. However, these approaches do not work well in case where an angle between the sagittal plane of the person and an image plane is relatively large and/or a case where the distance between the person and the camera is small compared with the size of the person, namely, when a weak perspective projection does not hold.

On the other hand, machine learning-based approaches exclude such geometric assumptions and exploit the statistics of the gait features between different views obtained from the training data (examples) instead. Unlike the straightforward multi-view gait recognition approaches, these approaches exploit multi-view gait features collected from cooperative training subjects (non-recognition target such as students in a laboratory, colleagues in a research institute, volunteers who join the data collection session) and find a transformation model or discriminative feature space using the training subjects. These kinds of approaches are further divided into two subcategories: (1) generative approaches and (2) discriminative approaches.

Generative approaches aims at generating a gait feature at one view (e.g., probe) from that at another view (e.g., gallery) so as to match a pair of gallery and probe under

[1] A vertical plane dividing an animal into left and right parts.

the same view. This family mainly employs a framework of a view transformation model (VTM) which represents how a gait feature from one view is transferred into that from another view. Several approaches have been used to construct the VTM such as matrix factorization using singular value decomposition (SVD) [33, 36, 37, 45], canonical correlation analysis (CCA) [29], regression [23, 24], and neural network [22].

Moreover, while the view transformation is limited between a pair of discrete views included in the training set in the above approaches, arbitrary VTM [37] enables view transformation between a pair of arbitrary views by constructing 3D gait volumes of the training subjects in advance and by projecting them into a specific pair of gallery and probe of the test subjects, which is analogous to [16, 28, 41]. An underlying assumption for this family of approaches is that gait features from different views are correlated each other to some extent and that such correlation is expressed in a common way among the training and test subjects, and hence the constructed VTMs are applicable not only to a training subject but also to a test subject.

On the other hand, if a gait feature of a test subject is not well represented by a set of gait features of the training subjects, transformation errors may become large. To cope with this, Muramatsu et al. [36] proposed a quality-dependent VTM (QVTM) which exploits the projection error from a gait feature of a test subject to a joint subspace spanned by the gait features of the training subjects, as a sort of quality measure and normalizes dissimilarity scores based on the quality measure. In other words, it adaptively sets the acceptance threshold for each of test subjects considering the degree how the test subject is well represented by the training subjects.

Since the generative approaches exploit view-dependent projection matrices or regression models, they require a view label for each of the gait feature when transforming the gait feature. In addition, because the generative approaches basically aims at minimizing errors between transformed gait feature and original gait features, they do not guarantee the optimality in terms of discrimination capability.

3. Multi-view discriminant analysis

We include a brief description of MvDA formulation to make this paper self-contained. MvDA finds a common discriminative subspace where within-class variation is minimized and between-class variation is maximized. In other words, MvDA finds v linear transforms $(w_1, \cdots, w_v)$ and j-th view is projected using w_j. Let $X^j = \{x_{ijk} | i = 1, \cdots, c; k = 1, \cdots, n_{ij}\}$ are the samples at j-th view, where c is the number of classes (subjects). $x_{ijk} \epsilon R^{d_j}$ is d_j-dimensional sample and n_{ij} is the number of samples for the i-th class at view j. Given the projection directions, let $Y = \{y_{ijk} = w_j^T x_{ijk} | i = 1, \cdots, c; j = 1, \cdots, v; k = 1, \cdots, n_{ij}\}$ is the projected samples in the common subspace.

In the discriminative subspace, the between-class variation from all views is maximized and the within-class variation from all views is minimized. Therefore, the objective function can be formulated using a generalized Rayleigh quotient as:

$$(w_1^*, \cdots, w_v^*) = \arg \max_{w_1, \cdots, w_v} \frac{Tr(S_B^y)}{Tr(S_W^y)}, \quad (1)$$

where the between-class scatter matrix S_B^y and within-class scatter matrix S_W^y are defined as

$$S_B^y = \sum_{i=1}^{c} n_i (\mu_i - \mu)(\mu_i - \mu)^T, \quad (2)$$

$$S_W^y = \sum_{i=1}^{c} \sum_{j=1}^{v} \sum_{k=1}^{n_{ij}} (y_{ijk} - \mu_i)(y_{ijk} - \mu_i)^T, \quad (3)$$

where $\mu_i = \frac{1}{n_i} \sum_{j=1}^{v} \sum_{k=1}^{n_{ij}} y_{ijk}$ is the mean of the i-th class (containing n_i samples) and $\mu = \frac{1}{n} \sum_{i=1}^{c} \sum_{j=1}^{v} \sum_{k=1}^{n_{ij}} y_{ijk}$ is the mean of all samples in the common subspace. Eq. 1 can be solved through generalized eigenvalue solution. For details please see [20].

