

Gait Recognition using a View Transformation Model in the Frequency Domain

Yasushi Makihara¹, Ryusuke Sagawa¹, Yasuhiro Mukaigawa¹, Tomio Echigo¹,
and Yasushi Yagi¹

Department of Intelligent Media, The Institute of Scientific and Industrial Research,
Osaka University, 567-0047, 8-1 Mihogaoka, Ibaraki, Osaka, JAPAN
{makihara, sagawa, mukaigaw, echigo, yagi}@am.sanken.osaka-u.ac.jp
<http://www.am.sanken.osaka-u.ac.jp/index.html>

Abstract. Gait analyses have recently gained attention as methods of identification of individuals at a distance from a camera. However, appearance changes due to view direction changes cause difficulties for gait recognition systems. Here, we propose a method of gait recognition from various view directions using frequency-domain features and a view transformation model. We first construct a spatio-temporal silhouette volume of a walking person and then extract frequency-domain features of the volume by Fourier analysis based on gait periodicity. Next, our view transformation model is obtained with a training set of multiple persons from multiple view directions. In a recognition phase, the model transforms gallery features into the same view direction as that of an input feature, and so the features match each other. Experiments involving gait recognition from 24 view directions demonstrate the effectiveness of the proposed method.

1 Introduction

There is a growing necessity in modern society for identification of individuals in many situations, such as from surveillance systems and for access control. For personal identification, many biometrics-based authentication methods are proposed using a wide variety of cues; fingerprint, finger or hand vein, voiceprint, iris, face, handwriting, and gait. Among these, gait recognition has recently gained considerable attention because gait is a promising cue for surveillance systems to ascertain identity at a distance from a camera.

Current approaches of gait recognition are mainly divided into model-based and appearance-based ones.

The model-based approaches extract gait features such as shape and motion by fitting the model to input images. Some methods [1][2] extracted periodical features of leg motion by Fourier analysis. Bobick et al. [3] extracted parameters of shape and stride. Wagg et al. [4] extracted static shape parameters and gait period with an articulated body model, and Urtasun et al. [5] extracted joint angles with an articulated body model. Those model-based approaches often face difficulties with model fitting or feature extraction.

Appearance-based approaches directly analyze images and extract features without body models. Sarkar et al. [6] proposed direct matching of silhouette image sequences as a baseline algorithm. Murase et al. [7] represented a gait image sequence as a trajectory in an eigen space and matched the trajectories. Ohara et

al. [8] and Niyogi et al. [9] constructed a spatio-temporal volume (x - y - t volume) by combining gait images and matched features extracted from the volume. Indeed, many gait features are proposed as being useful [10][11][12][13][14][15][16].

One of the difficulties facing appearance-based approaches is that appearance changes due to a change of the viewing or walking directions. In fact, BenAbdelkader [17] and Yu et al. [18] reported that view changes caused a drop in gait recognition performance.

To cope with the view changes, Shakhnarovich et al. [19] and Lee [20] proposed methods to synthesize an image for a virtual view direction using a visual hull. However, this method needs images taken synchronously from multiple view directions for all subjects and then necessitates the use of a multi-camera system or for there to be a solution to the troublesome problem of frame synchronization. Kale et al. [21] proposed a method to synthesize arbitrary-view images from a single-view image with perspective projection by assuming gait motion occurs in a sagittal plane. This method, however, does not work well because self occlusion occurs when an angle formed by an image plane and the sagittal plane is large.

To overcome these defects, we exploit a view transformation model (VTM) for appearance-based gait recognition. In the proposed method, once we obtain a VTM using a training set, made up of images of multiple subjects from multiple views, we can make images of a new subject taken from the multiple view directions by transforming a single-view image of the new subject.

In other computer vision areas, many methods have achieved adaptation to view direction changes with VTM. Mukaigawa et al. [22] applied the model to face image synthesis with pose and expression changes, and Utsumi et al. [23] applied it to transform images with pose and view changes.

However, these approaches just transform a static image into another static image; gait analysis, on the other hand, treats not a static image but a spatio-temporal volume. View transformation from a volume into another volume, though, causes troublesome problems such as frame synchronization. To overcome this, we first extract frequency-domain features from a spatio-temporal gait silhouette volume (GSV), and then we apply the VTM for frequency-domain features. Note that the use of the frequency-domain features releases us from the need for frame synchronization when view transformation and matching are performed.

The outline of this paper is as follows. We describe the construction of a GSV in section 2, and the matching of a GSV in section 3. Then, adaptation to view direction changes is addressed with the formulation of our VTM in section 4, and experiments of gait recognition from various view directions are shown in section 5. In section 6, we present our conclusions and indicate future works.

