Clothes-Invariant Gait Identification using Part-based Adaptive Weight Control

Md. Altab Hossain, Yasushi Makihara, Wang Junqui, Yasushi Yagi
Osaka University
{*hossain, makihara, wang,yagi*}@*am.sanken.osaka-u.ac.jp*

Abstract

This paper describes a method of part-based gait identification under substantial clothes variations. When clothes types between a gallery and a probe are different, silhouettes fairly change for some parts and subject discrimination capability decrease for those parts. Therefore, we exploit the discrimination capability as a matching weight for each part and control the weights adaptively based on a distribution of distances between a probe and all the galleries. As a result of experiments with our clothes-variation gait dataset, the proposed method achieved much better performance than a whole-based approach.

1. Introduction

Gait has recently gained considerable attention as a promising biometric cue due to the possibility of person authentication at a distance from a camera without subject's perception. However, appearance changes due to clothes variations including coats, down jackets, baggy pants, and skirts make the identification problem much difficult.

Working from appearance-based approaches, LDA-based method [3] reduces effects of within-class variations, namely, clothes variations on gait identification to some extent, it does, however, not work well when clothes variations overwhelm individual variations. Lee et al. [4] proposed a shape variation-based frieze pattern as a gait feature for robust recognition. Because the feature is obtained by projecting the 2-D silhouette into 1-D horizontal or vertical axis, meaningful 2-D position information is degraded in the feature.

On the other hand, we have an observation that different clothes combinations make impacts on different body parts in general. For example, hats, down jackets, and skirts have impacts on heads, upper body, and lower body respectively and the other parts can be still used for recognition efficiently. Boulgouris [1] proposed a component-based gait recognition considering unequal discrimination capability of each part; the method gives fixed weights of the parts obtained in a training phase.

Therefore, we propose a part-based adaptive weight control for clothes-invariant gait identification. In a training phase, distributions of distances between gait features are stored in both the same and different clothes cases for each part separately, and the corresponding discrimination capabilities are also measured in advance. In a test phase, given a probe, probabilities of the same and different clothes are calculated based on distances between the probe and galleries and on the trained distance distributions.Then a matching weight for each part is adaptively calculated as a probability-weighted discrimination capability.

2. Clothes-variation gait dataset

Previous gait dataset are short in subjects or clothes variations. Soton small database [7] have some clothes variations such as trainer, rain coat, and trench coat; however the number of subjects are only 12. On the other hand, Soton large database [7], HumanID dataset [6], and CASIA dataset [8] have more than 100 subjects, but the clothes variations are limited to casual wear and long coat. Therefore, we constructed our own large-scale clothes-variation gait dataset. It includes 52 subjects with at most 32 combinations of clothes types. Table 1. lists the clothes types and Tab. 2. shows the combination of clothes used in constructing the dataset. All those gait sequences were captured twice on the same day in an indoor environment. Thus, the total number of sequences in the dataset is 2120. Figure 1. shows sample images of clothes types 9, X, 6, P, V, and R from the dataset of a subject. The large number of subjects and clothes-variations of the new dataset allow us for an estimate of intra-subject variations together with inter-subjects variations for better assessment of the potential of gait.

3. Feature extraction

3.1. Frequency-domain feature

In this section, frequency-domain feature extraction is briefly addressed (see [5] for detail). Given an image

978-1-4244-2175-6/08/$25.00 ©2008 IEEE

Table 1. Clothes list in the dataset (Abbreviation: name)

NP: Normal Pants	HS: Half Shirt	CW: Casual Wear
BP: Baggy Pants	FS: Full Shirt	RC: Rain Coat
SP: Short Pants	LC: Long Coat	Ht: Hat
Sk: Skirt	Pk: Parkar	Cs: Casquette
CP: Casual Pants	DJ: Down Jacket	Mf: Muffler

Table 2. Clothes combination types

#	s_1	s_2	s_3	#	s_1	s_2	#	s_1	s_2
2	NP	HS	-	A	NP	Pk	T	Sk	FS
3	NP	HS	Ht	B	NP	DJ	U	Sk	Pk
4	NP	HS	Cs	I	BP	HS	V	Sk	DJ
9	NP	FS	-	K	BP	FS	D	CP	HS
X	NP	FS	Ht	J	BP	LC	F	CP	FS
Y	NP	FS	Cs	L	BP	Pk	E	CP	LC
5	NP	LC	-	M	BP	DJ	G	CP	Pk
6	NP	LC	Mf	N	SP	HS	H	CP	DJ
7	NP	LC	Ht	Z	SP	FS	0	CP	CW
8	NP	LC	Cs	P	SP	Pk	R	RC	RC
C	NP	DJ	Mf	S	Sk	HS			