4. Experiments

We test MvDA approach under cross-view gait recognition setting using OU-ISIR Large Population Dataset (OU-LP) [15] and CASIA dataset B [49].

OU-LP dataset: This dataset includes gait sequences from 1912 subjects. All the sequences are captured indoor with controlled lighting and background. For each subject, it contains gait image sequences at view angles, $\theta = \{55°, 65°, 75°, 85°\}$ and there are 2 sequences per view. In each sequence, background subtraction-based graph-cut segmentation [34] is applied to extract gait silhouette images and it is scaled to 64×44 pixel-sized images. We compute gait periods in each sequence and then compute gait energy image (GEI) [31]. GEIs at different views for a particular subjects is shown in Fig. 1. For dimensionality reduction, we use PCA and chose the eigenvectors so that 95% of the variance is preserved. The dimension reduced GEIs are used as feature. We equally divide this dataset into two sets[2] and one set is used for training while the other is used for testing.

CASIA-B: This dataset contains gait sequences from 124 subjects captured under three different covariate conditions (carrying, clothing, and view angle). However, to suit our problem setting, we used only the normal walking sequences (where a person does not carry a bag or wear a

[2]Available at: http://www.am.sanken.osaka-u.ac.jp/~mansur/files/list_train_test.txt

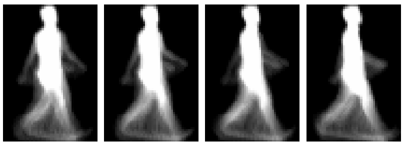


Figure 1. GEIs from OU-LP dataset at four different views.

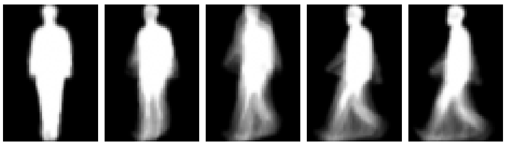


Figure 2. GEIs from CASIA-B at five different views.

coat) under different view angles. We use 62 subjects for training and remaining 62 subjects for testing with five random training and test sets. This dataset provides 11 different view angles ranging from frontal view (0°) to back view (180°) at an interval of 18°. We used view angles $\theta = \{0°, 18°, 36°, 54°, 72°\}$. In each view, each subject has 6 normal gait sequences. Example GEIs at five different views can be seen in Fig. 2. Here the GEI size is 64×44 pixel and again we used reduced dimensional GEIs as features.

Performance measures: MvDA is evaluated in both identification and verification scenarios. In the identification scenario, cumulative match characteristic (CMC) plot [27] is used which measures how well the identification system ranks each of the gallery subject with respect to the probe. In addition, we use Rank-1 identification rate picked up from the CMC for comparison. On the other hand, Receiver Operating Characteristics (ROC) curve [39] and equal error rate (EER) picked from the ROC curve is employed for comparison in verification scenario. ROC curve indicates a tradeoff between false rejection and false acceptance when the receiver changes an acceptance threshold.