2 Construction of a GSV

2.1 Extraction of gait silhouette images

The first step in constructing a GSV is to extract gait silhouette images; to do this, background subtraction is exploited. Background subtraction, however, sometimes fails because of cast shadows and illumination condition changes (see Fig. 1(a)(b)). To avoid such difficulties, we execute a temperature-based background subtraction using an infrared-ray camera (NEC TH1702MX) instead of

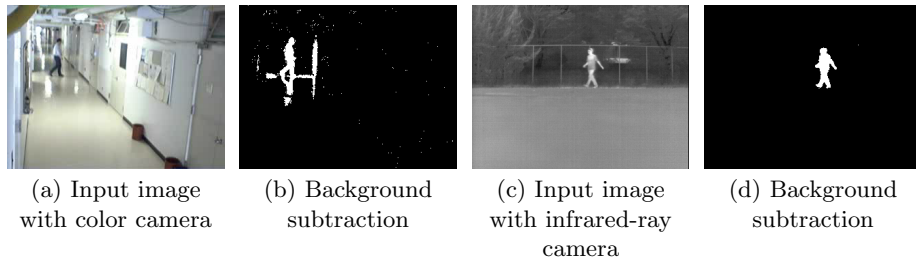


Fig. 1. Comparison of background subtraction between color camera and infrared-ray camera (In (c), brighter colors indicate higher temperature))

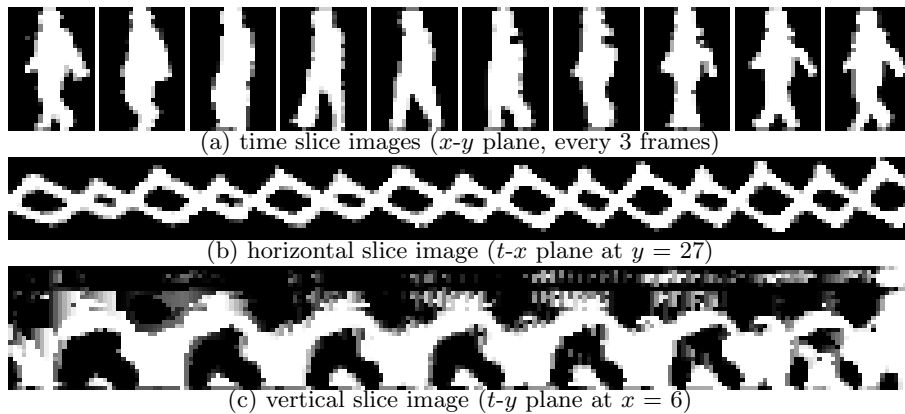


Fig. 2. An example of GSV

a conventional color camera. The infrared-ray camera captures 30 frames per second sized at 320×240 pixels. Figure 1(c) is an input image taken by the infrared-ray camera. In it we can see that the temperatures of a person are higher than those of the background; therefore we can extract clear regions as a gait silhouette image (see Fig. 1(d)). Here, for simplicity we assume only one person exists in the image, thus we keep only the largest connected region as the person.

2.2 Scaling and registration of silhouette images

The next step is scaling and registration of the extracted silhouette images. First, the top, the bottom, and horizontal center of the regions for each frame are obtained. The horizontal center is chosen as the median of horizontal positions belonging to the region. Second, a moving average filter of 30 frames is applied to those positions. Third, we scale the silhouette images so that the height can be just 30 pixels based on the averaged positions, and so that the aspect ratio of each region can be kept. Finally, we produce a 20×30 pixel-sized image in which the averaged horizontal median corresponds to the horizontal center of the image.

We show an example of a constructed GSV in Fig. 2 as time slice (x - y plane), horizontal slice (t - x plane), and vertical slice (t - y plane) images. We can confirm gait periodicity from Fig. 2(b), (c).

3 Matching of a GSV

3.1 Gait period detection

The first step for matching is gait period detection. We calculate the normalized autocorrelation of a GSV for the temporal axis as

$$C(N) = \frac{\sum_{x,y} \sum_{n=0}^{N_{total}-N-1} g_{gsv}(x,y,n)g_{gsv}(x,y,n+N)}{\sqrt{\sum_{x,y} \sum_{n=0}^{N_{total}-N-1} g_{gsv}(x,y,n)^2} \sqrt{\sum_{x,y} \sum_{n=0}^{N_{total}-N-1} g_{gsv}(x,y,n+N)^2}}, \quad (1)$$

where $C(N)$ is the autocorrelation for the N frame shift, $g_{gsv}(x,y,n)$ is the silhouette value at position (x,y) at the n th frame, and N_{total} is the number of total frames in the sequence. We set the domain of N to be $[20, 40]$ empirically for the natural gait period; this because various gait types such as running, brisk walking, and ox walking are not within the scope of this paper. Thus, the gait period N_{gait} is estimated as

$$N_{gait} = \arg \max_{N \in [20,40]} C(N). \quad (2)$$

3.2 Extraction of frequency-domain features

As mentioned in the introduction, we use frequency-domain features based on the gait period N_{gait} as gait features to avoid troublesome frame synchronization when matching and view transformations are executed. First we pick up the subsequences $\{\mathbf{S}_i\} (i = 1, 2, \dots, N_{sub})$ for every N_{gait} frames from a total sequence \mathbf{S} . Note that the frame range of the i th subsequence \mathbf{S}_i is $[iN_{gait}, (i+1)N_{gait} - 1]$. Then the Discrete Fourier Transformation (DFT) for the temporal axis is applied for each subsequence, and amplitude spectra are subsequently calculated as

$$G_i(x,y,k) = \sum_{n=iN_{gait}}^{(i+1)N_{gait}-1} g_{gsv}(x,y,n) e^{-j\omega_0 kn} \quad (3)$$

$$A_i(x,y,k) = |G_i(x,y,k)|, \quad (4)$$

where ω_0 is a base angular frequency for the gait period N_{gait} , $G_i(x,y,k)$ is the DFT of GSV for k -times the gait period, and $A_i(x,y,k)$ is an amplitude spectrum for $G_i(x,y,k)$.

