2012 IEEE International Conference on Systems, Man, and Cybernetics
October 14-17, 2012, COEX, Seoul, Korea

Gait Recognition using Images of Oriented Smooth Pseudo Motion

Yasushi Makihara, Betria Silvana Rossa, Yasushi Yagi
The Institute of Scientific and Industrial Research
Osaka University
Ibaraki, Osaka, Japan
{makihara, rossa, yagi}@am.sanken.osaka-u.ac.jp

***Abstract*—This paper proposes a method of gait recognition using not only shape feature but also motion feature from silhouette image sequences. The inner silhouette motion called pseudo motion is constructed by dividing the silhouette shape into small clusters and by computing many-to-many correspondence via earth mover's morphing framework. The raw pseudo motion, however, tends to be locally fluctuated in the spatio-temporal domain, and hence the spatio-temporal regularization is imposed to provide the smooth pseudo motion. The smooth pseudo motion image sequences are further partitioned into images with eight different orientations, and then averaged over each gait period to produce images of oriented smooth pseudo motion. Both shape and motion cues are integrated in score-level fusion framework based on linear logistic regression and the single-dimensional fused distance is returned by the learned optimal weights. The experiments with the publicly available gait database show the effectiveness of the proposed method compared with the case where the shape information is used alone.**

***Index Terms*—Gait recognition, Pseudo motion, Earth mover's distance, Score-level fusion**

I. Introduction

Gait recognition is one of the behavioral biometrics which has attracted much attention because it can be applied to uncooperative subjects and/or subjects at a distance from cameras. In fact, a burglar's gait captured by a CCTV camera was accepted as evidence in a UK court [1].

Due to such promising properties, many efforts on gait recognition have been done for the last decade in computer vision, pattern recognition, and biometrics communities. Approaches to gait recognition mainly fall into two categories: model-based and model-free approaches.

In the model-based approaches, a human body is usually modeled as articulated body composed of links and joints and the model parameters such as joint angles, joint position, link length and width are estimated by fitting the model to the observation (e.g., silhouette image or edge image). Bobick et al. [2] extracted parameters of shape and stride, Wagg et al. [3] extracted static shape parameters and gait period with an articulated body model. Urtasun et al. [4] extracted joint angles with a 3-D articulated body model. Cunado et al. [5] and Yam et al. [6] extracted periodical features of leg motion by Fourier analysis, and Goffredo et al. [7] extracts trajectories of joint positions. Although some of model-based approaches have the advantages of view-invariant and clothing-invariant properties, they often requires high computational cost and are error-prone in model fitting stage.

On the contrary, the model-free approaches, also known as appearance-based approaches directly analyze holistic images to extract features without the body models and have been widely used in gait recognition community. The model-free approaches are further divided into two categories: textured-based and silhouette-based approaches. In the textured-based approaches, Niyogi et al. [8] and Ohara et al. [9] constructed a spatio-temporal volume (x-y-t volume) by combining gait images and matched features extracted from the volume. BenAbdelkader et al. [10] utilizes self-similarity plot as motion features, while Little and Boyd [11] extracts optical flows as motion features. The textured-based approaches, however, have a disadvantage that clothes difference between a gallery and a probe may degrade the gait recognition performance.

Hence, the silhouette-based approaches are dominant in the gait recognition community. Sarkar et al. [12] proposes a baseline algorithm where silhouette sequences are directly matched to each other by synchronizing phase. Liu et al. [13], Veres et al. [14], and Han et al. [15] propose a simple averaged silhouette-based gait representation, also known as gait energy image (GEI), where the silhouettes from gait image sequences are averaged over one gait cycle. Bashira et al. [16] propose a gait entropy image (GEnI) and utilize it to mask out the static part of the GEI to enhance clothing-invariant or carrying-invariant properties. Wang et al. [17] propose a chrono-gait image where phase information is encoded as color information in the process of averaging over a quarter gait cycle.

Whereas the above approaches are based on silhouette sequence, several approaches try extracting the explicit motion information from the silhouette sequence. Lam et al. [18] directly exploits motion information, namely gait flow image (GFI), which is obtained by averaging the optical flow on the contour of the silhouette over one gait cycle. The method, however, relies only on contour motion and discards useful inner silhouette information.

