Are Intermediate Views Beneficial for Gait Recognition using a View Transformation Model?

Daigo Muramatsu, Yasushi Makihara, Yasushi Yagi*

Gait recognition is one of behavioral biometrics and has an advantage over the other biometrics in terms that it can be used even at a distance from a camera. Accuracies of the gait recognition, however, degrade if observation views of matching pairs of gaits are different. In order to suppress the accuracy degradation, a family of view transformation models (VTMs) for the gait recognition have been proposed, where a gait feature from one view is transformed to that from another different view so as to match the gait features under the same view. Although the transformation view of the VTM approaches can affect the authentication accuracy in general, its effect has not yet been well investigated in previous work. In this paper, we therefore evaluate the effect of the transformation view for gait recognition experimentally, and report the evaluation results using a publicly available large-population gait database.

Keywords: Gait recognition, observation view, view transformation model, biometrics

1. Introduction

In modern society, biometric person authentication has attracted increasing attention in many different applications, including surveillance, forensic, and access control. While the most of biometrics such as DNA, fingerprints, finger/hand vein, iris, and face, requires subject's cooperation and contact/approach to a sensor, gait biometrics have promising properties: person authentication at a distance from camera and without subject's cooperation [3]. In fact, the gait biometrics have been applied to the forensic science field [6] and also admitted as an evidence of a burglar in UK courts [32]. In addition, the first packaged software of a gait verification system for criminal investigation has been recently developed [11].

Approaches to the gait recognition roughly fall into two families: model-based approaches [1][5][7][33][36][37][40] and appearance-based approaches [2][8][10][20][25][27][29][30][35][39]. Because the model-based approaches often suffer from the human model fitting error and relatively high computational cost, the appearance-based approaches are currently dominant in the gait recognition community and achieve better performances than the model-based approaches in general.

The appearance-based approaches are, however, susceptible to various covariates (e.g., views, shoes, surfaces, clothing, carriages, and walking speeds). Among these covariates, view variation is one of the most important issues for the gait recognition since CCTVs in the street are installed at various positions and angles and also a pedestrian walks into various directions in real situations. The difference of views causes significant changes of the appearance-based gait features, and hence it makes the gait recognition much more difficult. In fact, BenAbdelkader [4] and Yu et al. [38] report that view changes cause significant drops of the gait recognition accuracies.

Approaches to such a cross-view gait recognition fall into two families: discriminative approaches and generative approaches. The discriminative approaches aim at extracting view-invariant gait features via uncorrelated discriminant simplex analysis [23] and joint subspace learning [22].

On the other hand, the generative approaches aims at generating a gait feature of a canonical view from that of another different view. They further fall into three approaches: 3D gait volume-based approaches, geometric approaches, and example-based approaches.

In the 3D gait volume-based approaches [13][21][31], a 3D gait volume is stored as a gallery by reconstructing from multi-view synchronous gait silhouettes by a visual intersection method at first, and then an arbitrary-view gait silhouette is generated by projecting the 3D gait volume so as that the generated gallery view can coincide with a probe view. It is, however, quite difficult to collect multi-view synchronous gait silhouette sequences for uncooperative subjects such as a perpetrator or a suspect in criminal investigation scenarios, and hence the 3D gait volume-based approaches are unsuitable for real applications such as surveillance and criminal investigation.

The geometric approaches often assume that a person (a 3D object) is well approximated by a planar object on a sagittal plane† in conjunction with the weak perspective hypothesis. Jean et al. [14] and Goffredo et al. [9] 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. [15] generate not trajectories but gait silhouettes themselves from side-view in a similar way. These approaches work well in case where an angle between the sagittal plane of the person and an image plane is small enough and also the person is observed at a sufficient distance, otherwise they fail.

The example-based approaches mainly employ a framework of a view transformation model (VTM) learnt from the examples. In this family of approaches, multi-view gait fea-

* Osaka University
8-1, Mihogaoka, Ibaraki, Osaka, Japan 567-0047,
{muramatsu, makihara, yagi}@am.sanken.osaka-u.ac.jp

† A vertical plane dividing an animal into left and right parts.