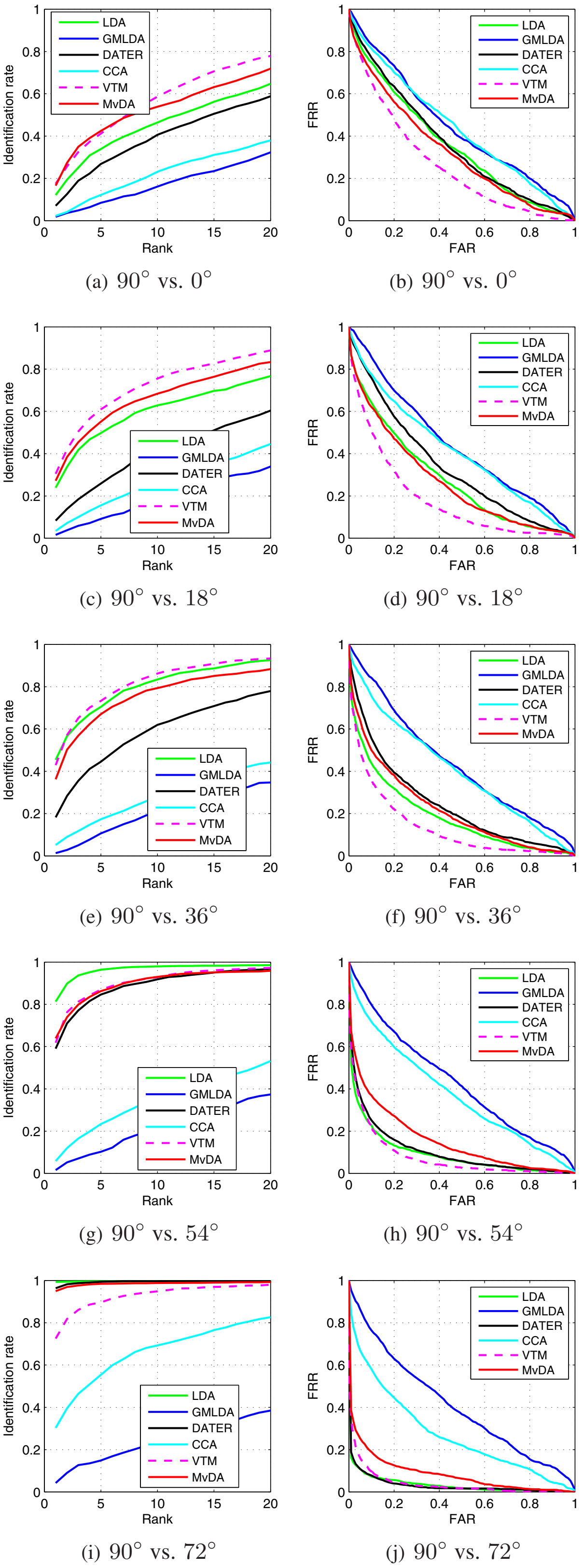
Benchmarks: We adopt both discriminative and generative approaches as benchmark methods. Discriminative approaches include LDA [10], GMLDA [43], DATER [48], CCA [29] and the generative approaches include VTM [33]. In addition we compare the results obtained using spatial transformation (ST) method (a geometric approach) [19]. We implemented GMLDA, DATER and ST while LDA and CCA codes were obtained from [44] and MATLAB toolbox respectively. VTM results were provided by the authors. For GMLDA, we set the parameters α and μ as used by the authors.

Gait recognition: Both OU-LP and CASIA-B were used for this experiment. In both datasets, the test sets contain gait sequences at different views. We used one view as gallery and the remaining views as probe. For OU-LP, we considered cross view gait recognition under fixed gallery view at 85° and several probe views at $\{55°, 65°, 75°\}$. In case of CASIA-B, we considered gallery at 90°, and probe at $\{0°, 18°, 36°, 54°, 72°\}$. In the training phase, two different views (corresponding to the test situation) v_1 and v_2 were used to find the view specific projection directions w_1 and w_2, respectively. During test phase, gallery and probe GEIs were projected using the corresponding projection directions. More specifically, GEIs in v_1 and v_2 were projected using w_1 and w_2, respectively. Euclidean distance-based 1-NN matching was used for matching between gallery and probe GEIs. Results are shown in Figs. 3, 4 and Table 1.

We can interpret the results as follows. In case of small training set (CASIA-B): (1) LDA performs the best in case of small view difference. Because GEIs are similar to some extent under such small view difference and hence even LDA with a single projection matrix can efficiently absorb the cross-view variations. In addition, LDA receives twice training samples in essence compared with MvDA, which relaxes generalization errors. (2) In case of extremely large view difference, a single projection matrix cannot project the data well for good classification. As a result, performance of LDA deteriorates and MvDA outperforms LDA. In addition, as reported in [4], discriminative approaches require plentiful of training data to achieve better performance. Therefore VTM (a generative approach) shows relatively better performance than MvDA. On the other hand, in case of large training set (OULP), MvDA outperforms the other methods because it receives sufficient training data in each domain.