Therefore, we try exploiting useful inner silhouette motion in this paper, which have potentially much more motion information than the contour-based optical flows [18]. For this purpose, a shape morphing technique so-called earth mover's morphing (EMM) [19] is employed to extract the inner silhouette motion. In this framework, a silhouette image is

978-1-4673-1714-6/12/$31.00 ©2012 IEEE

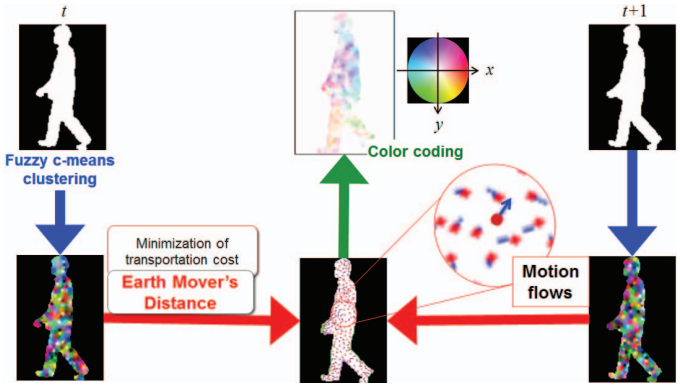
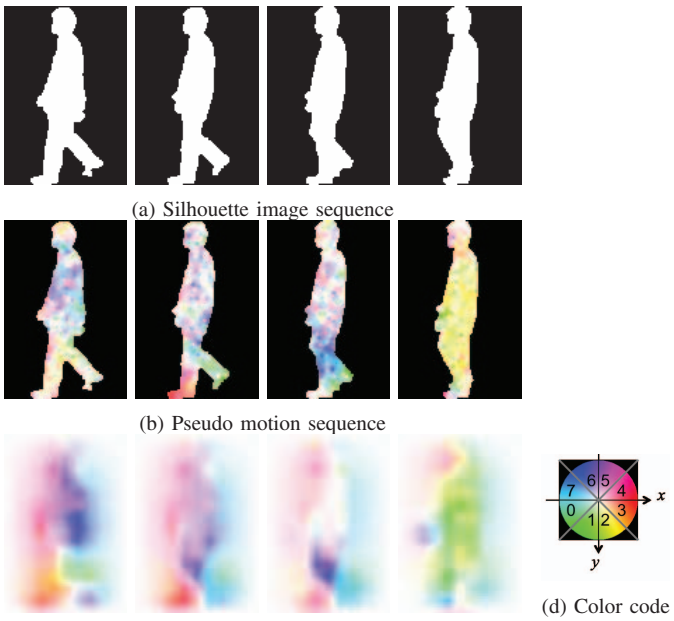


Fig. 1. Overview of pseudo motion extraction



(a) Silhouette image sequence

(b) Pseudo motion sequence

(c) Smooth pseudo motion sequence

(d) Color code

Fig. 2. Pseudo motion

clustered into small regions and many-to-many correspondence is computed so as to minimize the transportation cost via earth mover's distance (EMD). Because this many-to-many correspondence is regarded as a kind of motion information (but not actual motion information), this is called *pseudo motion* in this paper.

The pseudo motion, however, tends to be locally fluctuated in the spatio-temporal domain due to random clustering processes. Therefore, the spatio-temporal regularization is imposed to provide smooth pseudo motion. Moreover, due to the efficacy of feature averaging over one gait period (e.g., GEI), the smooth pseudo motion is further partitioned into eight different directions and the motion intensity is averaged over one gait period for individual directions, which construct images of oriented smooth pseudo motion (IOSPM). Finally, the matching scores by the baseline algorithm [12] and by the IOSPM are combined in the score-level fusion framework [20] to get better performance.

The remainder of this paper is organized as follows. Section II presents extraction of EMD-based pseudo motion and its variations. Section III discusses the methods to calculate the distances between gallery and probe and Section IV describes a score-level fusion technique of both shape and motion cues. Section V presents experiments using a publicly available gait database. Section VI gives conclusions of this paper and also future work.

II. Feature Extraction

A. EMD-based Pseudo Motion Extraction

In this section, EMD-based pseudo motion extraction is briefly presented as shown in Fig. 1. Because the pseudo motion corresponds to the intermediate results in the shape morphing process, the details are found in [19].

First, a silhouette image $I(\boldsymbol{x}, t)$ at t-th frame is defined as

$$I(\boldsymbol{x}, t) = \begin{cases} 1 & \text{for outside shape} \\ 0 & \text{for inside shape} \end{cases}, \quad (1)$$

where $\boldsymbol{x} = [x, y]^T$ is a two-dimensional position in the image. A set of points within the shape is indicated by $X_s(t) = \{\boldsymbol{x} | I(\boldsymbol{x}, t) = 1\}$.

