

Gait Recognition Based on Fusion of Multi-view Gait Sequences

Yuan Wang¹, Shiqi Yu¹, Yunhong Wang², and Tieniu Tan¹

¹ National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, P.O. Box 2728, Beijing 100080, China

² School of Computer Science and Engineering, Beihang University
{ywang, sqyu, wangyh, tnt}@nlpr.ia.ac.cn

Abstract. In recent years, many gait recognition algorithms have been developed, but most of them depend on a specific view angle. In this paper, we present a new gait recognition scheme based on multi-view gait sequence fusion. An experimental comparison of the fusion of gait sequences at different views is reported. Our experiments show the fusion of gait sequences at different views can consistently achieve better results. The Dempster-Shafer fusion method is found to give a great improvement. On the other hand, we also find that fusion of gait sequences with an angle difference greater than or equal to 90° can achieve better improvement than fusion of those with an acute angle difference.

1 Introduction

Gait has recently received an increasing interest from researchers. Gait is an attractive biometric feature for human identification at a distance, which is non-contact, non-invasive and easily acquired at a distance in contrast with other biometrics, so it has been considered as the most suitable biometric for human identification in visual surveillance.

Over the past years many gait recognition algorithms [1, 2, 3, 4, 5] have been proposed, but most of them are dependent on only one view, normally side view and have low recognition rate due to the influence of clothing, background, light, walker's mental state, etc. How to develop a robust and accurate gait recognition system has become an important direction. The purpose of this paper is therefore to present a new gait recognition scheme based on the fusion of multi-view gait sequences which can improve the performance of gait recognition system greatly and can be used in practice conveniently. Because in many surveillance environments, multiple cameras at different view angles are used, this makes it possible to get gait sequences from multi-view directions. Unlike most previous studies that focus on extracting good features, this paper is trying to construct a multi-view gait recognition system which is more robust and accurate.

The remainder of this paper is organized as follows. Section 2 briefly introduces the Key Fourier Descriptor (KFD) method for gait recognition and fusion rules for multi-view gait sequence fusion. Section 3 presents the CASIA multi-view gait database. Then the main scheme of the fusion system is presented in Section 4. Section 5 introduces our experimental results and analysis. Finally, this paper is concluded in Section 6.

2 KFD Gait Recognition Method and Fusion Rule

2.1 KFD Gait Recognition Method

Given a fixed camera, the human silhouette can be extracted by background subtraction and thresholding. We take advantage of the method given in [2] to segment human silhouettes from image sequences. Since the extracted silhouette sizes are not unique, the height of them is normalized to a fixed size.

To extract KFDs feature, all the contours and the gait cycle are normalized to have the same number (N) of samples and the same number (T) of frames, respectively. All Fourier descriptors $g(i)$ can be obtained by discrete Fourier transform. The KFDs are defined as in [5]:

$$G = \left[\frac{|g(2T)|}{|g(T)|}, \frac{|g(3T)|}{|g(T)|}, \dots, \frac{|g((N-1)T)|}{|g(T)|} \right] \tag{1}$$

where g is Fouries descriptors.

To measure the similarity between two gait sequences a and b , we use the metric shown in Equation (2) as the similarity measure:

$$D(a, b) = \frac{1}{M} \sum_{m=1}^M |G_a(m) - G_b(m)| \tag{2}$$

where G_a and G_b are, respectively, feature vectors of sequences a and b , and M is the feature vector length.

2.2 Overview of Fusion Rule

We used 4 traditional fusion methods in our experiments. For the verification mode in biometrics authentication system, the incomer will be compared with the template of the person he claims. We treat the outputs of each authenticaiton system as a feature vector $X = [x_1, x_2, \dots, x_N]$, where N is the number of subsystems and x_i is the output of each subsystem. Then we can use any known classifier to determine the separation bound between imposter and client. In our paper, we use the following 4 kinds of fusion rules:

1. Sum Rule

$$x = \sum_{i=1}^N x_i \tag{3}$$

2. Weighted Sum Rule

$$x = \sum_{i=1}^N w_i * x_i \tag{4}$$

Here, w_i is computed by the EER of each fusion system: $w_i = \frac{ERR_i^{-1}}{\sum_{j=1}^N ERR_j^{-1}}$

3. Product Rule

$$x = \prod_{i=1}^N x_i \tag{5}$$

4. Dempster-Shafer(D-S) rule

In this frame of the evidence theory, the best representation of support is a belief function rather than a Bayesian mass distribution. The theory embraces the familiar idea of assigning numbers between 0 and 1 to indicate the degree of support but, instead of focusing on how these numbers are determined, it concerns the combination of degrees of belief. Here, we use the algorithm proposed in [6].