Cross-dataset gait recognition: To overcome the problem of small training set in CASIA-B, we attempt to increase the training data using OU-LP dataset. As the walking directions are different in these datasets, we simply flip the GEIs of one dataset. Unlike the previous experiment, we consider both CASIA-B and OU-LP datasets for learning the projection directions and apply this projection directions on test data from CASIA-B. The training set contains 62 subjects from CASIA-B and 956 subjects from OU-LP dataset. On the other hand the test set contains 62 subjects from CASIA-B with gallery view at 90° and probe views at $\{54°, 72°\}$. However, available views in CASIA-B and OU-LP are not exactly same. For view correspondences, we pair 54° and 72° views of CASIA-B with 55° and 75° views of OU-LP, respectively.

Cross-dataset results of MvDA and other benchmark methods are shown in Fig. 5. It can be noticed that performances of LDA, DATER and VTM methods degraded with cross-dataset learning (compare with Fig. 3). On the other hand, performances of MVDA, GMLDA and CCA improved compared to training using CASIA-B alone. In particular, MvDA performs the best specially at the large



(a) 90° vs. 0° (b) 90° vs. 0°

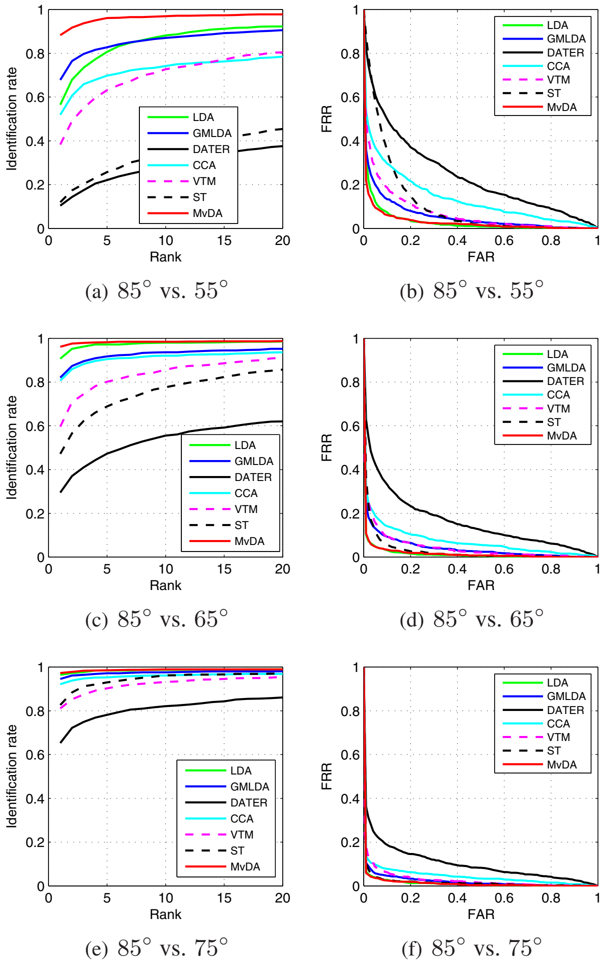
(c) 90° vs. 18° (d) 90° vs. 18°

(e) 90° vs. 36° (f) 90° vs. 36°

(g) 90° vs. 54° (h) 90° vs. 54°

(i) 90° vs. 72° (j) 90° vs. 72°

Figure 3. Results for CASIA-B. Left: CMC, right: ROC.



(a) 85° vs. 55° (b) 85° vs. 55°

(c) 85° vs. 65° (d) 85° vs. 65°

(e) 85° vs. 75° (f) 85° vs. 75°

Figure 4. Results for OU-LP dataset. Left: CMC, right: ROC.