The silhouette image $I(\boldsymbol{x}, t)$ is then decomposed into a number of small clusters by Fuzzy c-means clustering [21] (see Fig. 1, left-bottom and right-bottom). Let the mean position and weight of i-the cluster ($i = 0, \ldots, N_t - 1$) at t-th frame be $\bar{\boldsymbol{x}}_{t,i} = [\bar{x}_{t,i}, \bar{y}_{t,i}]^T$ and $w_{t,i}$, respectively. Next, the EMD d_{EMD} between the small clusters of t-th and $(t+1)$-th frames is calculated so as to minimize the transportation cost as

$$d_{EMD} = \min_{\{m_{i,j}\}} \sum_{i=0}^{N_t-1} \sum_{j=0}^{N_{t+1}-1} m_{i,j} \|\bar{\boldsymbol{x}}_{\boldsymbol{t},\boldsymbol{i}} - \bar{\boldsymbol{x}}_{\boldsymbol{t+1},\boldsymbol{j}}\| \quad (2)$$

$$\text{s.t.} \quad \sum_{i=0}^{N_t-1} m_{i,j} = w_{t+1,j} \quad (3)$$

$$\sum_{j=0}^{N_{t+1}-1} m_{i,j} = w_{t,i} \quad (4)$$

$$m_{i,j} \geq 0, \quad (5)$$

where $m_{i,j}$ is a flow amount from i-th cluster at t-th frame to j-th cluster at $(t+1)$-th frame. Note that only non-zero flow amount $m_{i,j} > 0$ corresponds to the weights for effective many-to-many correspondence between the small clusters. In this way, we obtain M_t effective flows composed of a triplets of the flow amount $m_{t,s}$, the original points in the spatio-temporal domain $\boldsymbol{x}_{t,s} = [x_{t,s}, y_{t,s}, t]^T$, and spatial displacement (pseudo motion) $\boldsymbol{f}_{t,s} = [f^x_{t,s}, f^y_{t,s}]^T$ as $\{m_{t,s}, \boldsymbol{x}_{t,s}, \boldsymbol{f}_{t,s}\}$ ($s = 0, \ldots, M_t - 1$). The resultant flow and its color-coding version are shown at bottom-middle and top-middle images in Fig. 1, respectively.

This process is repeated for all the adjacent frame pairs. As a result, a pseudo motion sequence is extracted from a silhouette image sequence as shown in Figs. 2(a) and (b).

B. Smooth Pseudo Motion Extraction

Because the raw pseudo motion tends to be locally fluctuated in the spatio-temporal domain as shown in Fig. 2(b), the spatio-temporal regularization is imposed. For this purpose, control points are placed on an image sequence, namely, a spatio-temporal three-dimensional space, and then finding the smooth pseudo motion at each control point.

Intervals of the control points for each x, y, and t axes are set to Δx, Δy, and Δt, respectively. The spatio-temporal position set of the control points are then defined as $X^{cp} = \{\boldsymbol{x}^{cp}_{i,j,k}\}$ $(i = 0, \cdots, N^{cp}_x - 1;\ j = 0, \cdots, N^{cp}_y - 1;\ k = 0, \cdots, N^{cp}_t - 1)$, where $\boldsymbol{x}^{cp}_{i,j,k}$ is a spatio-temporal position of i-th horizontal, j-th vertical, and k-th temporal control point, and N^{cp}_x, N^{cp}_y, and N^{cp}_t are the numbers of the control points for each horizontal, vertical, and temporal axes, respectively.

A set of the smooth pseudo motions $F^{cp} = \{\boldsymbol{f}^{cp}_{i,j,k}\}$ at control points are estimated by minimizing the energy function

$$S(F^{cp}) = S_D(F^{cp}) + S_S(F^{cp}) + S_B(F^{cp}) \quad (6)$$

$$F^{cp*} = \arg\min_{F^{cp}} S(F^{cp}), \quad (7)$$

where S_D, S_S, and S_B are respectively data term, smooth term, and background term.

The data term is defined by the sum of squared errors between a raw s-th pseudo motion $\boldsymbol{f}_{t,s}$ at t-th frame and its linearly interpolated flows $\hat{\boldsymbol{f}}(\boldsymbol{x}_{t,s}, X^{cp}, F^{cp})$ derived from the original point $\boldsymbol{x}_{t,s}$, the surrounding control points X^{cp} and their smooth pseudo motion F^{cp}, as

$$S_D(F^{cp}) = \sum_{t=0}^{N-1} \sum_{s=0}^{M_t-1} m_{t,s} \left\| \hat{\boldsymbol{f}}(\boldsymbol{x}_{t,s}, X^{cp}, F^{cp}) - \boldsymbol{f}_{t,s} \right\|^2 \quad (8)$$

The smooth term drives flows on the adjacent control points to be close to each other and is defined as

$$S_S(F^{cp}) = \lambda_S \sum_{i=0}^{N^{cp}_x-1} \sum_{j=0}^{N^{cp}_y-1} \sum_{k=0}^{N^{cp}_t-1} \{ \|\boldsymbol{f}^{cp}_{i+1,j,k} - \boldsymbol{f}^{cp}_{i,j,k}\|^2 \quad (9)$$

$$+ \|\boldsymbol{f}^{cp}_{i,j+1,k} - \boldsymbol{f}^{cp}_{i,j,k}\|^2 + \|\boldsymbol{f}^{cp}_{i,j,k+1} - \boldsymbol{f}^{cp}_{i,j,k}\|^2 \}, \quad (10)$$

where λ_S is the coefficient which controls the strength of the smoothness.