The decision of fusion system can be made based on x computed by these methods.

3 CASIA Multi-view Gait Database

In our experiments, we used the CASIA multi-view gait database, which contains gait sequences of 124 subjects (94 males, 30 females) taken from 11 cameras at 11 different views. All the subjects were asked to walk naturally on the concrete ground along a straight line in an indoor environment. The videos were captured by 11 cameras from different view directions. The view angle θ between the view direction and the walking direction took on the values of $0^\circ, 18^\circ, 36^\circ, \dots$, and 180° , as delineated in Fig. 1. Each subject walked along the straight line 10 times (6 for normal walking, 2 for walking with a bag, and 2 for walking with a coat), and 11 video sequences are captured each time. Thus, 110 sequences were recorded for each subject, and the database contains a total of $110 \times 124 = 13640$ video sequences. All the video sequences have the same resolution of 320×240 pixels. Some sample frames are shown in Fig. 2.

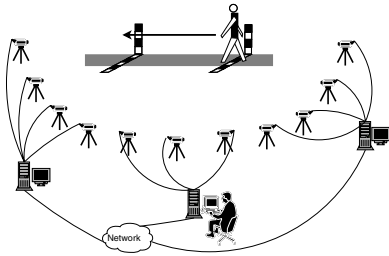


Fig. 1. The schematic diagram of gait data collection system

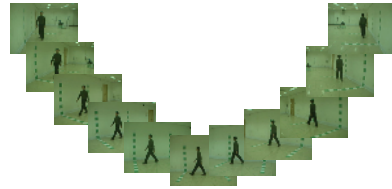


Fig. 2. Sample frames from 11 view directions

In the database those factors affecting gait recognition, such as view direction (11 views), clothing (with or without coat), and carrying condition (with or without bag) are included. Here only view direction is studied, though other factors are interesting to study too.

4 Fusion Scheme

Based on the KFD gait recognition method, we can get the similarity measures between the two gait sequences and the template as shown in Fig. 3. And before combining the

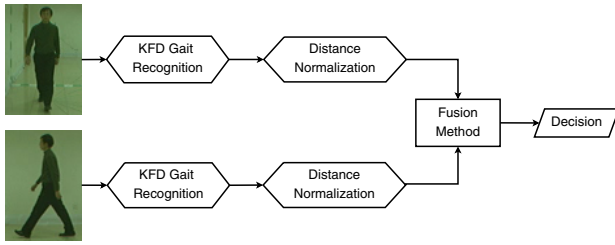


Fig. 3. The scheme of fusion system

two similarity measures by fusion, we should normalize those similarity measures to a common range $[0, 1]$. Here we use the Min-Max normalization method [7]:

$$s' = f(s) = \frac{s - \min}{\max - \min} \tag{6}$$

where, s' denotes the normalized score.

For the fusion methods, we use sum rule, weighted sum rule, product rule and Dempster-Shafer rule introduced in the previous section. The first two methods belong to the category of fixed rule. And for the other two rules which need training, we use 10% of score data as the training set.

5 Experiment Results and Analysis

5.1 Experimental Results

The EERs (Equal Error Rates) of the multi-view gait sequences fusion system are shown in Table 1 (only partial results because of length limitation), where the second column shows the EERs of gait recognition system of Angle 1 using KFD method, and so does Column 4. And for each view angle, the other 10 views are combined with it and so there are totally $C_{11}^2 = 55$ different combinations in our experiments.

5.2 Discussions

Based on the above results, we can draw some conclusions. When using sum rule as the fusion method, the average EER of 55 fusion experiments is 9.08%, and for product rule and weighted sum rule, it becomes 8.56% and 8.85%. D-S rule gives the lowest average EER along the four methods which is only 3.81%. Within the gait recognition system of 11 views, there are only 27% of them whose average EERs are less than 10%. But for the 55 fusion systems using sum rule, 75% of them give EERs less than 10% and if using D-S rule, 85% of fusion systems' EERs are less than 5%. On the other hand, within the 55 fusion experiments, 7 experiments fail to give improvement comparing the best single system when using sum rule and 6 experiments fail when using product rule. While for the trained rules, the number of failures becomes 2 and 0 corresponding to weighted sum rule and D-S rule respectively. So we can draw a conclusion that the trained rules are better than fixed rules from the view of whether fusion system can give improvement.