Direct-current elements ($k = 0$) of the DFT do not represent gait periodicity; therefore, they should be removed from the features. Moreover, high frequency elements ($k > k_{thresh}$) have less intensity than lower-frequency ones and mainly consist of noise, thus they also should be removed. In this paper, we decide $k_{thresh} = 5$ experimentally. As a result, $A_i(x,y,k) (k = 1, \dots, 5)$ is used as the gait feature and its dimension N_A sums up to $20 \times 30 \times 5 = 3000$.

Figure 3 shows extracted amplitude spectra for various view directions. The view direction is defined as the angle formed by an optical axis and a walking direction, as shown in Fig. 4, and in this paper the unit of the view direction is a degree. Amplitude spectra vary widely among view directions for each subject, and to some extent they also have individual variations for each view direction. Moreover, we can see that all the subjects have similar common tendencies for amplitude spectra variations across view direction changes. This fact indicates a real possibility that the variations across view direction changes are expressed with the VTM independently of individual variations.

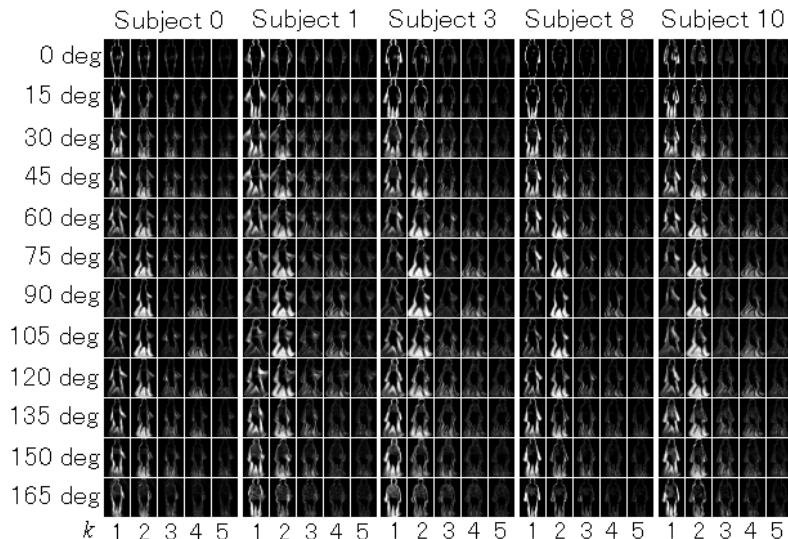


Fig. 3. Extracted features for every 15 degree view direction for some subjects

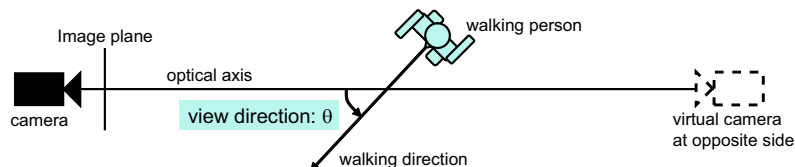


Fig. 4. Definition of view direction θ at top view

3.3 Matching measures

We first define a matching measure between two subsequences. Let $\mathbf{a}(\mathbf{S}_i)$ be a N_A dimensional feature vector composed of elements of the amplitude spectra $A_i(x, y, k)$. The matching measure $d(\mathbf{S}_i, \mathbf{S}_j)$ is simply chosen as the Euclidean distance:

$$d(\mathbf{S}_i, \mathbf{S}_j) = \|\mathbf{a}(\mathbf{S}_i) - \mathbf{a}(\mathbf{S}_j)\|. \quad (5)$$

Next, we define a matching measure between two total sequences. Let \mathbf{S}_P and \mathbf{S}_G be total sequences for probe and gallery, respectively, and let $\{\mathbf{S}_{P_i}\} (i = 1, 2, \dots)$ and $\{\mathbf{S}_{G_j}\} (j = 1, 2, \dots)$ be their subsequences, respectively. Gallery subsequences $\{\mathbf{S}_{G_j}\}$ have variations in general and probe subsequences $\{\mathbf{S}_{P_i}\}$ may contain outliers. A measure candidate $D(\mathbf{S}_P, \mathbf{S}_G)$ to cope with them is the median value of the minimum distances of each probe subsequence \mathbf{S}_{P_i} and gallery subsequences $\{\mathbf{S}_{G_j}\} (j = 1, 2, \dots)$:

$$D(\mathbf{S}_P, \mathbf{S}_G) = \text{Median}_i[\min_j\{d(\mathbf{S}_{P_i}, \mathbf{S}_{G_j})\}]. \quad (6)$$

4 Adaptation to view direction changes

We briefly describe the formulation of a VTM in a way similar to that in [23]. Note that we apply the model to the frequency-domain feature extracted from gait image sequences while that in [23] directly applied it to a static image.