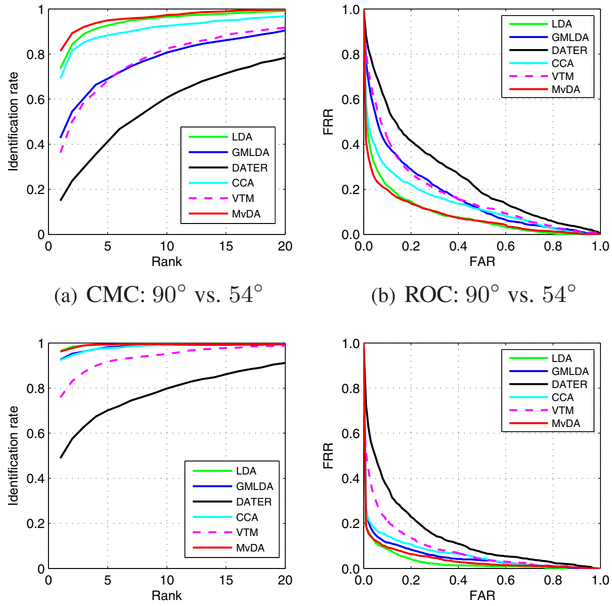
view difference between gallery and probe which is very desirable. For example, MvDA shows approximately 5% improvement in Rank-1 identification rate over LDA in 90° vs. 54°.

Sensitivity analysis: Finally, we show sensitivity analysis of MvDA using OU-LP dataset by varying the number of training subjects from 10 to 956 (see Fig. 6). We set the gallery at 85°, and probe view at $\{55°, 65°, 75°\}$ and identification rates (Rank-1 and Rank-5) and EER are used as performance measures. In case of CASIA-B, when MvDA is trained with 24 subjects and tested on 100 subjects obtained rank-1 identification rates are 0.06 for 90° vs. 72° and 0.05 for 90° vs. 54° which is quite low compared to [25] (0.96 and 0.68 respectively). This results indicates that generalization capability of MvDA is quite limited under small training set.

MvDA is little sensitive to the dimension used for projection. We used min((number of class - 1), dimension of data) as the MvDA dimension. When the number of training subject is $<$ 200, dimension is decided by the number

	Rank-1 identification rate								EER							
	CASIA-B					OULP			CASIA-B					OULP		
	0°	18°	36°	54°	72°	55°	65°	75°	0°	18°	36°	54°	72°	55°	65°	75°
LDA	0.12	0.24	**0.45**	**0.81**	**0.99**	0.56	0.91	0.96	0.39	0.35	0.26	0.16	**0.08**	0.08	**0.05**	**0.04**
GMLDA	0.02	0.02	0.01	0.02	0.04	0.68	0.82	0.95	0.44	0.44	0.44	0.46	0.43	0.12	0.09	0.05
DATER	0.07	0.08	0.18	0.59	0.96	0.10	0.29	0.65	0.4	0.37	0.3	0.18	**0.08**	0.3	0.22	0.16
CCA	0.02	0.03	0.05	0.06	0.30	0.52	0.81	0.92	0.46	0.44	0.44	0.41	0.32	0.21	0.13	0.08
VTM	**0.17**	**0.30**	0.43	0.62	0.83	0.38	0.60	0.81	**0.32**	**0.24**	**0.21**	**0.15**	0.1	0.14	0.09	0.07
ST	-	-	-	-	-	0.12	0.47	0.83	-	-	-	-	-	0.18	0.08	0.05
MvDA	**0.17**	0.27	0.36	0.64	0.95	**0.88**	**0.96**	**0.97**	0.38	0.33	0.29	0.24	0.15	**0.07**	**0.05**	**0.04**

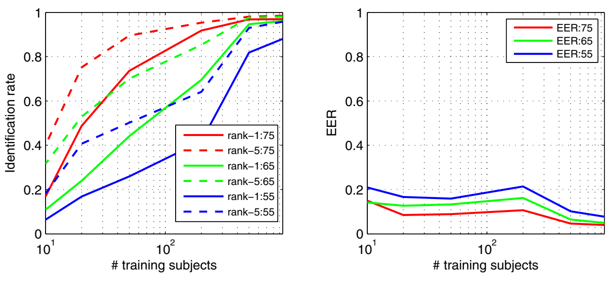
Table 1. Comparison of Rank-1 identification rates and EER for CASIA-B and OU-LP datasets.