#: Clothes combination type, s_i: ith clothes slot.

sequence, extracted silhouettes are size-normalized and registered to acquire a spatio-temporal Gait Silhouette Volume (GSV). Then, a gait period N_{gait} is detected by autocorrelation of the GSV, and amplitude spectra at each pixel is calculated by Discrete Fourier Transformation (DFT) based on the gait period N_{gait} as shown in Fig. 2. In this paper, the image size of the GSV is 128×88, and from 0- to 2-times frequencies are used. Thus, a gait feature vector composed of all the amplitude spectra becomes $128 \times 88 \times 3 = 33,792$ dimensional vector.

3.2. Part-based feature

Based on the known anatomical body properties [2], human body (height: H) is divided by the vertical positions of the neck ($0.87H$), waist ($0.535H$), pelvis (0.48), and knee (0.285) and 8 parts are defined as shown in Fig. 3. All parts are trained separately for the PCA subspace using the frequency-domain features from the corresponding regions, and dimension-reduced features are used as part-based gait features.

4. Part-based adaptive weight control

4.1. Overview

In a part-based method, the important issue is how to combine the individual distances into a single distance that quantify the overall similarity/dissimilarity between a probe and a gallery. We adopt the weighted

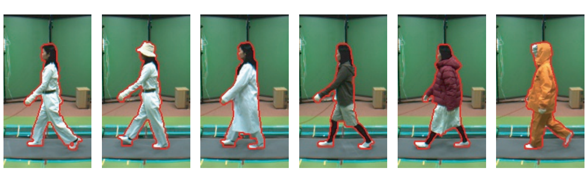


Figure 1. Sample clothes images

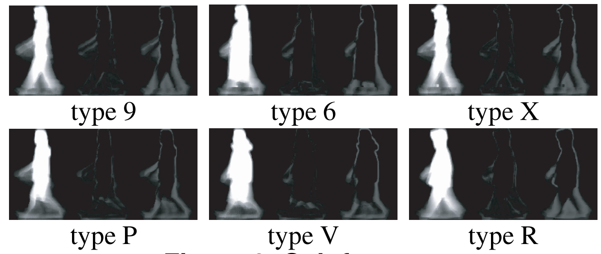


Figure 2. Gait features

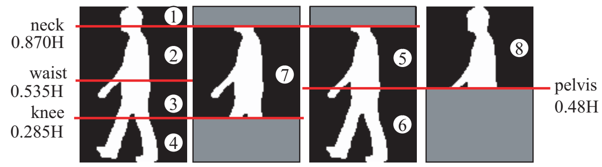


Figure 3. Definition of body parts

sum of distances corresponding to each part for matching measure here.

Let $\boldsymbol{x}^p$ and $\boldsymbol{x}^g$ are sequences of a probe and a gallery respectively. The matching measure between them is

$$D(\boldsymbol{x}^p, \boldsymbol{x}^g; \boldsymbol{X}^G) = \sum_{i=1}^{N_{parts}} w_i(\boldsymbol{x}^p; \boldsymbol{X}^G) d_i^z(\boldsymbol{x}^p, \boldsymbol{x}^g; \boldsymbol{X}^G), \qquad (1)$$

where N_{parts} is the number of the body parts, $d_i^z(\boldsymbol{x}^p, \boldsymbol{x}^g; \boldsymbol{X}^G)$ is a z-normalized distance between $\boldsymbol{x}^p$ and $\boldsymbol{x}^G$ for ith body part, and $\boldsymbol{X}^G$ is a set of sequences of all the galleries used for z-normalization and for the following probabilities calculation, and $w_i(\boldsymbol{x}^p; \boldsymbol{X}^G)$ is weight for ith body part. The weight $w_i(\boldsymbol{x}^p; \boldsymbol{X}^G)$ is adaptively controlled by the following equation:

$$w_i(\boldsymbol{x}^p; \boldsymbol{X}^G) = P_i^{Sc}(\boldsymbol{x}^p; \boldsymbol{X}^G) F_i^{Sc} + P_i^{Dc}(\boldsymbol{x}^p; \boldsymbol{X}^G) F_i^{Dc}, \qquad (2)$$

where F_i^{Sc} and $P_i^{Sc}(\boldsymbol{x}^p; \boldsymbol{X}^G)$ are discrimination capability and probability in case where the probe clothes is included in the gallery clothes category (call the case Sc), F_i^{Dc} and $P_i^{Dc}(\boldsymbol{x}^p; \boldsymbol{X}^G)$ are those in the other case (call the case Dc). In the following sections, each component is formulated.

4.2. Part-based matching measure

In this section, part-based matching measure is defined. Let $\{x_j^p\}(j = 1, 2, \ldots)$ and $\{x_k^g\}(k = 1, 2, \ldots)$ be subsequences with N_{gait} frames for a probe $\boldsymbol{x}^p$ and a gallery $\boldsymbol{x}^g$ respectively. Let $\boldsymbol{a}_i(x)$ be dimension-reduced feature vector for ith body part for a subsequence x. The matching measure for the subsequences (let it be $d_i^{sub}(x_j^p, x_k^g)$) is simply chosen as the Euclidean distance between feature vectors $\boldsymbol{a}_i(x_j^p)$ and $\boldsymbol{a}_i(x_k^g)$. Then, matching measure between complete sequences for ith body part is defined as mean value of the minimum distances of each probe subsequences and gallery subsequences:

$$d_i(\boldsymbol{x}^p, \boldsymbol{x}^g) = \text{Average}_j[\min_k\{d_i^{sub}(x_j^p, x_k^g)\}] \qquad (3)$$

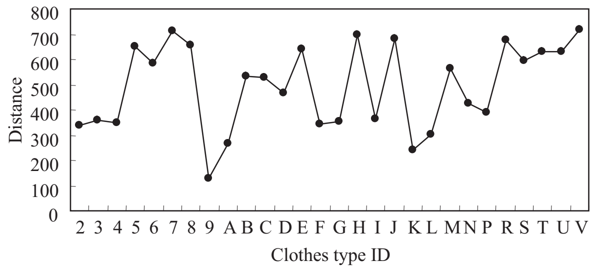


Figure 4. Distance between clothes type 9 and each clothes

Finally, the z-normalized distance is defined as

$$d_i^z(\boldsymbol{x}^p, \boldsymbol{x}^g; \boldsymbol{X}^G) = \frac{d_i(\boldsymbol{x}^p, \boldsymbol{x}^g) - \mu_i(\boldsymbol{x}^p; \boldsymbol{X}^G)}{\sigma_i(\boldsymbol{x}^p; \boldsymbol{X}^G)}, \quad (4)$$

where $\mu_i(\boldsymbol{x}^p; \boldsymbol{X}^G)$ and $\sigma_i(\boldsymbol{x}^p; \boldsymbol{X}^G)$ are mean and standard deviation of distances between a probe $\boldsymbol{x}^p$ and a set of all the galleries $\boldsymbol{X}^G$ for ith body part respecively.

4.3. Clothes categorization for each part

There are some similar/different clothes types for each part. For example, clothes type 9 (NP + FS) and type S (Sk + FS) are the same for body part 2 but different for body part 4, and vice versa in case of clothes type 9 (NP + FS) and type B (NP + DW). Thus, in order to define "the same clothes" and "different clothes" for each specific body part, we make clothes categorization. The distance distribution from standard clothes (clothes type 9) to all the other clothes is used for categorization. For example, Fig. 4. shows the distribution of the third body part. We divide the clothes into five categories: (i) 5, 6, 7, 8, J, and E, (ii) B, M, and H, (iii) S, T, U, and V, (iv) R, and (v) the others based on the distribution for this part.