We first quantize view directions into K directions. Let $\mathbf{a}_{\theta_k}^m$ be a N_A dimensional feature vector for the k th view direction of the m th subject. Supposing that the feature vectors for K view directions of M subjects are obtained as a training set, we can construct a matrix whose row indicates view direction changes and whose column indicates each subject; and so can decompose it by Singular Value Decomposition (SVD) as

$$\begin{bmatrix} \mathbf{a}_{\theta_1}^1 & \cdots & \mathbf{a}_{\theta_1}^M \\ \vdots & \ddots & \vdots \\ \mathbf{a}_{\theta_K}^1 & \cdots & \mathbf{a}_{\theta_K}^M \end{bmatrix} = USV^T = \begin{bmatrix} P_{\theta_1} \\ \vdots \\ P_{\theta_K} \end{bmatrix} [\mathbf{v}^1 \cdots \mathbf{v}^M], \quad (7)$$

where U is the $KN_A \times M$ orthogonal matrix, V is the $M \times M$ orthogonal matrix, S is the $M \times M$ diagonal matrix composed of singular values, P_{θ_k} is the $N_A \times M$ submatrix of US , and \mathbf{v}^m is the M dimensional column vector.

The vector \mathbf{v}^m is an intrinsic feature vector of the m th subject and is independent of view directions. The submatrix P_{θ_k} is a projection matrix from the intrinsic vector \mathbf{v} to the feature vector for view direction θ_k , and is common for all subjects, that is, it is independent of the subject. Thus, the feature vector $\mathbf{a}_{\theta_i}^m$ for the view direction θ_i of the m th subject is represented as

$$\mathbf{a}_{\theta_i}^m = P_{\theta_i} \mathbf{v}^m. \quad (8)$$

Then, feature vector transformation from view direction θ_j to θ_i is easily obtained as

$$\mathbf{a}_{\theta_i}^m = P_{\theta_i} P_{\theta_j}^+ \mathbf{a}_{\theta_j}^m, \quad (9)$$

where $P_{\theta_j}^+$ is the pseudo inverse matrix of P_{θ_j} . In practical use, transformation from one view direction may be insufficient because motions orthogonal to the image plane are degenerated in the silhouette image. For example, it is difficult for even us humans to estimate a feature $\mathbf{a}_{\theta_0}^m$ from $\mathbf{a}_{\theta_1}^m$ (see Fig. 3 for example). Therefore, when features for more than one view direction (let them be $\theta_j(1), \dots, \theta_j(k)$) are obtained, we can more precisely transform a feature for the view direction θ_i as

$$\mathbf{a}_{\theta_i}^m = P_{\theta_i} \begin{bmatrix} P_{\theta_j(1)} \\ \vdots \\ P_{\theta_j(k)} \end{bmatrix}^+ \begin{bmatrix} \mathbf{a}_{\theta_j(1)}^m \\ \vdots \\ \mathbf{a}_{\theta_j(k)}^m \end{bmatrix}. \quad (10)$$

In the above formulation, there are no constraints for view transformation, but each body point such as head, hands, and knees appears at the same height, respectively, for all view directions because of the height scaling as described in sec. 2. Therefore, we constrain transformation from a height y_i to another height $y_j (\neq y_i)$ and define the above transformation separately at each height y_i .

Moreover, we introduce a simple opposite view transformation. Let the range of a view direction $[\theta_i, \theta_j]$ be $R_{[\theta_i, \theta_j]}$. When a target subject is observed at a distance from a camera and weak perspective projection is assumed, the silhouette image observed with a virtual camera at the opposite side from the view direction¹ θ as shown in Fig. 4 (let the image be $I_{opp}(\theta)$), becomes a mirror image of

¹ Note that the view direction θ is defined for the actual camera and that it is used in common for both the actual and the virtual cameras.

the original silhouette image from view direction θ (let it be $I(\theta)$). In addition, it is clear that $I_{opp}(\theta)$ is the same as $I(\theta + 180)$. Hence, $I(\theta + 180)$ is transformed as a mirror image of $I(\theta)$. In the same way, once the amplitude spectra for $R_{[0,180)}$ are obtained, the remaining features for $R_{[180,360)}$ are obtained by transformation. Thus, a training set for VTM is only composed of features for $R_{[0,180)}$.