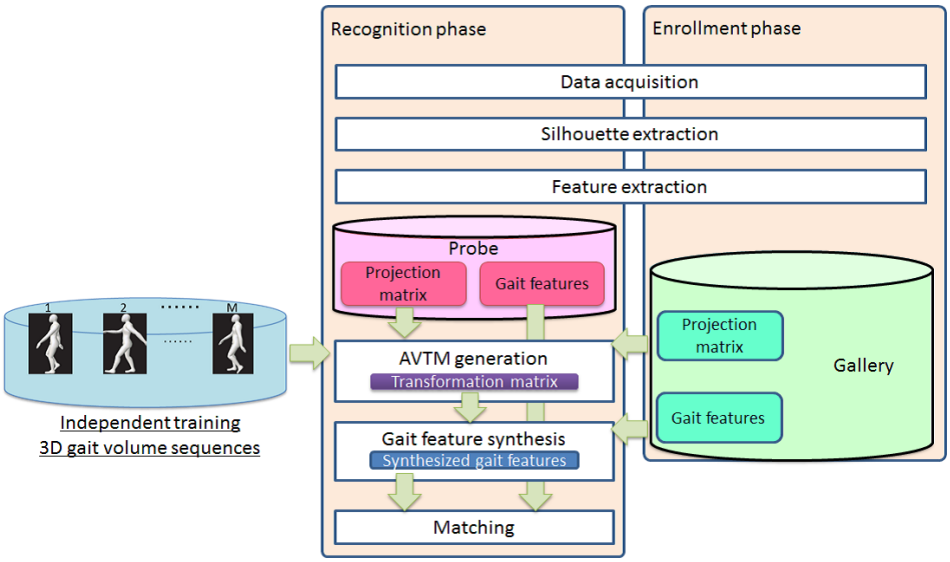

Fig. 1. Overview of the proposed framework

tures are collected not for uncooperative test subjects (recognition target such as a perpetrator and a suspect) but for cooperative training subjects (non-recognition target such as students in a laboratory, colleague in a research institute, volunteers to joint a data collection experiment), and a VTM which transforms a gait feature from one view into that from another view, is constructed using the training data. The VTMs are formulated in a variety of forms such as a matrix factorization by singular value decomposition (SVD) technique [16][25][28][34] and regression [17]~[19]. 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.

A key issue of the VTM family is degree of the view difference for transformation. It is essentially difficult to accurately transform gait features between largely different views, because the correlation of the gait features gets weaker as the view difference gets larger. For example, it is difficult even for us to infer a frontal-view gait feature from the side-view gait feature because the forward-backward motion observed in the side-view gait feature is degenerated in the frontal-view gait feature. In such a situation, the VTM suffers from large transformation errors.

We therefore propose a method of view transformation into *intermediate views* between two different views to reduce the transformation error in total. For example, given a pair of frontal-view gait feature as a probe and side-view gait feature as a gallery, we transform both the frontal-view gait feature and the side-view gait features into gait features from front-oblique view, rather than transform the frontal-view gait feature into the side-view gait feature or transform the side-view gait feature into the frontal-view gait feature. We then match the transformed probe and gallery gait features under the same front oblique-views with reduced transformation errors.

In addition, we investigate which intermediate view between the two different views achieves the best gait recognition accuracy. Since conventional discrete VTMs [16]~[19][25][34] generate gait features from limited number of discrete views included in the training set (e.g., 18 deg interval), they are insufficient for fine-grained analysis on the intermediate views. We therefore employ state-of-the-art VTM, namely, arbitrary VTM (AVTM) [28] which takes advantages of both the 3D gait volume-based approaches and the examples-based approaches, and which also enables us to generate gait features from arbitrary intermediate views for the fine-grained analysis.

2. Gait Recognition with AVTM

2.1 Overview of the Proposed Framework In this subsection, we describe an overview of the proposed framework along with Fig. 1. Once we extract gait features in each of enrollment and recognition phases as a gallery and a probe, respectively, we transform the gait features so as to match under the same view.

We employ the AVTM [28] for this purpose. More specifically, we construct 3D gait volume sequences for cooperative training subjects which are independent from test subjects. Given a pair of gait features of a test subject as well as camera calibration data in enrollment and recognition phases, respectively, we transform the 3D gait volume sequence of the training subjects into 2D gait silhouette sequence to train a custom-made VTM for a specific pair of gallery and probe views for the test subject. We then transform the gait features of the test subject from one view to another view with the learnt AVTM.

