

論文 / 著書情報
Article / Book Information

Title	Robust Gait Recognition against Speed Variation
Author	Muhammad Rasyid Aqmar, Koichi Shinoda, Sadaoki Furui
Journal/Book name	Proc. ICPR2010, Vol. , No. , pp. 2190-2193
発行日 / Issue date	2010, 8
権利情報 / Copyright	(c)2010 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Robust Gait Recognition against Speed Variation

Muhammad Rasyid Aqmar, Koichi Shinoda, and Sadaoki Furui

Department of Computer Science, Tokyo Institute of Technology, Tokyo 152-8552 Japan

E-mail: rasyid@ks.cs.titech.ac.jp, {shinoda, furui}@cs.titech.ac.jp

Abstract—Variations in walking speed have a strong impact on the recognition of gait. We propose a method of recognition of gait that is robust against walking-speed variations. It is established on a combination of Fisher discriminant analysis (FDA)-based cubic higher-order local auto-correlation (CHLAC) and the statistical framework provided by hidden Markov models (HMMs). The HMMs in this method identify the phase of each gait even when walking speed changes nonlinearly, and the CHLAC features capture the within-phase spatio-temporal characteristics of each individual. We compared the performance of our method with other conventional methods in our evaluation using three different databases, i.e., USH, USF-NIST, and TokyoTech DB. Ours was equal or better than the others when the speed did not change too much, and was significantly better when the speed varied across and within a gait sequence.

I. INTRODUCTION

People can walk at various speeds in any real-life situations, and human motion changes *nonlinearly* according to its speed (e.g., [1]). Variations can appear across and within a gait sequence that will significantly affect the performance of human-gait-based ID recognition systems. Therefore, many studies have been conducted to build a gait-recognition system that is robust against variations in walking speed.

For example, Lee *et al.* [6] proposed a shape-based frieze pattern and evaluated it using CMU MoBo database [3], which consisted of gait data at different speeds. Kobayashi *et al.* [5] proposed a three-way (x -, y -, and time-axis) method of autocorrelation for extracting features that effectively extracted spatio-temporal local geometric features to characterize motions. It was called cubic higher-order local auto-correlation (CHLAC). It was expected to be relatively robust against variations in walking speed, since it only used the sums of local features over a gait sequence, and thus did not explicitly use the phase information of the gait. To the best of our knowledge, this method performed the best among all gait-recognition methods. However, in most of these methods including CHLAC, researchers have assumed that walking speed does not change much within or across gait sequences. Their performance may degrade greatly when walking speed varies significantly because of misaligned gait cycles and/or phases.

A straightforward way of tackling this misalignment problem is to estimate the nonlinear time-warping function

between two time-sequence patterns with different speeds. Veeraraghavan *et al.* proposed an approach based on dynamic time warping (DTW), which used a set of DTW functions to represent the distribution of gait patterns using uniform and wrapped-Gaussian distributions [10] [11]. Kale *et al.* [4] used a hidden Markov model (HMM) for this purpose. HMM, which is a natural extension of DTW to a framework based on a probability theory, implicitly carries out time-alignment between different patterns.

In this paper, we propose a novel method of gait recognition, where we modify the extraction process of CHLAC to model each gait phase, and combine it with a statistical HMM framework. It uses speed-invariant features and can adapt to speed variations across and within a gait sequence. We expect that this combination can perform better than using any of them individually.

This paper is organized as follows. Section 2 reviews CHLAC-based features, and Section 3 explains our combined method of CHLAC features and HMMs. Section 4 reports the experimental results. Section 5 concludes the paper.

II. CHLAC FEATURES

CHLAC extracts shape and motion information based on local autocorrelation [5]. One of the most important properties of CHLAC features is their shift invariance, i.e., CHLAC features do not change if the position of a person varies inside a frame image. Let $f(x, y, t)$ represent spatio-temporal data (pixel intensity) depending on the image region, where x and y are pixel coordinates in one frame image, and t is the time index. Each of the N -th order autocorrelation functions is defined as:

$$\begin{aligned} R_N(\mathbf{a}_1, \dots, \mathbf{a}_N) \\ = \sum_{x, y, t \in D_s} f(x, y, t) f(x + a_{1x}, y + a_{1y}, t + a_{1t}) \\ \dots f(x + a_{Nx}, y + a_{Ny}, t + a_{Nt}), \end{aligned} \quad (1)$$

where \mathbf{a}_i ($i = 1, \dots, N$) is a displacement vector from the reference point, $\mathbf{r} = (x, y, t)$. The set $(\mathbf{r}, \mathbf{r} + \mathbf{a}_1, \dots, \mathbf{r} + \mathbf{a}_N)$ represents a local mask pattern, and its examples are shown in Fig. 1. D_s is a spatio-temporal region that sums up the correlation for each pixel. The size of D_s is $m \times n \times T$, where