5 Experiments

5.1 Datasets

We use a total of 719 gait sequences from 20 subjects for the experiments. The sequences include 24 view directions at every 15 degrees. The training set for the VTM is composed of 120 sequences of 10 subjects from 12 view directions: $\theta = 0, 15, 30, 45, 60, 75, 90, 105, 120, 135, 150,$ and 165 . Then, we prepare 5 gallery sets: $G_0, G_{45}, G_{90}, G_{135}, G_{0-90}$, where G_θ has 20 sequences from 20 subjects with view direction θ , and $G_{\theta_i-\theta_j}$ is a compound gallery of G_{θ_i} and G_{θ_j} ; that is, it has 40 sequences from 20 subjects with 2 views, θ_i and θ_j . A probe set (test set) is composed of the other sequences except for those of subjects included in the training set, and each sequence is indexed in advance with the view direction because view direction estimation is easily done using a walking person’s velocity in the image or by view direction classification with averaged features for each view direction. In the following subsections, for convenience, we represent a gallery transformed by eq. (9) or eq. (10), and probe with view direction θ as G_{s_θ} and Pr_θ , respectively.

5.2 Feature transformation

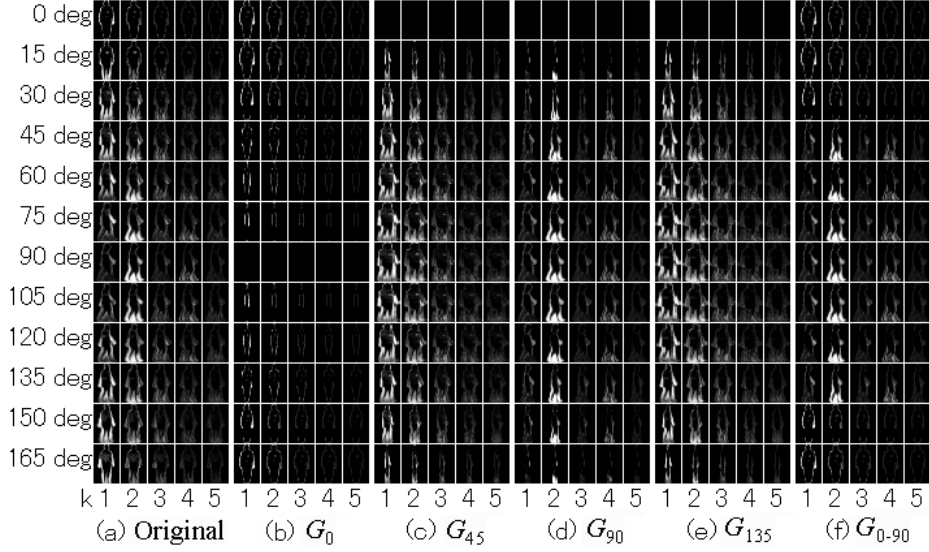
For comparison, we first briefly describe image transformation by perspective projection (PP) [21]. This method approximates that gait motion is represented in the sagittal plane when the person is observed at a distance from a camera. This method cannot transform images if G_0 is given, thus we substitute a longitudinal plane orthogonal to the sagittal plane in such case. Moreover, in the case of G_{0-90} , we use the sagittal plane for $R_{[45,135]}$ and $R_{[225,315]}$ and use the orthogonal plane for the other directions.

We show transformed features using PP in Fig. 5. We can see that the transformed features whose view directions are near those of the original galleries are relative fine (especially $G_{s_{75}}$ and $G_{s_{105}}$ for G_{90}) and that the other features differ a lot from the original features.

We show transformed features with our VTM in Fig. 6. Because G_0 contains relatively few features, the transformed features from G_0 are very poor (Fig. 6(b)). On the other hand, the other view directions contain relatively many features, and the transformed features (Fig. 6(c)-(f)) seem to be similar to the original ones (Fig. 6(a)).

5.3 Performance of gait recognition

We constructed a matching test using the transformed features by both PP and VTM from the 5 above gallery sets. A probe is assigned verification when eq. (6) is above a certain threshold value, and a Receiver Operating Characteristics (ROC) [24] curve is obtained by plotting pairs of verification rate and false alarm



(a): original feature, (b)-(f): transformed features from G_0 , G_{45} , G_{90} , G_{135} , and G_{0-90} respectively. **Fig. 5.** Transformed features with PP

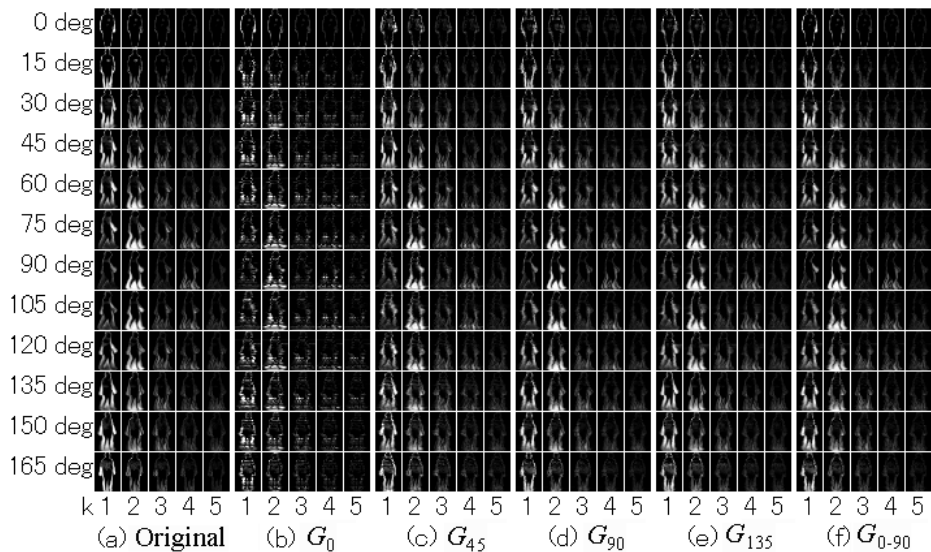
rate for various threshold values. The tests are repeated for different 20 training sets and the averaged performance evaluated by the ROC curve shown in Fig. 7. In this graph, probes are limited to Pr_0 , Pr_{45} , Pr_{90} , and Pr_{135} for visibility.