Table 1. The EERs of Multi-view Gait Sequence Fusion System

| Angle1 | EER | Angle2 | EER | Sum | Product | W-Sum | D-S |
|--------|--------|--------|--------|--------|---------|--------|-------|
| 0° | 13.98% | 18° | 15.31% | 12.00% | 11.97% | 11.43% | 5.86% |
| 0° | 13.98% | 54° | 11.34% | 10.28% | 8.48% | 10.08% | 4.70% |
| 0° | 13.98% | 90° | 8.05% | 7.38% | 7.89% | 6.66% | 2.69% |
| 0° | 13.98% | 126° | 11.80% | 9.54% | 8.63% | 9.25% | 4.11% |
| 0° | 13.98% | 162° | 10.44% | 9.02% | 8.81% | 9.04% | 3.90% |
| 18° | 15.31% | 36° | 12.16% | 12.70% | 12.17% | 12.66% | 5.21% |
| 18° | 15.31% | 72° | 8.08% | 8.94% | 8.20% | 7.46% | 3.21% |
| 18° | 15.31% | 108° | 9.96% | 8.79% | 7.80% | 8.95% | 3.12% |
| 18° | 15.31% | 144° | 11.50% | 11.08% | 10.12% | 10.39% | 4.67% |
| 18° | 15.31% | 180° | 13.75% | 10.94% | 9.84% | 11.38% | 4.46% |
| 36° | 12.16% | 54° | 11.34% | 10.73% | 9.33% | 10.78% | 4.65% |
| 36° | 12.16% | 90° | 8.05% | 8.16% | 8.48% | 7.81% | 3.00% |
| 36° | 12.16% | 126° | 11.80% | 9.36% | 8.74% | 9.37% | 3.98% |
| 36° | 12.16% | 162° | 10.44% | 9.30% | 8.69% | 9.30% | 3.56% |
| 54° | 11.34% | 72° | 8.08% | 7.96% | 8.88% | 7.45% | 3.40% |
| 54° | 11.34% | 108° | 9.96% | 8.97% | 8.27% | 8.53% | 3.58% |
| 54° | 11.34% | 144° | 11.50% | 9.02% | 8.35% | 9.13% | 3.76% |
| 54° | 11.34% | 180° | 13.75% | 9.52% | 8.41% | 9.82% | 4.06% |
| 72° | 8.08% | 90° | 8.05% | 6.61% | 6.59% | 6.62% | 2.46% |
| 72° | 8.08% | 126° | 11.80% | 8.53% | 8.36% | 8.05% | 2.99% |
| 72° | 8.08% | 162° | 10.44% | 6.41% | 6.85% | 6.49% | 2.47% |
| 90° | 8.05% | 108° | 9.96% | 7.42% | 7.08% | 7.10% | 3.43% |
| 90° | 8.05% | 144° | 11.50% | 7.39% | 6.83% | 6.88% | 2.59% |
| 90° | 8.05% | 180° | 13.75% | 7.54% | 6.98% | 7.45% | 2.68% |
| 108° | 9.96% | 126° | 11.80% | 8.71% | 7.84% | 8.60% | 5.06% |
| 108° | 9.96% | 162° | 10.44% | 7.48% | 7.45% | 7.33% | 3.55% |
| 126° | 11.80% | 144° | 11.50% | 10.66% | 9.87% | 10.66% | 5.23% |
| 126° | 11.80% | 180° | 13.75% | 10.30% | 10.03% | 10.24% | 4.84% |
| 144° | 11.50% | 162° | 10.44% | 9.40% | 8.90% | 9.37% | 4.53% |
| 144° | 11.50% | 180° | 13.75% | 9.96% | 9.19% | 9.78% | 4.68% |
| 162° | 10.44% | 180° | 13.75% | 9.48% | 9.31% | 9.21% | 4.46% |