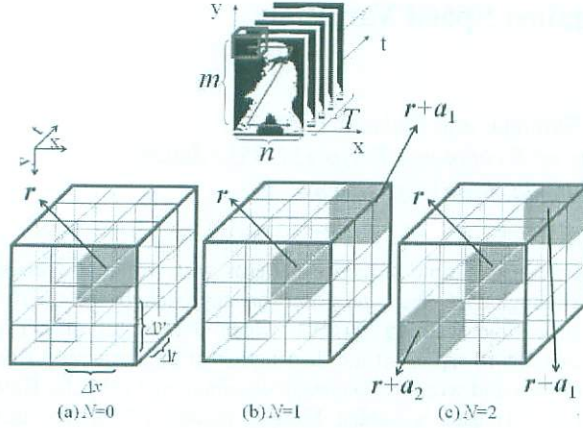


Figure 1. Extraction of CHLAC features.

T is the window width to be optimized in the experiments. For $N = 2$, the components of a_i for each i are limited to $\pm\Delta x$, $\pm\Delta y$, $\pm\Delta t$ or 0 corresponding to a_{ix} , a_{iy} , and a_{it} , where Δx and Δy denote the spatial displacement in pixels and Δt denotes the frame interval in frames. Here, we use the same value for Δx and Δy and denote this as Δr . When the order of correlation is $N = 0$, $N = 1$, and $N = 2$, the numbers of mask patterns (the dimensions of a CHLAC feature vector) correspond to 1, 14, and 251.

Then, the CHLAC features are mapped to the $(c - 1)$ -dimensional (c is the total number of classes) feature vector using Fisher discriminant analysis (FDA) to better separate classes in the feature space.

A k -nearest neighbor (k -NN) classifier was used in the original framework [5], where the Euclidean distance was used as the distance measure between training and testing feature vectors. The number of neighbors k was set to 30. The time window width T was set to 30, and the pair of parameters $(\Delta r, \Delta t)$ were set to be (2,1), (4,2), (6,3), (8,4), (10,5), (12,6), (14,7), and (16,8). In each frame, the CHLAC features for each parameter pair were extracted and mapped to the FDA space. Next, the number of nearest neighbors belonging to each ID i , $M_{(\Delta r, \Delta t)}(i)$, was counted for each of eight parameter pairs of Δr and Δt . Then, the ID i with the maximum number of neighbors over all parameter pairs was selected:

$$\hat{i} = \arg \max_i \max_{\Delta r, \Delta t} M_{(\Delta r, \Delta t)}(i). \quad (2)$$

The same process was repeated for all the frames. Finally, the ID i that most frequently appeared over all the frames was selected.

Since the parameters related to walking speed are automatically selected during the recognition process, CHLAC features are relatively robust against variations in speed

across samples when changes in speed are almost constant in a complete gait sequence. When walking speed varies within one gait sequence, however, the selected parameters are not optimal for some cycles and/or phases. This may degrade recognition.

III. COMBINATION OF CHLAC WITH HMMs

CHLAC discriminates accurately between classes, while HMMs have excellent properties to match sequences that have different speeds. By combining them, we expect to obtain a more robust method against speed variations, even when the speed varies within a gait sequence. We employ a *continuous* gait-recognition framework where gait cycles are segmented simultaneously with the recognition process.

Since CHLAC features do not have much gait phase/cycle information, it is difficult to train HMMs solely with CHLAC features. We not only extract CHLAC features, but also principal component analysis (PCA) features. We use PCA features to segment phases and cycles in each gait sequence, and use the segment information obtained to train the state output probability for CHLAC features.

Kobayashi *et al.* [5] set the window length, T , for CHLAC at 30 (frames), which roughly corresponded to the duration of a complete gait cycle. In our approach using HMMs, this window length should be modified to capture features at a certain phase in a gait cycle. We set T at five (frames) according to the results from our preliminary experiments.

We use a half-gait cycle as a unit to train the HMM because we assume symmetrical similarity between the first and the second half of the cycle. HMM topology is left-to-right without any skips. We use a mixture of Gaussian distributions as an output probability. We use an eight-state HMM that gave the best performance among different numbers of states.

