

# Adaptation to Walking Direction Changes for Gait Identification

Yasushi Makihara, Ryusuke Sagawa, Yasuhiro Mukaigawa, Tomio Echigo, Yasushi Yagi  
Osaka University, The Institute of Scientific and Industrial Research  
8-1 Mihogaoka, Ibaraki, Osaka, 567-0047, JAPAN  
{makihara,sagawa,mukaigaw,echigo,yagi}@am.sanken.osaka-u.ac.jp

## Abstract

*This paper describes adaptation to gradual changes of walking directions for gait identification. First, we propose a method of body tilt correction due to changes of walking directions when constructing a spatio-temporal gait silhouette volume. Next, we propose a view transformation model in the frequency domain to match gait features of different walking directions. Finally, experiments of gait identification for a circular path 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, hand vein, iris, face, and gait. Among these, gait identification has recently gained considerable attention because gait is a promising cue for surveillance systems to ascertain identity at a distance from a camera.

Many approaches of gait identification were proposed [1][2][7][5][4]. Generally, a gait motion is observed the best in a fronto-parallel view, hence most of the approaches assume the gait is observed in near fronto-parallel views. In addition, many approaches treated sequences with constant walking directions. However, these assumptions are unreasonable in actual surveillance scenarios such as case of walking on a curved road.

Therefore, we tackle with adaptation to gradual changes of walking directions for gait identification. First, we introduce a view transformation model (VTM) in the frequency-domain to match gait features of different walking directions. In the proposed method, once we obtain a VTM using a training set, made up of features of multiple subjects from multiple views, we can make features of a new subject taken from the multiple view directions by using features of the subject from a few views.

In addition, gradual changes of walking directions in a sequence induce other difficulties such as a body tilt by a centrifugal force. To overcome this, we introduce a mechanics-based body tilt correction using an estimated tra-

jectory of a walking person.

## 2 Construction of a gait silhouette volume

As the first step in gait identification, we describe construction of a gait silhouette volume (GSV) in this section. Moreover, we present a method of correcting body tilts induced by gradual changes of walking directions.

### 2.1 Trajectory estimation

First, gait silhouettes are simply extracted by background subtraction. Here, for simplicity we assume only one person exists in the image, thus we keep only the largest connected silhouette region as the person.

Next, the top, the bottom, and the horizontal center of the regions for each frame are obtained. The horizontal center is chosen as the median of horizontal positions in the region. Then, a moving average filter of 30 frames is applied to those positions. We simply obtain the position  $(Z, X)$  of the person in a camera coordinate (see Fig. 1) based on the object-to-image height ratio in the same way as [3],

$$X = (H/h)x, Z = (H/h)f, \quad (1)$$

where  $H$  and  $h$  are the heights of the person in the camera coordinate and in the image plane respectively, and  $f$  is the focal length. We assume that  $H$  is known for each person and that  $f$  is obtained by camera calibration in advance. An example of an estimated trajectory for a circular path is shown in Fig. 2.

### 2.2 Body tilt correction

When the person gradually changes his/her walking directions, he/she receives not only the gravitational force  $F_G (= mg)$  but also the centrifugal force  $F_C$ , where  $m$  is the mass of the person and  $g$  is the gravitational acceleration. Thus, these forces sum up to the pseudo gravitational

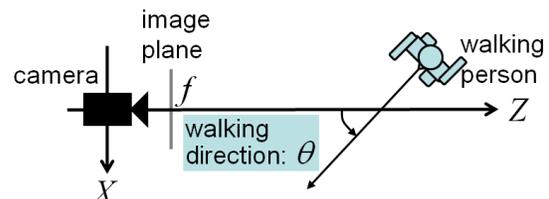


Figure 1. Definition of camera coordinate and walking direction  $\theta$  at top view