It is clear that the probes with the same view direction as the gallery have very high performances for all galleries. Then, as seen from the transformed features in the previous subsection, the performances for G_0 are very poor for both PP and VTM. In the other galleries, Pr_{135} for G_{45} and Pr_{45} for G_{135} in PP have relatively high performances; which is why the transformed features for the view directions θ and $(180 - \theta)$ become the same in the case that gait motion is completely symmetric with a phase shift of a half of the gait period. Except for this point, the performances of the VTM are better than those of PP, especially in G_{0-90} (Fig. 7(e)).

Figure 8 shows that the verification rate at a false positive rate (P_F) is 10 % in the ROC curves and the averaged verification rate. For view directions $R_{[180,360)}$, the mirror (horizontally reversed) features are transformed as described in sec. 4.

As shown in Fig. 7, performances for G_0 are very poor in both PP and VTM. As for PP, probes whose view directions are near to those of the gallery have relatively high performances (e.g. Pr_{75} and Pr_{105} for G_{90}) because the weak perspective projection to the sagittal plane works well. In addition, probes with advantages of symmetry (e.g. Pr_{135} for G_{45} and Pr_{45} for G_{135}) also have relatively high performances.

On the other hand, almost all of the other VTM performances except for the above probes are superior to those of PP, especially Pr_{45} , Pr_{135} , Pr_{225} , and Pr_{315} in G_{0-90} achieve fairly good performances compared with PP. As a result, the averaged performance of the VTM is superior to that of PP, except for G_0 .



(a): original feature, (b)-(f): transformed features from G_0 , G_{45} , G_{90} , G_{135} , and G_{0-90} respectively.

Fig. 6. Transformed features with VTM

6 Conclusion and future works

In this paper, we proposed a gait recognition method using amplitude spectra for the temporal axis and our view transformation model (VTM). First, a walking person is extracted utilizing temperature-based background subtraction using an infrared-ray camera, and the gait silhouette volume (GSV) is constructed by scaling and registering the silhouette images. Then the gait period is detected by normalized autocorrelation, and the amplitude spectra of the GSV are calculated by Fourier analysis based on the gait period. After the VTM is obtained with a training set of multiple subjects from multiple view directions, the features of various view directions can be made by transformation from features of one or a few of the view directions. We made experiments using 719 sequences from 20 subjects of the 24 view directions. As a result, the proposed methods achieve higher performance than the previously proposed perspective projection (PP) method.

Future works are as follows.

- Combination of VTM and PP for better view change adaptation.
- Experiments for a general database, such as the HumanID Gait Challenge Problem Datasets [6].

References

1. Cunado, D., Nixon, M.S., and Carter, J.N.: Automatic Extraction and Description of Human Gait Models for Recognition Purposes, *Computer Vision and Image Understanding*, Vol. 90, No. 1, (2003) 1–41
2. Yam, C., Nixon, M.S., and Carter, J.N.: Automated Person Recognition by Walking and Running via Model-based Approaches, *Pattern Recognition*, Vol. 37, No. 5, (2004) 1057–1072