The background term is introduced to attenuate the pseudo motion outside the silhouette, where there should not be any motion flow, it forces the flows in the background area remain close to zero as

$$S_B(F^{cp}) = \lambda_B \sum_{i=0}^{N^{cp}_x-1} \sum_{j=0}^{N^{cp}_y-1} \sum_{k=0}^{N^{cp}_t-1} \|\boldsymbol{f}^{cp}_{i,j,k}\|^2 \delta_{I(i\Delta x, j\Delta y, k\Delta t), 0}, \quad (11)$$

where the coefficient λ_B determines how strong the constrain should be in controlling the pseudo motions in the background area, and δ is Kronecker's delta, that returns 1 when the control points $\boldsymbol{x}^{cp}_{i,j,k}$ are outside the silhouette, namely, background $I(i\Delta x, j\Delta y, k\Delta t) = 0$.

Figure 2(c) shows the result of the smooth pseudo motion extraction. While the local fluctuation are found in the raw pseudo motion (Fig. 2(b)), the smooth pseudo motion is much more consistent in the spatio-temporal domain.

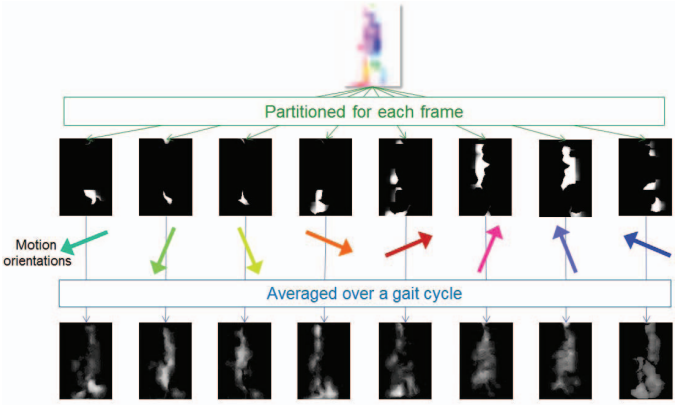


Fig. 3. Image of oriented smooth pseudo motion. Each column indicates each orientation from orientation code 0 to 7.

C. Image of Oriented Smooth Pseudo Motion

Due to the efficacy of feature representation which average over one gait period (e.g., GEI [15] and GFI [18]), the smooth pseudo motion is also averaged over one gait period.

First, the gait period P is detected by maximizing normalized autocorrelation of the silhouette sequence [22]. The whole sequence $I = \{I(\boldsymbol{x}, t)\}(t = 0, \ldots, N-1)$ is divided into subsequences, where i-th subsequence is denoted as $I^{sub}_i = \{I(\boldsymbol{x}, t)\}(t = iP, \ldots, (i+1)P - 1\}$ $(i = 0, \ldots, N_{period} - 1)$, where $N_{period} = \lfloor \frac{N}{P} \rfloor$.

We then consider the averaging the smooth pseudo motion. Because simple averaging of the smooth pseudo motion may cancel each other out (e.g., left and right motions), the smooth pseudo motion is first partitioned into eight different orientations (see middle row of Fig. 3) according to the color code shown in Fig. 2(d). Here, we denote the intensity image of k-th oriented smooth pseudo motion at t-th frame as $F_k(\boldsymbol{x}, t)$. Consequently, an image of oriented smooth pseudo motion (IOSPM) for i-th period and k-th orientation is constructed by averaging the intensity image $F_k(\boldsymbol{x}, t)$ over one gait period as

$$F_{i,k}(\boldsymbol{x}) = \frac{1}{P} \sum_{t=iP}^{(i+1)P-1} F_k(\boldsymbol{x}, t) \quad (12)$$

The extracted IOSPMs are shown at the bottom of Fig. 3. We can see that the intensity for every orientation is well preserved without canceling out each other.

III. Matching

A. Shape Matching

As for the shape cue, we simply use raw silhouette image sequences and adopt a variant of the baseline algorithm [12] for matching the silhouette image sequences. Given gallery silhouette image sequence $I^G(\boldsymbol{x}, t), (t = 0, \ldots, N_G - 1)$ and probe one $I^P(\boldsymbol{x}, t), (t = 0, \ldots, N_P - 1)$, gait period P for the probe silhouette image sequence is detected by maximizing normalized autocorrelation [22]. The probe silhouette image sequence is then divided into subsequences, where i-th subsequence is defined as $I^P_i(\boldsymbol{x}, t), (t = iP, \ldots, (i+1)P - 1)$.