IV. EXPERIMENTS

A. Experimental Conditions

In our experiments, we first compared our method with other conventional approaches under a condition where walking speed did not change too much. Second, we evaluated it under the condition where the walking speed was changed.

For the first evaluation, we used the University of Southampton's (USH) large database [8] (115 subjects) and the University of South Florida's (USF)-NIST database Probe A [7] (71 subjects). Variations in both testing sets was only in the camera-views, and other conditions including walking speed were the same.

For the second evaluation, we constructed a TokyoTech database containing 30 subjects walking at various fixed speeds. A treadmill was used to ensure that all subjects walked at exactly the same speed for all speed categories. The gait data were categorized into four types: slow, normal, fast, and mixed (Table I). We divided the slow data into two

Table I
TOKYO TECH DATABASE.

Speed type	Slow	Normal	Fast	Mixed
Speed (km/hr)	2	3	4.5	3 & 4.5
No. of samples	605	550	447	300

parts, 60% for training (363 samples) and 40% for testing (242 samples). The training set only consisted of slow data. The testing set consisted of the rest of the slow data and data with the other three speeds. In each evaluation, one testing sample was a gait sequence that contained five gait cycles from one subject. To evaluate the effect of variation within one gait cycle, we created mixed data, where each test sample was constructed from the concatenation of three gait cycles and two gait cycles from the normal and fast data respectively (150 samples), and also from the concatenation of three gait cycles and two gait cycles from the fast and normal data respectively (150 samples).

We assumed that a background image was available in the preprocessing stage. After setting a certain threshold for the intensity of each pixel, the background and foreground pixels were classified to produce a silhouette image. Then, the bounding box around the silhouette was resized into a fixed $m \times n$ pixels. In this study, the size was set to $m = 128$ and $n = 88$ following the NIST standard.

We used the 0th to 2nd order CHLAC features and applied FDA. In our method, CHLAC-HMM, Δr was 12 and Δt was 6 for the USH and USF-NIST databases, and these correspond to 4 and 2 for the TokyoTech database.

We used 114 CHLAC-FDA feature dimensions in our experiments for USH, 70 for USF-NIST, and 29 for the TokyoTech database. These dimensions were automatically determined by the number of subjects to be classified. We also used 60-dimensional PCA features, to compare them with the CHLAC features.

We used a single Gaussian distribution in each state of the HMMs to evaluate the USH and USF-NIST databases due to the limited number of training samples. We used a Gaussian mixture distribution with 16 mixtures in the TokyoTech database, which was relatively large. We used the HMM-Toolkit (HTK) [12] to produce the gait-sequence HMMs.

The segmentation results in a sequence were compared to the manual transcriptions. The frame mis-alignments from all databases before and after using the segment information from PCA features to train the HMMs on average correspond to 12.6 and 10.3 frames. We confirmed that the segmentation was better when using the segment information from PCA features.

B. Results

Table II lists the results of our experiment using the USH and USF-NIST databases. For USH database, we have

Table II
GAIT RECOGNITION ACCURACY (%) FOR USH AND USF-NIST PROBE A DATABASE.

	USH	USF-NIST
Foster <i>et al.</i> [2]	75.0	—
Tolliver <i>et al.</i> [9]	—	85.0
Kale <i>et al.</i> [4]	—	99.0
CHLAC- k -NN	98.3	100.0
CHLAC-HMM (Proposed method)	98.3	100.0

presented the results of the area-based mask pattern- k -NN proposed by Foster *et al.* [2] and CHLAC- k -NN [5] for comparison. The proposed CHLAC-HMM was better than the area-based mask pattern- k -NN and equal to CHLAC- k -NN. For USF-NIST database, we compared our method with the other three methods, i.e., Shape-1-NN [9], Silhouette frame-to-exemplar distance (FED)-HMM [4], and CHLAC- k -NN [5]. The results for these methods were taken from those published in their corresponding papers. We found that our method was significantly better than Shape-1-NN and almost equal to (FED)-HMM and CHLAC- k -NN.

Next, we evaluated the robustness of our proposed approach against speed differences using the TokyoTech database. The recognition results are listed in Figure 2. The results reveal that CHLAC-HMM yielded better results than the other three methods including CHLAC- k -NN. When walking speed was “mixed”, 96.7% accuracy was achieved while the accuracy obtained by the CHLAC- k -NN was 92.0%. The robustness of the proposed approach against walking-speed variations across and within sequences has been confirmed.