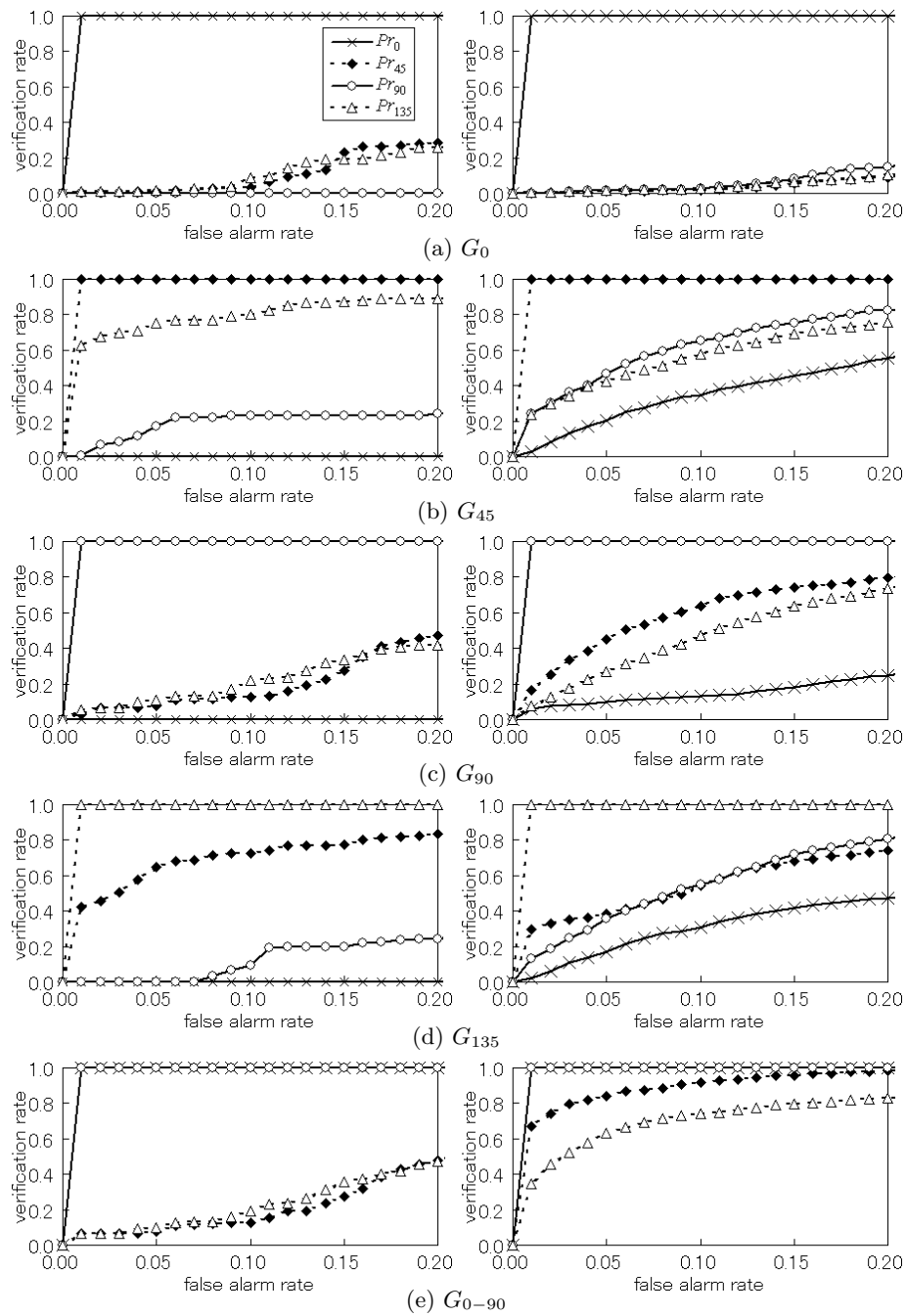


Fig. 7. ROC curves of gait recognition performance for PP (left side) and VTM (right side). Legend marks are common in all graphs.

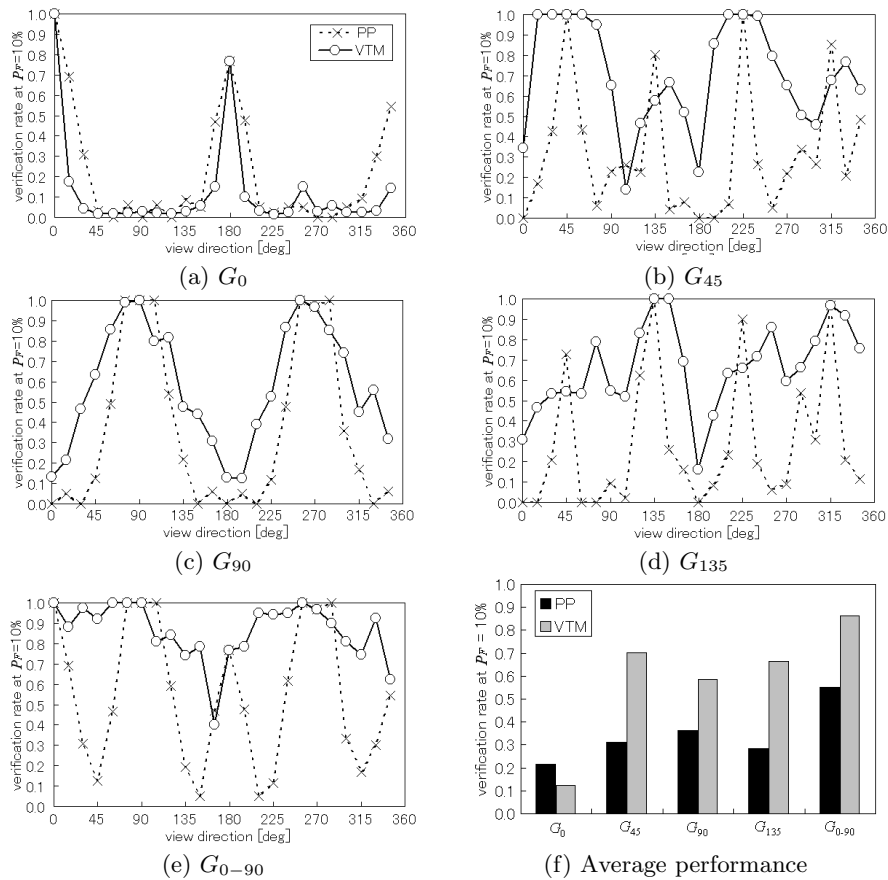


Fig. 8. Performance comparison of PP and VTM with verification rate at $P_F = 10\%$. Legend marks are common in (a)-(e).