(a) CMC: 90° vs. 54° (b) ROC: 90° vs. 54°

(c) CMC: 90° vs. 72° (d) ROC: 90° vs. 72°

Figure 5. Results for cross-dataset experiments on CASIA-B.



(a) (b)

Figure 6. Performance of MvDA under different number of training subjects: (a) identification rates at different ranks (b) equal error rate.

of classes. However, beyond 200, MvDA dimension is decided by the dimension of data. Therefore there is a switching at 200 and this is the reason for non-monotonicity in the sensitivity profiles. If we use the same dimension for all the experiments, non-monotonicity disappears.

5. Conclusions

In this paper, we proposed a cross-view gait recognition method using view-dependent projection matrices via MvDA framework. We confirmed that the proposed method achieved state-of-the-art results through experiments with CASIA-B with large view variations as well as OU-LP data set with large subjects variations. We showed that MvDA can get advantage from transfer learning across two gait databases and also provided the rigorous sensitivity analysis with respect to the number of training subjects to investigate the generalization capability of the proposed method.

The results of the sensitivity analysis revealed that generalization errors of the proposed method increases as the number of training subjects decreases and the proposed method works quite well with sufficient number of training subjects. It is reported that recent discriminant analysis approaches which keep the higher-order data structure (e.g., matrix or tensor) [48] can overcome the small sample size problem as well as curse of dimensionality problem. In addition, they are shown effective in gait recognition [47]. We plan to enhance the generalization capability of the proposed method through tensor representation [48] in future.

Acknowledgment

This work was supported by JSPS Grant-in-Aid for Scientific Research(S) 21220003, "R & D Program for Implementation of Anti-Crime and Anti-Terrorism Technologies for a Safe and Secure Society", Strategic Funds for the Promotion of Science and Technology of the Ministry of Education, Culture, Sports, Science and Technology, the Japanese Government, and the JST CREST "Behavior Understanding based on Intention-Gait Model" project.

References

[1] G. Ariyanto and M. Nixon. Marionette mass-spring model for 3d gait biometrics. In *Proc. of the 5th IAPR International*

Conference on Biometrics, pages 354–359, March 2012.

[2] K. Bashir, T. Xiang, and S. Gong. Gait recognition without subject cooperation. *Pattern Recognition Letters*, 31(13):2052–2060, Oct. 2010.

[3] C. BenAbdelkader. *Gait as a Biometric For Person Identification in Video*. PhD thesis, Maryland Univ., 2002.

[4] C. M. Bishop and J. Lasserre. Generative or Discrimative? Getting the Best of Both Worlds. In *Bayesian Statistics 8*, pages 3–24, 2007.

[5] A. Bobick and A. Johnson. Gait recognition using static activity-specific parameters. In *Proc. of the 14th IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 423–430, 2001.

[6] I. Bouchrika, M. Goffredo, J. Carter, and M. Nixon. On using gait in forensic biometrics. *Journal of Forensic Sciences*, 56(4):882–889, 2011.

[7] I. Bouchrika and M. Nixon. Exploratory factor analysis of gait recognition. In *Proc. of the 8th IEEE International Conference on Automatic Face and Gesture Recognition*, pages 1–6, Amsterdam, The Netherlands, Sep. 2008.

[8] T. Connie, G. K. O. Michael, and A. T. B. Jin. Grassmannian locality preserving discriminant analysis to view invariant gait recognition with image sets. In *Proceedings of the 27th Conference on Image and Vision Computing New Zealand*, IVCNZ '12, pages 400–405, 2012.

[9] D. Cunado, M. Nixon, and J. Carter. Automatic extraction and description of human gait models for recognition purposes. *Computer Vision and Image Understanding*, 90(1):1–41, 2003.

[10] R. Duda, P. Hart, and D. Stork. *Pattern Classification*. Wiley, 2001.

[11] M. Goffredo, I. Bouchrika, J. Carter, and M. Nixon. Self-calibrating view-invariant gait biometrics. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 40(4):997–1008, 2010.