Next, the distance between the gallery silhouette image sequence and i-th probe silhouette image subsequences is calculated by shifting the gallery frame so as to minimize the sum of Euclidean distance as

$$d_i^I = \min_s \sum_{t=0}^{P-1} \sum_{\boldsymbol{x}} |I^G(\boldsymbol{x}, t+s) - I^P(\boldsymbol{x}, t+iP)|^2. \quad (13)$$

The whole distance D^I is finally computed as a median of $\{d_i\}$, namely,

$$D^I = \mathrm{Median}_i d_i^I. \quad (14)$$

B. Motion Matching

As for the motion cue, the IOSPMs described in II-C are utilized. Since the IOSPMs have already averaged over one gait cycle, phase information is degenerated, which means that phase synchronization is unnecessary. Let the IOSPM for i-th period of gallery and j-th period of probe be $\boldsymbol{F}_i^G = \{F_{i,k}^G\}$ and $\boldsymbol{F}_j^P = \{F_{j,k}^P\}$, $(k = 0, \ldots, 7)$, respectively, and the number of periods for the gallery and probe be N_{period}^G and N_{period}^P, respectively. A distance between them is defined as Euclidean distance as

$$d_{i,j,k}^F = \sum_{\boldsymbol{x}} |F_{i,k}^G - F_{j,k}^P|^2. \quad (15)$$

The whole distance D_k^F for k-th orientation is finally computed as the minimum of all the combination $\{d_{i,j}^F\}$, namely,

$$D_k^F = \min_{i,j} d_{i,j,k}^F. \quad (16)$$

IV. Score-level Fusion

For the purpose of better performance, score-level fusion techniques are applied to the distances derived from shape and motion cues. In particular, we adopt linear logistic regression (LLR) [23] as a training-based score-level fusion method in this paper.

Assume that N training samples are given and that i-th training sample is composed of a pair of positive/negative label and nine-dimensional distance vector $\boldsymbol{d}_i = [D^I, D_0^F, \ldots, D_7^F]^T$ derived from both shape and motion cues. Figure 4(a) shows the distributions between shape-based distance D^I and motion-based distance D_k^F. Because the shape and motion cues have different properties, we can see that the distributions are not completely correlated each other.

In addition, non-overlapped test samples are also shown in Fig. 4(b). We can see the similar tendencies between the training and test samples distributions to some extent, which indicates the possibility of performance improvement by score-level fusion.

Hence, the training-based score-level fusion find a kind of the optimal weights $\boldsymbol{w} = [w^I, w_0^F, \ldots, w_7^F]^T$ for each distance component in the distance vector $\boldsymbol{d}$, and this is done by minimizing the logistic loss function of the training sample in the LLR framework in conjunction with Z-normalization [24] for each distance component as preprocess. Then, the learnt optimal weights is applied to the test samples and the single-dimensional fused distance is returned as a weighted linear combination of the distance components as

$$d = w^I D^I + \sum_{k=0}^{7} w_k^F D_k^F. \quad (17)$$

Finally, if the fused distance d is less than an acceptance threshold T, a trial is accepted as a genuine, otherwise rejected as an imposter.

V. Experiments

A. Data Set

The proposed method was tested on the publicly available gait database, that is, the OU-ISIR Gait Database, the Treadmill Dataset D, DB_{high} [25][1].

The DB_{high} consists of 100 subjects with relatively high stability, which means it has higher normalized autocorrelation. Each subject consists of a pair of gallery and probe silhouette image sequence, where each of them has been already size-normalized and registered with regard to silhouette centroids. Sample silhouette image sequences are shown in Fig. 5.

In order to train the optimal weights for score level fusion, 50 subjects are randomly chosen as a training set and the other 50 subjects are employed as a test set. These random trials are repeated five times and averaged performance is evaluated.

B. Results

For evaluating the performance in the verification scenario (one-to-one matching), Receiver Operating Characteristics (ROC) curves is employed. The ROC curve denotes a trade-off curve between False Rejection Rate (FRR) and False Acceptance Rate (FAR) [26] when the operator changes acceptance thresholds. FRR denotes the rate when the system mistakenly recognizes a genuine to be an impostor, while FAR denotes the rate when the system mistakenly recognizes an impostor to be a genuine.

The resultant ROC curves for one of the five random trials are shown both for the training and test set in Figs. 6(a) and (b), respectively, with comparison to the case where the baseline algorithm is used alone. The ROC curves of the training data indicate that the proposed method outperforms the baseline algorithm for all the operating points, while that the proposed method is sometimes comparable to the baseline algorithm for the test set. One of major reasons of performance degradation is so-called generalization errors derived from the distance distribution difference between the training and test sets.