2.2 Preprocessing Since we employ an appearance-based gait feature, more specifically, silhouette-based gait feature, we extract a gait silhouette sequence from an input gait image sequence. First, given an input gait image sequence as well as a background image, a gait silhouette for each frame is extracted as a foreground region by background subtraction-based graph-cut segmentation [26]. Next, a normalized gait silhouette sequence of pre-determined size is generated by image size-normalization and registration based on the height and the center of gravity of each silhouette.

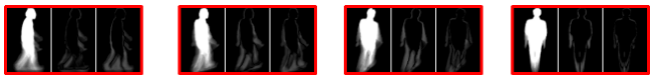

Fig. 2. FDFs from various views. Left: 0-times frequency, middle: 1-time frequency, right: 2-times frequency.

2.3 Frequency-domain Feature Among a large variety of appearance-based gait features [2][8][10][20][25][27][29][30][35][39], we employ a frequency-domain feature (FDF) [25] since the FDF achieved state-of-the-art gait recognition accuracy in performance evaluation with the world largest gait database [12]. In this subsection, we briefly describe a feature extraction process of the FDF.

We first detect a gait period, namely, a time of a pair of left and right steps, by maximizing autocorrelation of the size-normalized silhouette sequence along the time axis. We then apply one-dimensional Fourier transformation to the size normalized silhouette sequence again along the time axis and compute the amplitude spectra for 0-, 1-, and 2-times frequency elements for each pixel independently. Figure 2 shows extracted FDFs from various views.

2.4 Training AVTM A training process of the AVTM is composed of two steps: generation of training data and construction of view transformation matrices. At first, let θ_n and $P_{\theta_n} \in \mathbb{R}^{3\times 4} (n = 1,\ldots,N)$ be the n-th target view for view transformation and a projection matrix for the n-th target view, respectively, where N is the number of target views. Moreover, let $S^m(t)(m = 1,\ldots,M, t = 1,\ldots,T_m)$ be a 3D gait volume of the m-th training subject at the t-th frame, where M and T_m is the number of training subjects and the number of frames of one gait period for the m-th training subject, respectively. We train the AVTM with these data as follows.

Step 1: Generation of training data

We generate a gait silhouette $\xi^m_{\theta_n}(t)$ of the m-th training subject at the t-th frame from the n-th target view with the 3D gait volume and the projection matrix as

$$\xi^m_{\theta_n}(t) = \mathrm{Project}(S^m(t); P_{\theta_n}), \tag{1}$$

where $\mathrm{Project}(\cdot;P)$ is a mapping function from a 3D gait volume into a 2D gait silhouette via the projection matrix P.

We then extract the FDF $\mathbf{x}^m_{\theta_n} \in \mathbb{R}^K$ of the m-th training subject from the n-th target view with the projected 2D gait silhouette sequence $\{\xi^m_{\theta_n}(t)\}(t = 1,\ldots,T_m)$ as described in Subsection 2.3, where K is the dimension of the FDF.

Step 2: Construction of view transformation matrices

We first set up a training matrix composed of the FDFs of M training subjects from N target views and apply SVD to it as

$$\begin{bmatrix} \mathbf{x}^1_{\theta_1} & \cdots & \mathbf{x}^M_{\theta_1} \\ \vdots & \ddots & \vdots \\ \mathbf{x}^1_{\theta_N} & \cdots & \mathbf{x}^M_{\theta_N} \end{bmatrix} = U\Sigma V^T = \begin{bmatrix} R_{\theta_1} \\ \vdots \\ R_{\theta_N} \end{bmatrix} \begin{bmatrix} \mathbf{v}^1 \cdots \mathbf{v}^M \end{bmatrix}, \tag{2}$$

where $U \in \mathbb{R}^{NK\times NK}$ and $V \in \mathbb{R}^{M\times M}$ are orthogonal matrices and $\Sigma \in \mathbb{R}^{NK\times M}$ is a matrix whose diagonal components are singular values. In addition, $\mathbf{v}^m \in \mathbb{R}^M$ is an intrinsic column vector for the m-th subject and $R_{\theta_n} \in \mathbb{R}^{K\times M}$ is a submatrix of $U\Sigma$ which projects the intrinsic column vector $\mathbf{v}$ into a gait feature from the n-th target view. Now, we can write the FDF of the m-th subject from the n-th view as

$$\mathbf{x}^m_{\theta_n} = R_{\theta_n}\mathbf{v}^m. \tag{3}$$

2.5 View Transformation In this subsection, we describe view transformation for a test subject. Let an FDF of the test subject from view θ_n be $\mathbf{x}_{\theta_n}$ and we try transforming it into that from view θ_l. At first, we estimate the intrinsic column vector for the test subject by the following least square

$$\hat{\mathbf{v}}_{\theta_n} = \arg\min_{\mathbf{v}} ||R_{\theta_n}\mathbf{v} - \mathbf{x}_{\theta_n}||^2 \quad \cdots\cdots (4)$$

$$= \left((R_{\theta_n})^T R_{\theta_n}\right)^{-1} (R_{\theta_n})^T \mathbf{x}_{\theta_n}. \quad \cdots\cdots (5)$$