3. Bobick, A.F. and Johnson, A.Y.: Gait Recognition using Static Activity-specific Parameters, Proc. of Computer Vision and Pattern Recognition, Vol. 1, (2001) 423–430
4. Wagg, D.K. and Nixon, M.S: On Automated Model-Based Extraction and Analysis of Gait, Proc. of the 6th IEEE Int. Conf. on Automatic Face and Gesture Recognition, (2004) 11–16
5. Urtasun, R. and Fua, P.: 3D Tracking for Gait Characterization and Recognition, Proc. of the 6th IEEE Int. Conf. on Automatic Face and Gesture Recognition, (2004) 17–22
6. Sarkar, S., Phillips, J.P., Liu, Z., Vega, I.R., Grother, P., and Bowyer, K.W.: The HumanID Gait Challenge Problem: Data Sets, Performance, and Analysis, Trans. of Pattern Analysis and Machine Intelligence, Vol. 27, No. 2, (2005) 162–177
7. Murase, H. and Sakai, R.: Moving Object Recognition in Eigenspace Representation: Gait Analysis and Lip Reading, Pattern Recognition Letters, Vol. 17, (1996) 155–162
8. Ohara, Y., Sagawa, R., Echigo, T., and Yagi, Y.: Gait volume: Spatio-temporal analysis of walking, Proc. of the 5th Workshop on Omnidirectional Vision, Camera

- Networks and Non-classical cameras, (2004) 79–90
9. Niyogi, S and Adelson, E.: Analyzing and recognizing walking figures in xyt, Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, (1994) 469–474
 10. Liu, Z and Sarkar, S.: Simplest Representation Yet for Gait Recognition: Averaged Silhouette, Proc. of the 17th Int. Conf. on Pattern Recognition, Vol. 1, (2004) 211–214
 11. BenAbdelkader, C., Culter, R., Nanda, H., and Davis, L.: Eigengait: Motion-based recognition people using image self-similarity, Proc. of Int. Conf. on Audio and Video-based Person Authentication, (2001) 284–294
 12. Cuntoor, N, Kale, A, and Chellappa, R: Combining Multiple Evidences for Gait Recognition, Proc. of IEEE Int. Cont. Acoustics, Speech, and Signal Processing, Vol. 3, (2003) 33–36
 13. Liu, Y., Collins, R.T., and Tsin, Y.: Gait sequence analysis using frieze patterns, Proc. of the 7th European Conf. on Computer Vision, Vol. 2, (2002) 657–671
 14. Kobayashi, T. and Otsu, N.: Action and Simultaneous Multiple-Person Identification Using Cubic Higher-Order Local Auto-Correlation, Proc. of the 17th Int. Conf. on Pattern Recognition, Vol. 3, (2004) 741–744
 15. Mowbray, S.D and Nixon, M.S.: Automatic Gait Recognition via Fourier Descriptors of Deformable Objects, Proc. of IEEE Conf. on Advanced Video and Signal Based Surveillance, (2003) 566–573
 16. Zhao, G., Chen, R., Liu, G., and Li, H.: Amplitude Spectrum-based Gait Recognition, Proc. of the 6th IEEE Int. Conf. on Automatic Face and Gesture Recognition, (2004) 23–30
 17. BenAbdelkader, C.: Gait as a Biometric For Person Identification in Video, Ph.D. thesis in Maryland Univ., (2002)
 18. Yu, S., Tan, D., and Tan, T.: Modelling the Effect of View Angle Variation on Appearance-Based Gait Recognition, Proc. of the 7th Asian Conf. on Computer Vision, (2006), 807–816
 19. Shakhnarovich, G., Lee, L, and Darrell, T.: Integrated Face and Gait Recognition from Multiple Views, Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, Vol. 1, (2001) 439–446
 20. Lee, L.: Gait Analysis for Classification, Ph.D. thesis in Massachusetts Institute of Technology, (2002)
 21. Kale, A., Chowdhury, K.R., and Chellappa, R.: Towards a View Invariant Gait Recognition Algorithm, Proc. of IEEE Conf. on Advanced Video and Signal Based Surveillance, (2003) 143–150
 22. Mukaigawa, Y., Nakamura, Y, and Ohta, Y.: Face Synthesis with Arbitrary Pose and Expression from Several Images - An integration of Image-based and Model-based Approach - Proc. of the 3rd Asian Conf. on Computer Vision, Vol. 1, (1998), 680–687.
 23. Utsumi, A., Tetsutani, N.: Adaptation of appearance model for human tracking using geometrical pixel value distributions, Proc. of the 6th Asian Conf. on Computer Vision, (2004)
 24. Phillips, P.J., Moon, H., Rizvi, S, and Rauss, P: The FERET Evaluation Methodology for Face-Recognition Algorithms, Trans. of Pattern Analysis and Machine Intelligence, Vol. 22, No. 10, (2000), 1090–1104