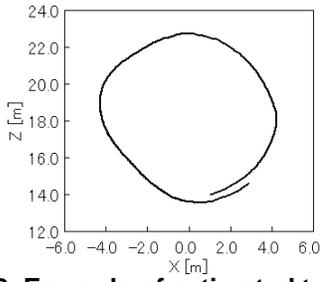


Figure 2. Example of estimated trajectory

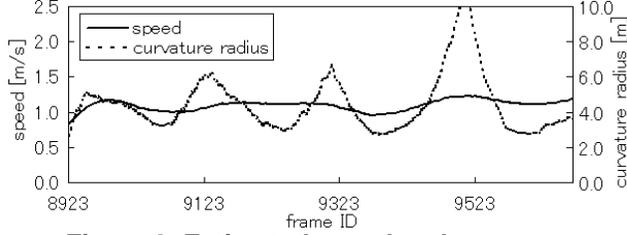


Figure 3. Estimated speed and curvature

force  $\mathbf{F}'_G$  and make his/her body tilt into the same direction as  $\mathbf{F}'_G$  (see Fig. 4(a)) to keep his/her balance. When matching gait sequences with and without walking direction changes in the same condition, we should correct these body tilts.

First, we try to obtain body tilt angle  $\phi$ . When the gait trajectory is assumed to consist of sets of arc trajectories, the centrifugal force  $\mathbf{F}_C$  is

$$|\mathbf{F}_C| = m|\mathbf{v}|^2/r, \quad (2)$$

where  $\mathbf{v}$  is the walking velocity, and  $r$  is the signed curvature radius (positive for counterclockwise direction), which are simply obtained by the walking trajectories. Estimated walking speed and curvature radius for the circular path in Fig. 2 are shown in Fig. 3. Thus, the body tilt angle  $\phi$  is

$$\phi = \tan^{-1}(|\mathbf{F}_C|/|\mathbf{F}'_G|) = \tan^{-1}(|\mathbf{v}|^2/r|g|). \quad (3)$$

Figure 5 shows estimated walking direction  $\theta$  and body tilt  $\phi$  for the circular path in Fig. 2. Note that the walking direction  $\theta$  is defined as Fig. 1 and easily obtained by the walking velocity  $\mathbf{v}$ .

Next, we modify the silhouette images based on the obtained body tilt angle  $\phi$  and walking direction  $\theta$ . In the image plane, the horizontal shift of the top position after body tilt is obtained as  $h \sin \phi \cos \theta$  by assuming weak perspective projection (see Fig. 4(b)). Then an apparent height, that is, the distance between the bottom and the top after tilt is obtained as  $h \sqrt{\cos^2 \phi + \sin^2 \phi \cos^2 \theta}$  in the same way. In

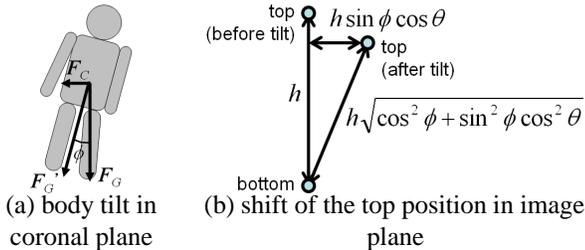


Figure 4. Body tilt by centrifugal force

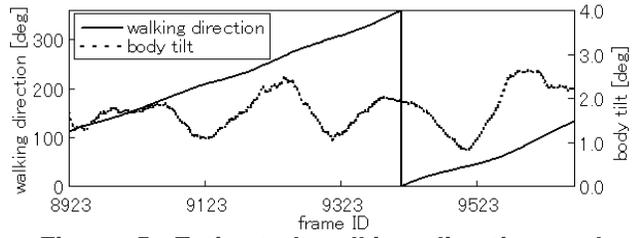


Figure 5. Estimated walking direction and body tilt for circular path

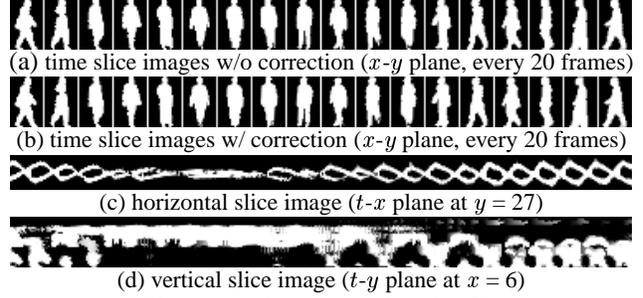


Figure 6. An example of GSV

order to transform so that the top after tilt can return to that before tilt, the region is rotated by  $\phi'$  and is scaled by  $s_y$  for the vertical direction, where

$$\phi' = -\tan^{-1}(\tan \phi \cos \theta) \quad (4)$$

$$s_y = 1/\sqrt{\cos^2 \phi + \sin^2 \phi \cos^2 \theta}. \quad (5)$$

### 2.3 Scaling and registration

The last step is scaling and registration of the extracted silhouette images. The silhouette images are scaled so that the height  $h$  can be just 30 pixels, and so that the aspect ratio of each region can be kept. Finally, we produce a  $20 \times 30$  pixel-sized image in which the horizontal median corresponds to the horizontal center of the image.