Once we obtain the estimated intrinsic column vector $\hat{\mathbf{v}}_{\theta_n}$, we can estimate an FDF $\hat{\mathbf{x}}_{\theta_l\leftarrow\theta_n}$ transformed from view θ_n to view θ_l based on Eq. (3) as

$$\hat{\mathbf{x}}_{\theta_l\leftarrow\theta_n} = R_{\theta_l}\hat{\mathbf{v}}_{\theta_n}. \tag{6}$$

2.6 Matching In this subsection, we consider matching a pair of gallery FDF $\mathbf{x}^G$ from view θ^G and probe FDF $\mathbf{x}^P$ from view θ^P. We generate gallery and probe FDFs transformed from views θ^G and θ^P into the same intermediate target view ϕ, respectively, by the trained AVTM. We then compute a dissimilarity score d between gallery and probe FDFs under the same view ϕ as

$$d(\mathbf{x}^G, \mathbf{x}^P; \phi) = \|\hat{\mathbf{x}}^G_{\phi\leftarrow\theta^G} - \hat{\mathbf{x}}^P_{\phi\leftarrow\theta^P}\|_2$$

$$= \|R_\phi(\hat{\mathbf{v}}^G_{\theta^G} - \hat{\mathbf{v}}^P_{\theta^P})\|_2, \quad \cdots\cdots (7)$$

where $\hat{\mathbf{x}}^G_{\phi\leftarrow\theta^G}$ and $\hat{\mathbf{x}}^P_{\phi\leftarrow\theta^P}$ are the transformed gallery and probe FDFs from view θ^G and θ^P into the same view ϕ. As we can see, the dissimilarity score between the gallery and probe FDFs is a function of the target view ϕ, and hence the gait recognition accuracy may change as the target view ϕ changes. While this kind of target view has not been actively discussed in the previous VTMs due to a limited number of discrete views, we can arbitrarily select the intermediate target views so as to improve the gait recognition accuracy.

Note that previous VTMs only have two options: transformation from probe view to gallery view or from gallery view to probe view. Under such a situation, dissimilarity scores with the previous VTMs are computed as either

$$d_{P2G}(\mathbf{x}^G, \mathbf{x}^P) = \|\mathbf{x}^G - \hat{\mathbf{x}}^P_{\theta^G\leftarrow\theta^P}\|_2$$

$$= \|\mathbf{x}^G - R_{\theta^G}\hat{\mathbf{v}}^P_{\theta^P}\|_2, \quad \cdots\cdots (8)$$

or

$$d_{G2P}(\mathbf{x}^G, \mathbf{x}^P) = \|\hat{\mathbf{x}}^G_{\theta^P\leftarrow\theta^G} - \mathbf{x}^P\|_2$$

$$= \|R_{\theta^P}\hat{\mathbf{v}}^G_{\theta^G} - \mathbf{x}^P\|_2. \quad \cdots\cdots (9)$$

3. Experiments

3.1 Data Sets We used the OU-ISIR Gait Database for our experiments. We drew 103 training subjects to construct 3D gait volume sequences for the AVTM from treadmill data set [24]. Moreover, we drew a subset of the large population data set [12] composed of over 4,000 subjects for