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

[13] M. Hu, Y. Wang, Z. Zhang, J. Little, and D. Huang. View-invariant discriminative projection for multi-view gait-based human identification. *Information Forensics and Security, IEEE Transactions on*, 8(12):2034–2045, Dec 2013.

[14] H. Iwama, D. Muramatsu, Y. Makihara, and Y. Yagi. Gait verification system for criminal investigation. *IPSJ Transactions on Computer Vision and Applications*, 5:163–175, Oct. 2013.

[15] H. Iwama, M. Okumura, Y. Makihara, and Y. Yagi. The ou-isir gait database comprising the large population dataset and performance evaluation of gait recognition. *IEEE Transactions on Information Forensics and Security*, 7(5):1511–1521, Oct. 2012.

[16] Y. Iwashita, R. Baba, K. Ogawara, and R. Kurazume. Person identification from spatio-temporal 3d gait. In *Proc. of the Int. Conf. on Emerging Security Technologies*, pages 30–35, September 2010.

[17] A. K. Jain, P. Flynn, and A. A. Ross. *Handbook of Biometrics*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2007.

[18] F. Jean, R. Bergevin, and A. B. Albu. Computing and evaluating view-normalized body part trajectories. *Image and Vision Computing*, 27(9):1272–1284, Aug. 2009.

[19] A. Kale, A. Roy-Chowdhury, and R. Chellappa. Towards a view invariant gait recognition algorithm. In *Proc. of IEEE Conf. on Advanced Video and Signal Based Surveillance*, pages 143–150, 2003.

[20] M. Kan, S. Shan, H. Zhang, S. Lao, and X. Chen. Multi-view discriminant analysis. In *Proc. European Conference on Computer Vision*, volume 7572 of *Lecture Notes in Computer Science*, pages 808–821. Springer, 2012.

[21] T.-K. Kim, J. Kittler, and R. Cipolla. Learning discriminative canonical correlations for object recognition with image sets. In *Proc. European Conference on Computer Vision*, volume 3953 of *Lecture Notes in Computer Science*, pages 251–262, 2006.

[22] W. Kusakunniran, Q. Wu, J. Zhang, and H. Li. Multi-view gait recognition based on motion regression using multilayer perceptron. In *Prof. of the 20th Int. Conf. on Pattern Recognition*, pages 2186–2189, Istanbul, Turkey, Aug. 2010.

[23] W. Kusakunniran, Q. Wu, J. Zhang, and H. Li. Support vector regression for multi-view gait recognition based on local motion feature selection. In *Proc. of IEEE conference on Computer Vision and Pattern Recognition 2010*, pages 1–8, San Francisco, CA, USA, Jun. 2010.

[24] W. Kusakunniran, Q. Wu, J. Zhang, and H. Li. Gait recognition under various viewing angles based on correlated motion regression. *IEEE Transactions on Circuits and Systems for Video Technology*, 22(6):966–980, 2012.

[25] W. Kusakunniran, Q. Wu, J. Zhang, Y. Ma, and H. Li. A new view-invariant feature for cross-view gait recognition. *IEEE Transactions on Information Forensics and Security*, 8(10):1642–1653, 2013.

[26] 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*, 44:973–987, April 2011.

[27] H. Lee, S. Hong, and E. Kim. An efficient gait recognition with backpack removal. *EURASIP J. Adv. Sig. Proc.*, 2009, 2009.

[28] L. Lee. *Gait Analysis for Classification*. PhD thesis, Massachusetts Institute of Technology, 2002.

[29] N. Liu, J. Lu, and Y.-P. Tan. Joint subspace learning for view-invariant gait recognition. *IEEE Signal Processing Letters*, 18(7):431–434, 2011.

[30] N. Liu, J. Lu, G. Yang, and Y.-P. Tan. Robust gait recognition via discriminative set matching. *Journal of Visual Communication and Image Representation*, 2013, 2013.

[31] Z. Liu and S. Sarkar. Simplest representation yet for gait recognition: Averaged silhouette. In *Proc. of the 17th International Conference on Pattern Recognition*, volume 1, pages 211–214, Aug. 2004.