We show an example of a constructed GSV in Fig. 6 as time slice ( $x$ - $y$  plane) without and with body tilt correction, horizontal slice ( $t$ - $x$  plane), and vertical slice ( $t$ - $y$  plane) images. We can see that counterclockwise body tilts in Fig. 6(a) are corrected in Fig. 6(b). Moreover we can see gait periodicity from Fig. 6(c), (d).

## 3 Matching of a GSV

In this section, we present extraction of frequency-domain features normalized with gait period and matching measure between the features.

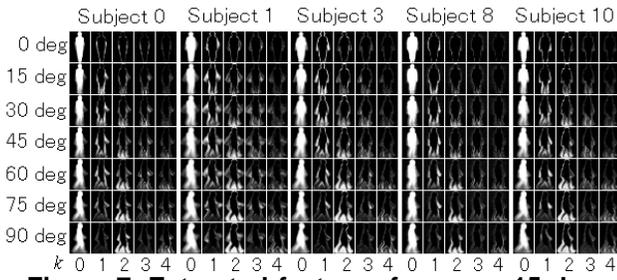
### 3.1 Gait period detection

First, gait period  $N_{gait}$  is detected by the normalized autocorrelation of a GSV for the temporal axis as

$$N_{gait} = \arg \max_{N \in [N_{min}, N_{max}]} C(N) \quad (6)$$

$$C(N) = \frac{\sum_{x,y} \sum_{n=0}^{T(N)} g(x,y,n)g(x,y,n+N)}{\sqrt{\sum_{x,y} \sum_{n=0}^{T(N)} g(x,y,n)^2} \sqrt{\sum_{x,y} \sum_{n=0}^{T(N)} g(x,y,n+N)^2}}, \quad (7)$$

where  $C(N)$  is the autocorrelation for the  $N$  frame shift,  $g(x, y, n)$  is the silhouette value at position  $(x, y)$  at the  $n$ th



**Figure 7. Extracted features for every 15-degree view direction for some subjects**

frame,  $T(N)$  is defined as  $(T_{total} - N - 1)$ , and  $T_{total}$  is the number of total frames in the sequence. We set  $N_{min} = 20$  and  $N_{max} = 40$  empirically as the natural gait period.

### 3.2 Frequency-domain features

First, we pick up the subsequences  $\{S_i\}$  ( $i = 1, 2, \dots, N_{sub}$ ) for every  $N_{gait}$  frames from a total sequence  $S$ . Note that the frame range of the  $i$ th subsequence  $S_i$  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 normalized with the gait period are subsequently calculated as

$$G_i(x, y, k) = \sum_{n=iN_{gait}}^{(i+1)N_{gait}-1} g(x, y, n) e^{-j\omega_0 kn} \quad (8)$$

$$A_i(x, y, k) = \frac{1}{N_{gait}} |G_i(x, y, k)|, \quad (9)$$

where  $\omega_0$  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 frequency, and  $A_i(x, y, k)$  is an amplitude spectrum for  $G_i(x, y, k)$ . In this paper,  $A_i(x, y, k)$  ( $k = 0, \dots, 4$ ) is used as the gait feature and its dimension  $N_A$  sums up to  $20 \times 30 \times 5 = 3000$ .

Figure 7 shows extracted amplitude spectra of straight-walk sequences for various directions. Note that the unit of the walking direction is a degree in this paper. Amplitude spectra vary widely among directions for each subject, and to some extent they also have individual variations for each direction. Moreover, we can see that all the subjects have similar common tendencies for amplitude spectra variations across direction changes. This fact indicates a real possibility that the variations across direction changes are expressed with the VTM independently of individual variations (see sec. 4).