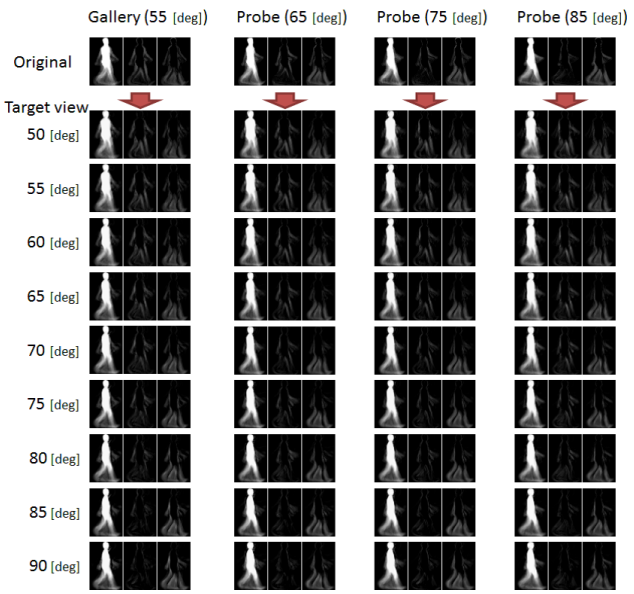

Fig. 3. Transformed FDF with the AVTM

performance evaluation. More specifically, we chose walking image sequences of 1,912 subjects whose camera calibration data was available. Each of the subject was asked to walk twice along with a specified course and was captured by a camera. This data set was further divided into subsets depending on the observation views; 55, 65, 75, and 85 deg.

3.2 Setup In the first experimental setup, we assigned 55-deg view to a gallery view and the other views (65, 75, and 85 deg) to the probe views. In the second experimental setup, we assigned 85-deg view to a gallery view and the other views (55, 65, and 75 deg) to the probe views. In both cases, the intermediate views changed from 50 deg to 90 deg at 5 deg interval.

3.3 Transformed FDF In this subsection, we shows the transformed FDFs into the intermediate target views from each of gallery and probe views with the trained AVTM in Fig. 3. As a result, we can see that the transformed FDFs are smoothly transited from front-oblique view (50 deg) to side view (90 deg) for each of gallery and probe FDFs and also FDFs under the same target views are similar to each other among the gallery and probes.

3.4 Quantitative Evaluation in Verification Scenario We evaluated the proposed framework in a verification scenario, namely, one-to-one matching. The verification performance is generally evaluated with a so-called receiver operating characteristics (ROC) curve which indicates a tradeoff between false acceptance rate (FAR) of imposter (false match pair) and false rejection rate (FRR) of genuine (true match pair) when changing an acceptance threshold for the dissimilarity score. Figures 4 and 5 show ROC curves for matching of 85-deg gallery and 55-deg probe, and 55-deg gallery and 85-deg probe, respectively. In both cases, we assign the target views to just intermediate view (70 deg) as well as the gallery and probe views (55 deg and 85 deg). As a result, we can see that intermediate target view yielded better performance in the both cases.

Moreover, we picked up an equal error rate (EER) of the FAR and the FRR as a typical evaluation measure for the verification scenario and summarized them as shown in Table 1. As a result, we can see that intermediate target views yielded better performance as a whole. More specifically, while the performance improvement is limited in case of relatively small view difference between the gallery and probe (e.g., 55 deg vs. 65 deg), the performance clearly improve in case of relatively large view difference (e.g., 55 deg vs. 85 deg). For example, in case of 55-deg gallery and 85-deg

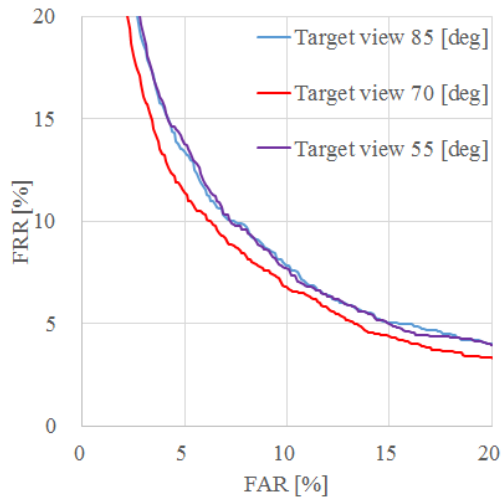

Fig. 4. ROC curves for 55-deg gallery

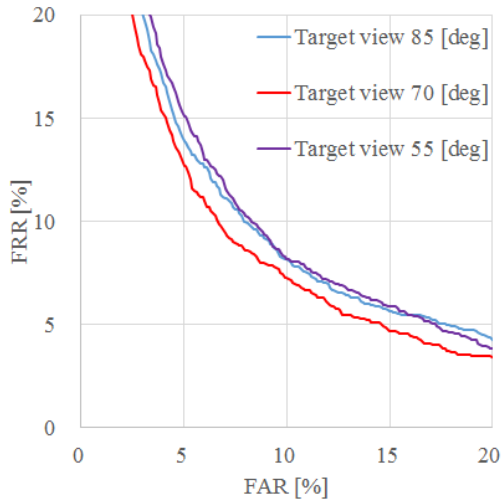

Fig. 5. ROC curves for 85-deg gallery

Table 1. EERs [%]. Bold indicate the best performance for each setup, while the same views as either gallery or probe views are underlined.