[32] J. Lu and Y.-P. Tan. Uncorrelated discriminant simplex analysis for view-invariant gait signal computing. *Pattern Recognition Letters*, 31(5):382–393, 2010.

[33] Y. Makihara, R. Sagawa, Y. Mukaigawa, T. Echigo, and Y. Yagi. Gait recognition using a view transformation model in the frequency domain. In *Proc. of the 9th European Conference on Computer Vision*, pages 151–163, Graz, Austria, May 2006.

[34] Y. Makihara and Y. Yagi. Silhouette extraction based on iterative spatio-temporal local color transformation and graph-cut segmentation. In *Proc. of the 19th International Conference on Pattern Recognition*, Tampa, Florida USA, Dec. 2008.

[35] R. Martín-Félez and T. Xiang. Gait recognition by ranking. In *Proceedings of the 12th European conference on Computer Vision - Volume Part I*, ECCV'12, pages 328–341, Berlin, Heidelberg, 2012. Springer-Verlag.

[36] D. Muramatsu, Y. Makihara, and Y. Yagi. Quality-dependent view transformation model for cross-view gait recognition. In *Proc. of the IEEE 6th International Conference on Biometrics: Theory, Applications and Systems*, number Paper ID: 12, pages 1–8, Washington D.C., USA, Sep. 2013.

[37] D. Muramatsu, A. Shiraishi, Y. Makihara, and Y. Yagi. Arbitrary view transformation model for gait person authentication. In *Proc. of the 5th IEEE Int. Conf. on Biometrics: Theory, Applications and Systems*, number Paper ID 39, pages 1–6, Washington D.C., USA, Sep. 2012.

[38] M. S. Nixon, T. N. Tan, and R. Chellappa. *Human Identification Based on Gait*. Int. Series on Biometrics. Springer-Verlag, Dec. 2005.

[39] P. Phillips, H. Moon, S. Rizvi, and P. Rauss. The feret evaluation methodology for face-recognition algorithms. *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 22(10):1090–1104, 2000.

[40] S. Sarkar, J. Phillips, Z. Liu, I. Vega, P. G. ther, and K. Bowyer. The humanid gait challenge problem: Data sets, performance, and analysis. *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 27(2):162–177, 2005.

[41] G. Shakhnarovich, L. Lee, and T. Darrell. Integrated face and gait recognition from multiple views. In *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, volume 1, pages 439–446, 2001.

[42] A. Sharma and D. W. Jacobs. Bypassing synthesis: Pls for face recognition with pose, low-resolution and sketch. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 593–600, 2011.

[43] A. Sharma, A. Kumar, H. D. III, and D. W. Jacobs. Generalized multiview analysis: A discriminative latent space. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 2160–2167, 2012.

[44] Statistical pattern recognition toolbox. `http://cmp.felk.cvut.cz/cmp/software/stprtool/`.

[45] A. Utsumi and N. Tetsutani. Adaptation of appearance model for human tracking using geometrical pixel value distributions. In *Proc. of the 6th Asian Conf. on Computer Vision*, volume 2, pages 794–799, 2004.

[46] C. Wang, J. Zhang, L. Wang, J. Pu, and X. Yuan. Human identification using temporal information preserving gait template. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2164 –2176, nov. 2012.

[47] D. Xu, S. Yan, D. Tao, L. Zhang, X. Li, and H. jiang Zhang. Human gait recognition with matrix representation. *IEEE Trans. Circuits Syst. Video Technol*, 16:896–903, 2006.

[48] S. Yan, D. Xu, Q. Yang, L. Zhang, X. Tang, and H.-J. Zhang. Discriminant analysis with tensor representation. In *Proc. of the IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, pages 526–532, Jun. 2005.

[49] S. Yu, D. Tan, and T. Tan. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. In *Proc. of the 18th Int. Conf. on Pattern Recognition*, volume 4, pages 441–444, Hong Kong, China, Aug. 2006.

[50] S. Yu, D. Tan, and T. Tan. Modelling the effect of view angle variation on appearance-based gait recognition. In *Proc. of 7th Asian Conf. on Computer Vision*, volume 1, pages 807–816, Jan. 2006.