For the further analysis, Error Equal Rates (EERs) of FRR and FAR are confirmed. The statistics of EERs both for the training and test sets for the five random trials are summarized in Table I. As a result, the proposed method shows better performance than the baseline algorithm for the training set, and it shows slightly better performance than the baseline algorithm.

[1]This database is distributed at http://www.am.sanken.osaka-u.ac.jp/GaitDB/index.html

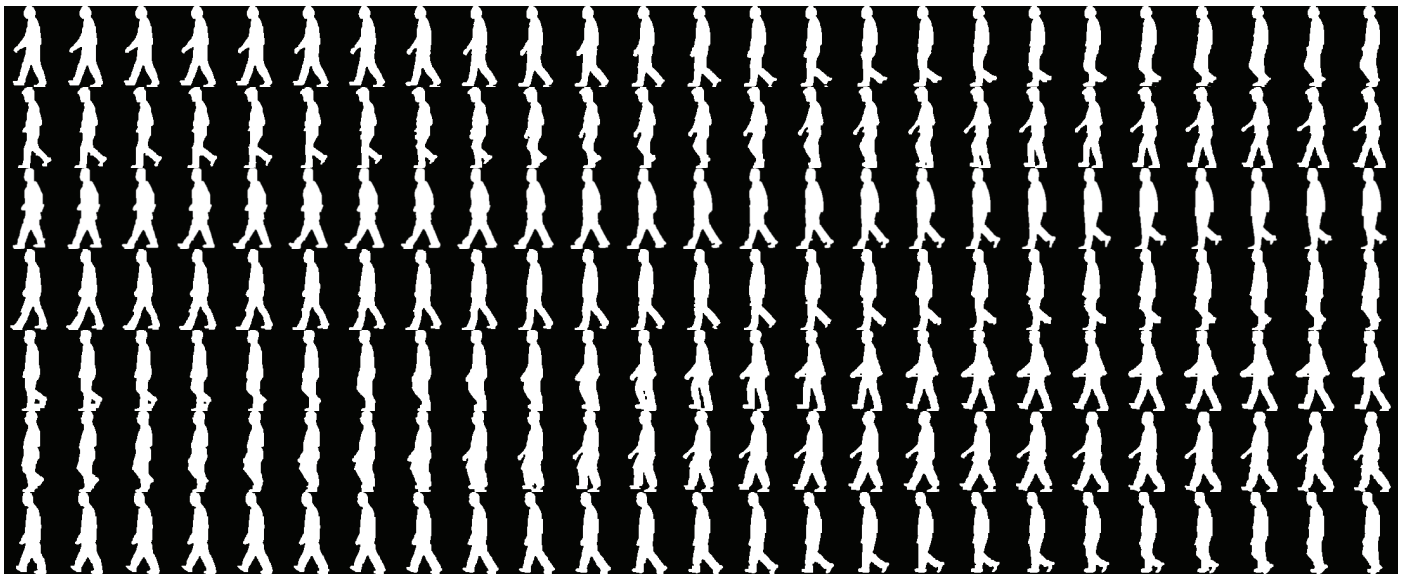


Fig. 5. Sample silhouette image sequences of the Treadmill Dataset D, DB_{high}. Each row shows each different subject.

TABLE I
AVERAGED EERS (%) (BRACKET: STANDARD DEVIATION).

Data set	Baseline algorithm [12]	Proposed method
Training data	4.2 ($\pm$ 0.5)	**3.1** ($\pm$ 1.0)
Test data	4.4 ($\pm$ 0.4)	**4.1** ($\pm$ 1.1)

From these results, although it is an essential future work to cope with generalization errors, the effectiveness of the proposed method is confirmed to some extent.

VI. CONCLUSION

In this paper, a method of gait recognition that uses not only shape feature but also motion feature was proposed.

In order to obtain inner silhouette motion (*pseudo motion*) unlike the previous method only utilized contour motion, a technique for shape morphing, EMM [19] was introduced. The raw pseudo motion, however, tended to be locally fluctuated due to random clustering process in the EMM framework, the spatio-temporal regularization is imposed to obtain the smooth pseudo motion. The smooth pseudo motion image sequences is further partitioned into eight images according to the orientation of the smooth pseudo motion and averaged over every gait cycle to construct images of oriented smooth pseudo motion (IOSPMs).

Finally, motion features of IOSPMs were fed into the score-level fusion framework in conjunction with shape features via the baseline algorithm, to achieve better performance. More specifically, LLR for the nine distance vector was applied to obtain the optimal weights for each distance component and the fused distance was defined as a weighted linear combination of them.