Gallery view [deg]	55			85		
Probe view [deg]	65	75	85	55	65	75
Target view [deg]						
50	**4.8**	6.6	9.5	9.2	7.0	5.3
55	**4.8**	**6.4**	9.1	8.8	6.7	5.4
60	4.9	6.5	8.8	8.5	6.7	5.4
65	5.2	**6.4**	8.7	8.3	6.4	5.3
70	5.4	6.5	8.4	**8.2**	6.3	5.2
75	5.4	6.6	**8.3**	**8.2**	**6.0**	5.0
80	5.5	7.0	**8.3**	**8.2**	6.2	**4.9**
85	5.9	7.5	9.1	8.8	6.7	**4.9**
90	6.0	8.1	9.9	9.4	7.1	5.0

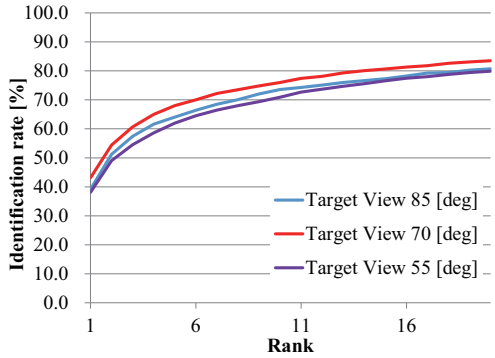

Fig. 6. CMC curves for 55-deg gallery

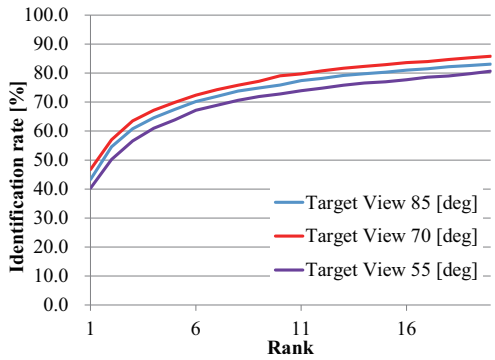

Fig. 7. CMC curves for 85-deg gallery

probe, 0.9% improvement from 9.1% EER for 55-deg and 85-deg target views to 8.3% EER for 75-deg target view.

3.5 Quantitative Evaluation in Identification Scenario We also evaluated the proposed framework in an identification scenario, namely, one-to-many matching. The identification performance is generally evaluated with a cumulative matching characteristics (CMC) curve which indicates a ratio of correct match pairs within each specific rank in the horizontal axis. Figures 6 and 7 shows CMC curves for matching of 85-deg gallery and 55-deg probe, and 55-deg gallery and 85-deg probe, respectively. In both cases, we assign the target views to just intermediate view (70 deg) as well as the gallery and probe views (55 deg and 85 deg). As a result, we can see that intermediate target view yielded better performance in the both cases.

Moreover, we picked up rank-1, rank-5, and rank-10 identification rates as typical evaluation measures for the identification scenario and summarized them as shown in Tables 2, 3, and 4. As a result, we can draw similar conclusions to the verification scenario. While the performance improvement is limited in case of relatively small view difference between the gallery and probe (e.g., 55 deg vs. 65 deg), the performance clearly improve in case of relatively large view difference (e.g., 55 deg vs. 85 deg). For example, in case of 55-deg gallery and 85-deg probe, 5.6% and 4.5% improvements from 38.1% rank-1 identification rate for 55-deg target view and 39.2% rank-1 identification rates for 85-deg target view to 43.7% rank-1 identification rate for 75-deg target view.

4. Conclusion

We proposed an extended framework of the previously proposed VTM. Whereas the previous VTMs transform the gait features either from gallery view to probe view or from probe view to gallery view, we rather transform the gait features into *intermediate views* between gallery and probe views to reduce the transformation error in total. As a result of experiments with large population gait database, we confirmed that the proposed transformation into intermediate target views yielded better performance than previous transformation into gallery or probe views.

In future, we further investigate a way how to automatically select the optimal intermediate target views given a pair of gallery and probe views.