V. CONCLUSION

We proposed a method of robust gait recognition against speed variations, based on the combination of CHLAC and HMM. By using USH, USF-NIST, and TokyoTech databases, we confirmed that the approach we propose performed well across different rates of speed across and within sequences. In particular, it successfully reduce the errors more than half from CHLAC- k -NN method when the walking speed varied within a gait sequence.

In future work, we plan to apply an adaptation scheme to the HMM-based framework, to further improve recognition.

ACKNOWLEDGMENTS

This work was supported by a Grant-in-Aid for Scientific Research (B) 20300063. We also thank Professor Otsu and Dr. Kobayashi of AIST, Japan, for their helpful support regarding the use of the USF-NIST gait database.

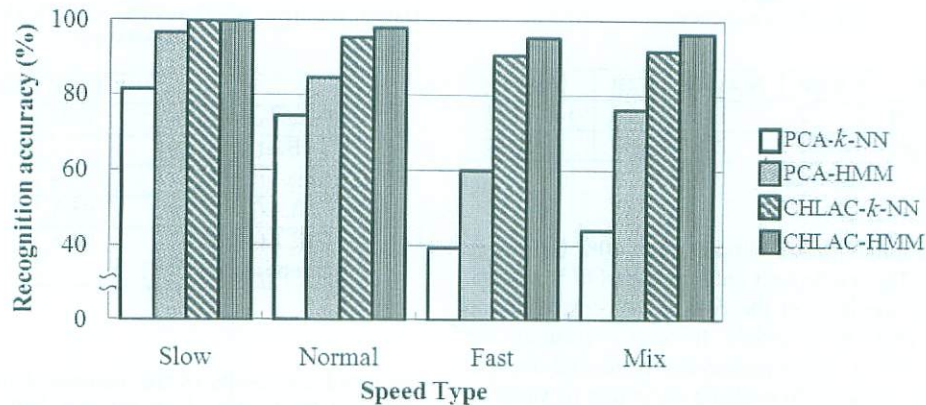


Figure 2. Gait recognition accuracy (%) for TokyoTech database.

REFERENCES

- [1] A. Bobick and R. Tanawongsuwan. Performance analysis of time-distance gait parameters under different speeds. *4th Int. Conf. on AVBPA*, June 2003.
- [2] Jeff P. Foster, Mark S. Nixon, and Adam Prugel-Bennett. Automatic gait recognition using area-based metrics. *Pattern Recognition Letters*, 24(14):2489–2497, 2003.
- [3] R. Gross and J. Shi. The CMU Motion of Body (mobo) database. *Technical Report CMU-RI-TR-01-18*, Robotics Inst., Carnegie Mellon Univ., 2001.
- [4] A. Kale, A. Sundaresan, A. Rajagopalan, A.N. Cuntoor, N.P. Roy-Chowdhury, A.K. Kruger, and R. V. Chellappa. Identification of humans using gait. *IEEE Transactions on Image Processing*, 13(9):1163–1173, 2004.
- [5] Takumi Kobayashi and Nobuyuki Otsu. Three-way auto-correlation approach to motion recognition. *Pattern Recognition Letters*, 30:212–221, 2009.
- [6] Seungkyu Lee, Yanxi Liu, and R. Collins. Shape variation-based frieze pattern for robust gait recognition. *Proc. IEEE CVPR*, pages 1–8, 2007.
- [7] P. J. Phillips, S. Sarkar, I. Robledo, P. Grother, and K. W. Bowyer. The gait identification challenge problem: data sets and baseline algorithm. *Proc. ICPR*, pages 385–388, 2002.
- [8] J. D. Shutler, M. G. Grant, M. S. Nixon, and J. N. Carter. On a large sequence-based human gait database. *Proc. 4th International Conf. on Recent Advances in Soft Computing*, 2002.
- [9] David Tolliver and Robert T. Collins. Gait shape estimation for identification. *Proc. AVBPA*, pages 734–742, 2003.
- [10] A. Veeraraghavan, R. Chellapa, and A.K. Roy-Chowdury. The function space of an activity. *Proc. IEEE CVPR*, pages 959–968, 2006.
- [11] A. Veeraraghavan, R. Chellapa, and A.K. Roy-Chowdury. Rate-invariant recognition of humans and their activities. *IEEE Trans. on Image Processing*, vol. 18, pages 1326–1339, 2009.
- [12] S. Young. *The HTK book (for HTK version 3.2)*. Entropic Ltd., 2002.