### 3.3 Matching measures

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

$$d(S_i, S_j) = \|\mathbf{a}(S_i) - \mathbf{a}(S_j)\|. \quad (10)$$

Next, we define a matching measure between two total sequences. Let  $S_P$  and  $S_G$  be total sequences for probe and gallery, respectively, and let  $\{S_{P_i}\}$  ( $i = 1, 2, \dots$ ) and  $\{S_{G_j}\}$  ( $j = 1, 2, \dots$ ) be their subsequences, respectively. Gallery subsequences  $\{S_{G_j}\}$  have variations in general and

probe subsequences  $\{S_{P_i}\}$  may contain outliers. A measure candidate  $D(S_P, S_G)$  to cope with them is the median value of the minimum distances of each probe subsequence  $S_{P_i}$  and gallery subsequences  $\{S_{G_j}\}$  ( $j = 1, 2, \dots$ ):

$$D(S_P, S_G) = \text{Median}_i[\min_j\{d(S_{P_i}, S_{G_j})\}]. \quad (11)$$

## 4 VTM in the frequency domain

We briefly describe the formulation of a VTM in a way similar to that in [8]. Note that we apply the model to the frequency-domain feature extracted from gait image sequences while that in [8] directly applied it to a static image.

We first quantize view (walking) 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

$$\begin{bmatrix} \mathbf{a}_{\theta_1}^1 & \dots & \mathbf{a}_{\theta_1}^M \\ \vdots & \ddots & \vdots \\ \mathbf{a}_{\theta_K}^1 & \dots & \mathbf{a}_{\theta_K}^M \end{bmatrix} = USV^T = \begin{bmatrix} P_{\theta_1} \\ \vdots \\ P_{\theta_K} \end{bmatrix} [\mathbf{v}^1 \dots \mathbf{v}^M], \quad (12)$$

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,  $P_{\theta_k}$  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  $P_{\theta_k}$  is a projection matrix from the intrinsic vector  $\mathbf{v}$  to the feature vector for view direction  $\theta_k$ , 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  $\theta_i$  of the  $m$ th subject is represented as

$$\mathbf{a}_{\theta_i}^m = P_{\theta_i} \mathbf{v}^m. \quad (13)$$

Then, feature vector transformation from view direction  $\theta_j$  to  $\theta_i$  is easily obtained as

$$\hat{\mathbf{a}}_{\theta_i}^m = P_{\theta_i} P_{\theta_j}^+ \mathbf{a}_{\theta_j}^m, \quad (14)$$

where  $P_{\theta_j}^+$  is the pseudo inverse matrix of  $P_{\theta_j}$ . 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  $\mathbf{a}_{90}^m$  from  $\mathbf{a}_0^m$  (see Fig. 7 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  $\theta_i$  as

$$\hat{\mathbf{a}}_{\theta_i}^m = P_{\theta_i} \begin{bmatrix} P_{\theta_j(1)} \\ \vdots \\ P_{\theta_j(k)} \end{bmatrix}^+ \begin{bmatrix} \mathbf{a}_{\theta_j(1)}^m \\ \vdots \\ \mathbf{a}_{\theta_j(k)}^m \end{bmatrix}. \quad (15)$$

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  $y_i$ .

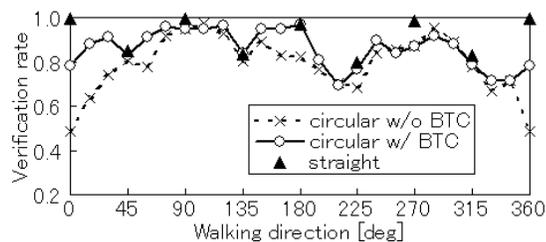


Figure 8. Verification rate for each direction

## 5 Experiments

In this section, experiments of gait identification for straight-walk and circular-walk sequences are described.

### 5.1 Datasets

We use a total of 744 straight-walk sequences from 20 subjects and 15 sequences of 3 rounds of approximately circular-walk sequences from 15 subjects. The straight-walk sequences are quantized at every 15 directions, and then the number  $K$  of directions is 24. The training set for the VTM is composed of 480 straight-walk sequences of 20 subjects from 24 view directions. The gallery sets from 24 directions for each subject are constructed by transforming straight-walk sequences from 3 directions: 0, 90, and 180 degrees based on the trained VTM.