Evaluations were done by using the publicly available gait database showed that the proposed method worked well compared with the case where the baseline algorithm was used alone.

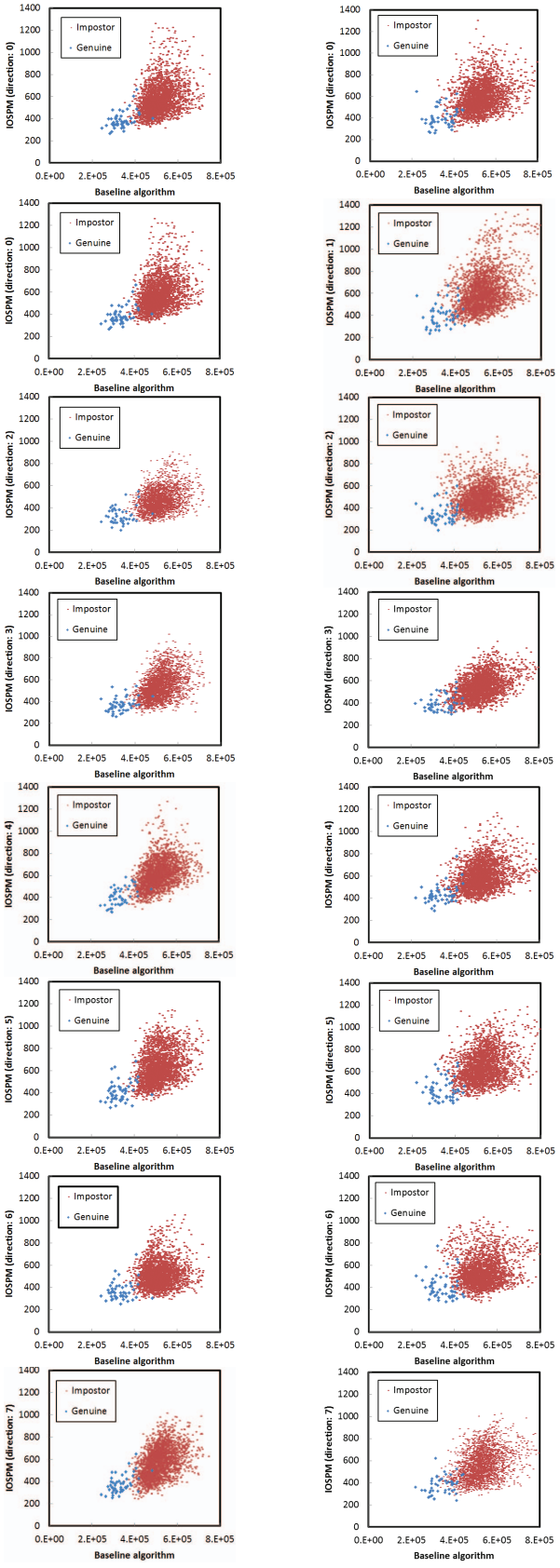
Future work related to this work is to cope with generalization error to improve the performance of the test set.

ACKNOWLEDGMENT

This work was supported by JSPS Grant-in-Aid for Scientific Research(S) Grant Number 21220003.