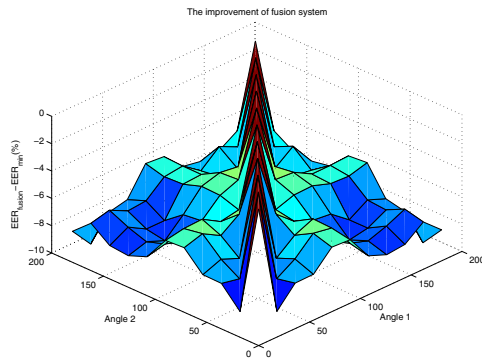


Fig. 4. The Improvement of Combination System Using D-S Rule

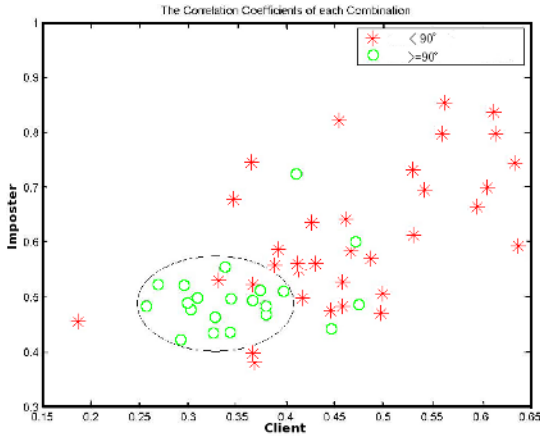


Fig. 5. The Correlation Coefficients of Each Combination

According to Table 1, we can also find that the view differences of those systems without improvements are all less than 90° .

It means that fusion of gait sequences with an angle difference greater than or equal to 90° can achieve better improvement than fusion of those with an acute angle difference. This conclusion is also illustrated in Fig. 4, where the z axis shows the difference between the EERs of fusion system and the lowest EERs of the single systems (the lower of the surface, the greater improvement). It is very clear that the difference values between the two EERs are much larger in the acute angle zone. This is mainly because that the information contained in the two gait sequences with an acute angle difference is more correlated than those with larger angle difference. In other words, much more information is contained in case of an obtuse angle (including a right angle) than an acute angle. In Fig. 5, we computed the correlation coefficients of the output score of two single view gait recognition systems. The x axis denotes the correlation coefficients of the score of client data and y axis denotes that of imposter data. We can find that the correlation coefficients of the gait sequences with an acute angle differences are scattered in the whole area and for those with larger angle differences, the coefficients are gathered into a small ellipse as shown in the figure.

6 Conclusion and Future Work

In this paper, we have presented a new gait recognition scheme based on the fusion of multi-view gait sequences. Experimental results show that the proposed system can help improve the performance of gait recognition system. Specifically, when using Dempster-Shafer fusion rule, the combination EERs mostly drop to around 5%, which achieves a great improvement when compared to the EERs of single view gait recognition.

Multi-view gait recognition is a new direction in this field and there are many open questions to address. For multi-view gait sequence taken from one subject, how to

extract some common features existing in all views is an interesting direction. Our current work is focused on score level fusion and further work should include the feature level fusion of multi-view gait sequences. In addition the database should include some outdoor data which is more similar to the data in practical application.

Acknowledgement

This work is partly supported by National Natural Science Foundation of China (Grant No. 60332010 and 60335010) and the National Basic Research Program of China (Grant No. 2004CB318100).

References

1. Kale, A., Sundaresan, A., Rajagopalan, A.N., Cuntoor, N.P., roy Chowdhury, A.K., Krüger, V., Chellappa, R.: Identification of humans using gait. *IEEE Transactions on Image Processing* **13** (2004) 1163–1173
2. Wang, L., Tan, T., Ning, H., Hu, W.: Silhouette analysis-based gait recognition for human identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **25** (2003) 1505–1518
3. Wang, L., Ning, H., Tan, T., Hu, W.: Fusion of static and dynamic body biometrics for gait recognition. *IEEE Transactions on Circuits and Systems for Video Technology* **14** (2004) 149–158
4. Yam, C.Y., Nixon, M.S., Carter, J.N.: On the relationship of human walking and running: automatic person identification by gait. In: *Proc. of International Conference on Pattern Recognition, Quebec, Canada* (2002) 287–290
5. Yu, S., Wang, L., Hu, W., Tan, T.: Gait analysis for human identification in frequency domain. In: *Proc. of the 3rd International Conference on Image and Graphics, Hong Kong, China* (2004) 282–285
6. Y, S., T, K.: Media-integrated biometric person recognition based on the dempster-shafer theory. In: *16th International Conference on Pattern Recognition*. (2002) 381–384
7. Jain, A.K., Nandakumar, K., Ross, A.: Score normalization in multimodal biometric systems. to appear in *Pattern Recognition* (2005)