Probe sets (test sets) are composed of the other straight-walk sequences and the circular-walk sequences. For the circular-walk sequences, we extracted two types of features with and without body tilt correction. In the following subsections, for convenience, we denote probe sets of the straight-walk sequence, and the circular-walk sequences with and without body tilt correction as "straight", "circular w/ BTC", and "circular w/o BTC", respectively.

### 5.2 Performance of gait identification

In matching process for circular-walk sequences, we calculate an average walking direction  $\bar{\theta}$  for each subsequence, and match to galleries whose direction is  $\bar{\theta}$  or  $(\bar{\theta} \pm 15)$  degrees.

We evaluate a performance by a verification rate at 10% false positive rate in a Receiver Operating Characteristics (ROC) [6] curve. We first show the verification rate when each subsequence including one gait period is matched separately for each walking direction in Fig. 8. In this time, directions of "straight" are limited for every 45 degrees. We can see that body tilt correction makes the performance better, especially for near frontal-view subsequences which affected severely by body tilt rotation in the image. In addition, the performance of "circular w/ BTC" is competitive to that of "straight".

Next, we show average verification rates when multiple periods of subsequences are used to matching in Fig. 9. As a result, the performance improves as the number of periods increase, and the performance of "circular w/ BTC" is always superior to that of "circular w/o BTC". Moreover the performance of "circular w/ BTC" is competitive to that of "straight" when the number of periods is more than 3.

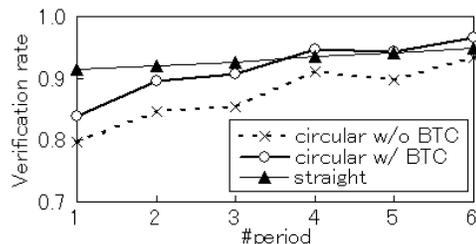


Figure 9. Verification rate for each #period

## 6 Conclusion and future works

In this paper, we proposed a method of adaptation to gradual changes of walking directions for gait identification. First, we proposed a body tilt correction based on an estimated centrifugal force when constructing a gait silhouette volume (GSV). Second, we introduced a view transformation model (VTM) in the frequency domain to transform the gallery features into the same walking direction as that of an probe feature. Finally, we demonstrated the effectiveness of the proposed method by experiments of gait identification for straight-walk and circular-walk sequences.

Future works are as follows.

- Adaptation to appearance changes due to camera tilt
- Experiments for a general database, such as the HumanID Gait Challenge Problem Datasets [7].

## References

- [1] A. Bobick and A. Johnson. Gait recognition using static activity-specific parameters. In *Proc. of Computer Vision and Pattern Recognition*, volume 1, pages 423–430, 2001.
- [2] 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.
- [3] M. Hild. Estimation of 3d motion trajectory and velocity from monocular image sequences in the context of human gait recognition. In *Proc. of the 17th Int. Conf. on Pattern Recognition*, volume 1, pages 231–235, 2004.
- [4] T. Kobayashi and N. Otsu. Action and simultaneous multiple-person identification using cubic higher-order local auto-correlation. In *Proc. of the 17th Int. Conf. on Pattern Recognition*, volume 3, pages 741–744, 2004.
- [5] Y. Ohara, R. Sagawa, T. Echigo, and Y. Yagi. Gait volume: Spatio-temporal analysis of walking. In *Proc. of the 5th Workshop on Omnidirectional Vision, Camera Networks and Non-classical cameras*, pages 79–90, 2004.
- [6] P. Phillips, H. Moon, S. Rizvi, and P. Rauss. The feret evaluation methodology for face-recognition algorithms. *Trans. of Pattern Analysis and Machine Intelligence*, 22(10):1090–1104, 2000.
- [7] S. Sarkar, J. Phillips, Z. Liu, I. Vega, P. Grother, and K. Bowyer. The humanid gait challenge problem: Data sets, performance, and analysis. *Trans. of Pattern Analysis and Machine Intelligence*, 27(2):2005, 162-177.
- [8] 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.