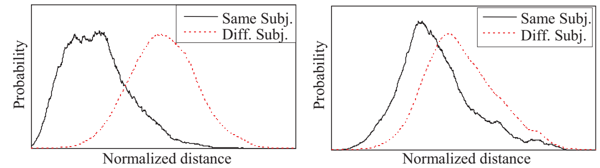
4.4. Discrimination capability

Suppose that a training set composed of a set of probes and galleries is given and that it is divided into 4 sets for ith body part: $U_{Ss,i}^{Sc}$ (the same subjects with the same clothes), $U_{Ds,i}^{Sc}$ (different subjects with the same clothes), $U_{Ss,i}^{Dc}$ (the same subjects with different clothes), and $U_{Ds,i}^{Dc}$ (different subjects with different clothes).

Then, a distribution of z-normalized distances between all pairs of a probe and a gallery in each set is calculated, and Fisher's discriminant ratio between the distributions for the same and different subjects is exploited as discrimination capability. For example, Fisher's discriminant ration for the same clothes F_i^{Sc} is defined as

$$F_i^{Sc} = \frac{(\sigma_{B,i}^{Sc})^2}{(\sigma_{T,i}^{Sc})^2} \quad (5)$$

where $(\sigma_{B,i}^{Sc})^2$ and $(\sigma_{T,i}^{Sc})^2$ are between-class and total variances for the distributions for $U_{Ss,i}^{Sc}$ and $U_{Ds,i}^{Sc}$ respectively. Similarly, Fisher's discriminant ratio for different clothes F_i^{Dc} is calculated.



(a) The same clothes (b) Different clothes

Figure 5. Distance distribution between the same and different subjects

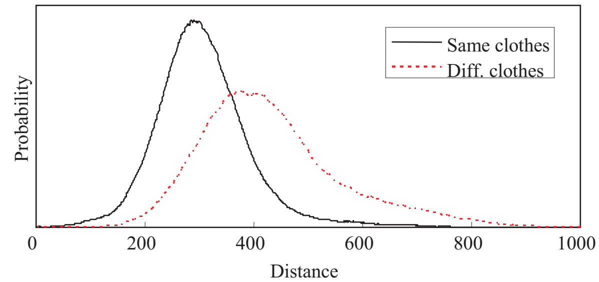


Figure 6. Distance distribution between the same and different clothes

Figure 5. shows, for an example of the third body part, distributions of z-normalized distances for each set. We can see the discrimination capability for the same clothes is superior to that for different clothes.

4.5. Probabilities of the same and different clothes

Generally, because the clothes variation overwhelms the individual variation, distance distributions for the same and different clothes have different properties. First, the distance distributions for the same and different clothes are calculated using training sets $U_i^{Sc}(= U_{Ss,i}^{Sc} \cup U_{Ds,i}^{Sc})$ and $U_i^{Dc}(= U_{Ss,i}^{Dc} \cup U_{Ds,i}^{Dc})$ respectively, and probability distribution functions (PDFs) are obtained by normalizing with the number of pairs in each set (let them be $P_i(d_i|Sc)$ and $P_i(d_i|Dc)$ respectively). For example, PDFs for the third body part are shown shown in Fig. 6. and the difference between them can be clearly observed.

Thus, based on the distribution difference, probabilities for the same and different clothes for a given probe are formulated. Let $\boldsymbol{d}_i(\boldsymbol{x}^p; \boldsymbol{X}^G)$ is a set of distances between $\boldsymbol{x}^p$ and galleries in $\boldsymbol{X}^G$ for ith body part. A posterior for the same clothes is written by Bayesian rule:

$$P_i(Sc|\boldsymbol{d}_i(\boldsymbol{x}^p;\boldsymbol{X}^G)) = \frac{P_i(\boldsymbol{d}_i(\boldsymbol{x}^p; \boldsymbol{X}^G)|Sc)P_i(Sc)}{\sum_{c \in \{Sc,Dc\}} P_i(\boldsymbol{d}_i(\boldsymbol{x}^p;\boldsymbol{X}^G)|c)P_i(c)} \quad (6)$$

$$P_i(\boldsymbol{d}_i(\boldsymbol{x}^p; \boldsymbol{X}^G)|c) = \prod_{\boldsymbol{x}^g \in \boldsymbol{X}^G} P_i(d_i(\boldsymbol{x}^p, \boldsymbol{x}^g)|c), \quad (7)$$

where $P_i(Sc)$ and $P_i(Dc)$ are priors for the same and different clothes which are obtained by ratios of the numbers of pairs in the set U_i^{Sc} and U_i^{Dc} respectively.