Table 2. Rank-1 identification rates [%]

Gallery view [deg]	55			85		
Probe view [deg]	65	75	85	55	65	75
Target view [deg]						
50	74.3	55.4	36.2	37.2	51.8	68.1
55	**74.6**	56.6	38.1	40.3	55.0	69.4
60	**74.6**	57.8	40.2	43.1	57.9	70.8
65	73.9	58.4	41.4	45.4	59.5	71.8
70	74.3	**58.9**	43.1	**46.7**	61.9	72.7
75	72.8	58.2	**43.7**	45.4	**62.3**	73.4
80	71.5	55.9	42.3	45.3	**62.3**	**74.2**
85	70.1	53.5	39.2	43.3	61.3	74.1
90	67.8	50.8	35.3	39.3	58.2	73.3

Table 3. Rank-5 identification rates [%]

Gallery View [deg]	55			85		
Probe View [deg]	65	75	85	55	65	75
Target View [deg]						
50	87.3	74.8	60.4	62.5	73.1	83.0
55	87.8	76.3	62.0	63.9	75.8	83.9
60	**87.9**	77.0	64.0	66.4	77.5	84.5
65	87.6	**78.0**	66.0	69.0	79.6	85.4
70	87.2	77.7	**68.0**	69.9	81.0	86.0
75	86.6	76.9	67.8	**70.6**	81.4	87.0
80	86.2	75.7	66.6	69.8	**81.8**	**87.6**
85	84.7	73.8	64.1	67.5	81.0	87.0
90	84.3	71.1	60.8	64.2	79.4	86.9

Table 4. Rank-10 identification rates [%]

Gallery View [deg]	55			85		
Probe View [deg]	65	75	85	55	65	75
Target View [deg]						
50	91.0	81.7	69.0	70.9	80.8	87.2
55	**91.2**	82.4	70.9	72.8	82.6	87.8
60	90.8	82.8	73.2	74.9	84.4	88.4
65	90.8	82.9	75.1	77.6	85.6	89.4
70	90.4	**83.1**	**76.0**	**79.1**	86.2	90.7
75	89.5	82.6	75.7	77.9	86.7	90.6
80	89.4	82.0	75.5	77.4	**86.9**	**91.3**
85	88.9	80.6	73.6	75.9	86.3	90.5
90	87.9	78.6	69.6	73.6	84.8	90.5

Acknowledgement

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 Intension-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 representation using flow fields. In *Proc. of the 20th British Machine Vision Conference*, London, UK, Sep. 2009.