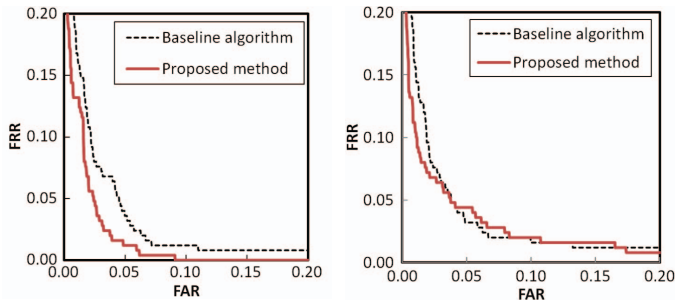
REFERENCES

[1] "How biometrics could change security," http://news.bbc.co.uk/2/hi/programmes/click_online/ 7702065.stm.
[2] A. Bobick and A. Johnson, "Gait recognition using static activity-specific parameters," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 1, 2001, pp. 423–430.
[3] D. Wagg and M. Nixon, "On automated model-based extraction and analysis of gait," in *Proc. of the 6th IEEE Int. Conf. on Automatic Face and Gesture Recognition*, 2004, pp. 11–16.
[4] R. Urtasun and P. Fua, "3d tracking for gait characterization and recognition," in *Proc. of the 6th IEEE Int. Conf. on Automatic Face and Gesture Recognition*, 2004, pp. 17–22.
[5] D. Cunado, M. Nixon, and J. Carter, "Automatic extraction and description of human gait models for recognition purposes," *Computer Vision and Image Understanding*, vol. 90, no. 1, pp. 1–41, 2003.
[6] C. Yam, M. Nixon, and J. Carter, "Automated person recognition by walking and running via model-based approaches," *Pattern Recognition*, vol. 37, no. 5, pp. 1057–1072, 2004.
[7] M. Goffredo, I. Bouchrika, J. N. Carter, and M. S. Nixon, "Performance analysis for automated gait extraction and recognition in multi-camera surveillance," *Multimedia Tools Appl.*, vol. 50, no. 1, pp. 75–94, 2010.
[8] S. Niyogi and E. Adelson, "Analyzing and recognizing walking figures in xyt," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, 1994, pp. 469–474.
[9] Y. Ohara, R. Sagawa, T. Echigo, and Y. Yagi, "Gait volume: Spatio-temporal analysis of walking," in *Proc. of the 5th Workshop on Omni-directional Vision, Camera Networks and Non-classical cameras*, 2004, pp. 79–90.
[10] C. BenAbdelkader, R. Culter, H. Nanda, and L. Davis, "Eigengait: Motion-based recognition people using image self-similarity," in *Proc. of Int. Conf. on Audio and Video-based Person Authentication*, 2001, pp. 284–294.



(a) Training data (b) Test data

Fig. 4. Score distribution between baseline algorithm and individual images of oriented smooth pseudo motion from top to bottom.



(a) Training data (b) Test data

Fig. 6. ROC curves

[11] J. Little and J. Boyd, "Recognizing people by their gait: The shape of motion," *Videre: Journal of Computer Vision Research*, vol. 1, no. 2, pp. 1–13, 1998.

[12] S. Sarkar, J. Phillips, Z. Liu, I. Vega, P. G. ther, and K. Bowyer, "The humanid gait challenge problem: Data sets, performance, and analysis," *Trans. of Pattern Analysis and Machine Intelligence*, vol. 27, no. 2, pp. 162–177, 2005.

[13] Z. Liu and S. Sarkar, "Simplest representation yet for gait recognition: Averaged silhouette," in *Proc. of the 17th Int. Conf. on Pattern Recognition*, vol. 1, Aug. 2004, pp. 211–214.

[14] G. Veres, L. Gordon, J. Carter, and M. Nixon, "What image information is important in silhouette-based gait recognition?" in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 2, 2004, pp. 776–782.

[15] J. Han and B. Bhanu, "Individual recognition using gait energy image," *Trans. on Pattern Analysis and Machine Intelligence*, vol. 28, no. 2, pp. 316– 322, 2006.

[16] K. Bashira, T. Xiang, and S. Gong, "Gait recognition without subject cooperation," *Pattern Recognition Letters*, vol. 31, no. 13, pp. 2052–2060, Jun. 2010.

[17] C. Wang, J. Zhang, J. Pu, X. Yuan, and L. Wang, "Chrono-gait image: A novel temporal template for gait recognition," in *Proc. of the 11th European Conf. on Computer Vision*, Heraklion, Crete, Greece, 2010, pp. 257–270.

[18] 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*, vol. 44, pp. 973–987, April 2011. [Online]. Available: http://dx.doi.org/10.1016/j.patcog.2010.10.011

[19] Y. Makihara and Y. Yagi, "Earth mover's morphing: Topology-free shape morphing using cluster-based emd flows," in *Proc. of the 10th Asian Conf. on Computer Vision*, Queenstown, New Zealand, Nov. 2010, pp. 2302–2315.

[20] N. Poh and S. Bengio, "Database, protocols and tools for evaluating score-level fusion algorithms in biometric authentication," *Pattern Recognition*, vol. 39, pp. 223–233, 2006.

[21] F. Hoppner, F. Klawonn, R. Kruse, and T. Runkler, *Fuzzy Cluster Analysis*. John Wiley and Sons, 1999.

[22] 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 Conf. on Computer Vision*, Graz, Austria, May 2006, pp. 151–163.

[23] F. Alonso-Fernandez, J. Fierrez, D. Ramos, and J. Ortega-Garcia, "Dealing with sensor interoperability in multi-biometrics: the upm experience at the biosecure multimodal evaluation 2007," in *Proc. of SPIE 6994, Biometric Technologies for Human Identification IV*, Orlando, FL, USA, Mar. 2008.

[24] R. Auckenthaler, M. Carey, and H. Lloyd-Thomas, "Score normalization for text-independant speaker verification systems," *Digital Signal Processing*, vol. 10, no. 1-3, pp. 42–54, 2000.

[25] 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 Trans. on Computer Vision and Applications*, vol. 4, pp. 53–62, Apr. 2012.

[26] N. Poh and S. Bengio, "Eer of fixed and trainable fusion classifiers: A theoretical study with application to biometric authentication tasks," in *Multiple Classier Systems*, 2005, pp. 74–85.