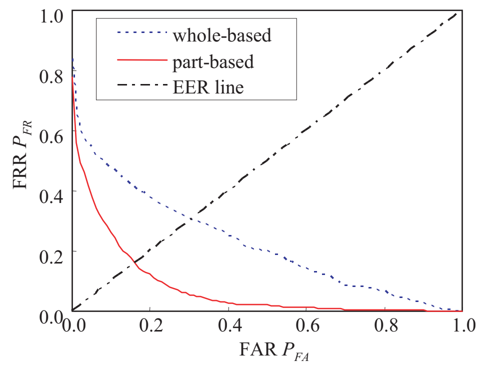


Figure 7. ROC curves

5. Experiments

We use a subset consisting 513 sequences of 20 subjects (10 males and 10 females) from our clothes-variation gait dataset for training the PCA subspace. The sequences consist of different clothes types ranging from 18 to 28 for each subject. These sequences are also used for learning the discrimination capabilities and distance distribution between the same and different clothes.

For testing, a gallery set consists of standard clothes (clothes type 9) sequences of 32 subjects other than the 20 training subjects from the dataset. A probe set is composed of 546 sequences from these 32 subjects with clothes other than clothes type 9.

The proposed method was compared with the whole-based method to confirm the effectiveness. The whole-based method used the entire frequency-domain features which are dimension-reduced by PCA in the same way as the part-based method.

The gait identification performances were evaluated by the Receiver Operating Characteristics (ROC) curves. The ROC curves shows relation between the False Rejection Rate (FRR) and the False Acceptance Rate (FAR) when the receiver changes the acceptance thresholds. In the ROC evaluation, a curve near the left bottom corner indicates high performance because it realizes both low false rejection and low false acceptance.

ROC curves for the experiments are illustrated in Fig. 7. The ROC curves for the part-based method are closer to the left bottom than the whole-based method. Then, Equal Error Rate (EER) is picked up by calculating an intersection of the ROC curve and EER line as a typical performance measure for the ROC curve. As a result, EER decreases from 30.5% for the whole-based method to 16.0% for the part-based method and it turns out the effectiveness of the proposed method.

6. Conclusions

This paper describes a method of part-based gait identification under substantial clothes variations. First, a clothes-variation gait dataset including 52 subjects with at most 32 combinations of clothes is constructed. Next, 8 body parts are defined based on the anatomical statistics, and dimension-reduced frequency-domain features are used as part-based gait features. Then, discrimination capabilities and probability distribution functions (PDFs) of distances between a probe and a gallery are acquired both for the same and different clothes for the 8 parts separately in the training phase. Subsequently, given a probe in test phase, matching weights for the 8 parts are adaptively determined by distances between the probe and all the galleries based on the trained discrimination capabilities and PDFs. As a result of experiments using the constructed dataset, it turned out that the proposed method reduced Equal Error Rate (EER) as much as half of a whole-based method.

Acknowledgements

This work is supported by the Special Coordination Funds for Promoting Science and Technology of Ministry of Education, Culture, Sports, Science and Technology.

References

[1] N. Boulgouris and Z. Chi. Human gait recognition based on matching of body components. *Pattern Recognition*, 40(6):1763–1770, 2007.

[2] W. Dempster and G. Gaughran. Properties of body segments based on size and weight. *American Journal of Anatomy*, 120(1):33–54, 1967.

[3] J. Han and B. Bhanu. Individual recognition using gait energy image. *Trans. on Pattern Analysis and Machine Intelligence*, 28(2):316– 322, 2006.

[4] S. Lee, Y. Liu, and R. Collins. Shape variation-based frieze pattern for robust gait recognition. In *Proc. of the 2007 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pages 1–8, Minneapolis, USA, Jun. 2007.

[5] 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*, volume 3, pages 151–163, Graz, Austria, May 2006.

[6] 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):162–177, 2005.

[7] J. Shutler, M. Grant, M. Nixon, and J. Carter. On a large sequence-based human gait database. In *Proc. of the 4th Int. Conf. on Recent Advances in Soft Computing*, pages 66–71, Nottingham, UK, Dec. 2002.

[8] S. Yu, D. Tan, and T. Tan. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. In *Proc. of the 18th Int. Conf. on Pattern Recognition*, volume 4, pages 441–444, Hong Kong, China, Aug. 2006.