(3) K. Bashir, T. Xiang, and S. Gong. Gait recognition without subject coopera-

tion. *Pattern Recognition Letters*, 31(13):2052–2060, Oct. 2010.
(4) C. BenAbdelkader. *Gait as a Biometric For Person Identification in Video*. PhD thesis, Maryland Univ., 2002.
(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) 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.
(8) N. Cuntoor, A. Kale, and R. Chellappa. Combining multiple evidences for gait recognition. In *Proc. of the 28th IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 3, pages 33–36, 2003.
(9) 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.
(10) J. Han and B. Bhanu. Individual recognition using gait energy image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(2):316– 322, 2006.
(11) 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.
(12) 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.
(13) 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.
(14) F. Jean, R. Bergevin, and A. B. Albu. Computing and evaluating view-normalized body part trajectories. *Image Vision Comput.*, 27(9):1272–1284, Aug. 2009.
(15) 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.
(16) W. Kusakunniran, Q. Wu, H. Li, and J. Zhang. Multiple views gait recognition using view transformation model based on optimized gait energy image. In *Proc. of the 2nd IEEE International Workshop on Tracking Humans for the Evaluation of their Motion in Image Sequences*, pages 1058–1064, Oct. 2009.
(17) 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.
(18) 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 computer society conference on Computer Vision and Pattern Recognition 2010*, pages 1–8, San Francisco, CA, USA, Jun. 2010.
(19) 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.
(20) 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.
(21) L. Lee. *Gait Analysis for Classification*. PhD thesis, Massachusetts Institute of Technology, 2002.
(22) 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.
(23) J. Lu and Y.-P. Tan. Uncorrelated discriminant simplex analysis for view-invariant gait signal computing. *Pattern Recognition Letters*, 31(5):382–393, 2010.
(24) Y. Makihara, H. Mannami, A. Tsuji, M. Hossain, K. Sugiura, A. Mori, and Y. Yagi. The ou-isir gait database comprising the treadmill dataset. *IPSJ Transactions on Computer Vision and Applications*, 4:53–62, Apr. 2012.
(25) 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.
(26) 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.
(27) S. Mowbray and M. Nixon. Automatic gait recognition via fourier descriptors of deformable objects. In *Proc. of the 1st IEEE International Conference on Advanced Video and Signal Based Surveillance*, pages 566–573, 2003.
(28) 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.
(29) S. Niyogi and E. Adelson. Analyzing and recognizing walking figures in xyt. In *Proc. of the 7th IEEE Conference on Computer Vision and Pattern Recognition*, pages 469–474, 1994.
(30) 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.
(31) 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.
(32) How biometrics could change security. `http://news.bbc.co.uk/2/hi/programmes/click_online/7702065.stm`, Oct. 2008.
(33) R. Urtasun and P. Fua. 3d tracking for gait characterization and recognition. In *Proc. of the 6th IEEE International Conference on Automatic Face and Gesture Recognition*, pages 17–22, 2004.
(34) 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.
(35) 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.
(36) L. Wang, H. Ning, T. Tan, and W. Hu. Fusion of static and dynamic body biometrics for gait recognition. In *Proc. of the 9th International Conference on Computer Vision*, volume 2, pages 1449–1454, 2003.
(37) C. Yam, M. Nixon, and J. Carter. Extended model based automatic gait recognition of walking and running. In *Proc. of the 3rd International Conference on Audio and Video-based Person Authentication*, pages 278–283, Halmstad, Sweden, Jun. 2001.
(38) 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.
(39) S. Yu, L. Wang, W. Hu, and T. Tan. Gait analysis for human identification in frequency domain. In *Proc. of the 3rd International Conference on Image and Graphics*, pages 282–285, 2004.
(40) G. Zhao, G. Liu, H. Li, and M. Pietikainen. 3d gait recognition using multiple cameras. In *Proc. of the 7th International Conference on Automatic Face and Gesture Recognition*, pages 529–534, Apr. 2006.

Yasushi Makihara received the BS, MS, and PhD degrees in engineering from Osaka University in 2001, 2002, and 2005, respectively. He is currently an assistant professor of the Institute of Scientific and Industrial Research, Osaka University. His research interests are computer vision and pattern recognition, in particular, image segmentation, image morphing, temporal super resolution, score-level fusion, and their applications to gait recognition. He has 50 publications in major international conferences (e.g., CVPR, ECCV, ACCV, ICPR, FG) as well as journal papers. He has also written eight reports of expert evidence on gait-based person verification at the request of police.

Daigo Muramatsu received the B.S., M.E., and Ph.D. degrees in Electrical, Electronics, and Computer Engineering from Waseda University in 1997, 1999, and 2006, respectively. He is currently a Specially Appointed Associate Professor of the Institute of Scientific and Industrial Research, Osaka University. His research interests are gait recognition, signature verification, and biometric fusion. He is a member of IEEE, IEICE, and IIEEJ.

Yasushi Yagi is the Director of the Institute of Scientific and Industrial Research, Osaka University, Ibaraki, Japan. He received his Ph.D. degrees from Osaka University in 1991. In 1985, he joined the Product Development Laboratory, Mitsubishi Electric Corporation, where he worked on robotics and inspections. He became a Research Associate in 1990, a Lecturer in 1993, an Associate Professor in 1996, and a Professor in 2003 at Osaka University. International conferences for which he has served as Chair include: FG1998 (Financial Chair), OMINVIS2003 (Organizing chair), ROBIO2006 (Program co-chair), ACCV2007 (Program chair), PSVIT2009 (Financial chair), ICRA2009 (Technical Visit Chair), ACCV2009 (General chair), ACPR2011 (Program co-chair) and ACPR2013 (General chair). He has also served as the Editor of IEEE ICRA Conference Editorial Board (2007–2011). He is the Editorial member of IJCV and the Editor-in-Chief of IPSJ Transactions on Computer Vision & Applications. He was awarded ACM VRST2003 Honorable Mention Award, IEEE ROBIO2006 Finalist of T.J. Tan Best Paper in Robotics, IEEE ICRA2008 Finalist for Best Vision Paper, MIRU2008 Nagao Award, and PSIVT2010 Best Paper Award. His research interests are computer vision, medical engineering and robotics. He is a fellow of IPSJ and a member of IEICE, RSJ, and IEEE.