



An Optimized Algorithm on Multi-view Transform for Gait Recognition

Lingyun Chi, Cheng Dai, Jingren Yan, and Xingang Liu^(✉)

School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China
Hanksliu@uestc.edu.cn

Abstract. Gait is one of the common used biometric features for human recognition, however, for some view angles, it is difficult to exact distinctive features, which leads to hindrance for gait recognition. Considering the challenge, this paper proposes an optimized multi-view gait recognition algorithm, which creates a Multi-view Transform Model (VTM) by adopting Singular Value Decomposition (SVD) on Gait Energy Image (GEI). To achieve the goal above, we first get the Gait Energy Image (GEI) from the gait silhouette data. After that, SVD is used to build the VTM, which can convert the gait view-angles to 90° to get more distinctive features. Then, considering the image matrix is so large after SVD in practice, Principal Component Analysis (PCA) is used in our experiments, which helps to reduce redundancy. Finally, we measure the Euclidean distance between gallery GEI and transformed GEI for recognition. The experimental result shows that our proposal can significantly increase the richness of multi-view gait features, especially for angles offset to 90° .

Keywords: Gait recognition · Gait energy image · View transform model
Principal component analysis

1 Introduction

With the development of information technology, human identification has been widely studied. Recently, human identification bases on biological feature has been the hottest topic in this field, due to the uniqueness of biological characteristics. Biometric based human identification technology refers to identifying the human identity using the body's different inherent physiological characteristics or behavioral characteristics. In our past research, the common used physiological biological features are fingerprint recognition [1], face recognition [2], iris recognition [3]. (1) Fingerprint recognition: fingerprint refers to the ridges on the frontal skin of human fingers. The starting point, the ending point, the joining point and the bifurcation point of the ridge line are feature points, which are different from people to people. The fingerprint recognition is wildly used in many situations, such as unlock the phone and unlock the door. (2) Face recognition, the human face consists of several parts, such as eyes, nose, mouth, and chin. The geometric description of these parts and the structural relationship between them can be used as important features to identify faces. The face recognition is now used in mobile payment and many other fields. (3) Iris recognition. The eye-structure of

human consists of sclera, iris, pupil lens, retina and other parts. The iris is an annular portion between the black pupil and the white sclera, which contains many detailed features such as staggered spots, filaments, coronal, stripes and crypts. All these details are unique for human beings, so iris can be used to identify people. Iris is often used in high security requirement situations. In short, all these technologies bring great convenience in our daily life. However, the methods mentioned above are based on the humans' cooperation and the identification task is only suitable for short distance recognition, which may cause masquerade and hidden problems. To overcome these problems, researchers started to focus on the human behavioral characteristics, and the human gait identification is proposed.

After several decades of development, human gait identification has already developed many mature frameworks, in this case, some common problems in gait recognition are solved to some extent such as human with different clothes, human carrying different things, and human in different light conditions. These problems are set in certain condition, one of the limits is they often conduct the research in a certain view angle, however, in practice, the camera is often fixed, and it only get certain gait view angles when the human walks through the captured area in different directions, which makes it difficult to acquire the overall information of human gait, especially for parallel conditions. To solve this problem caused by multi-view, researchers proposed several methods. View Transform Model (VTM), which can realize the mutual transformation between different view angles, makes it possible to obtain more abundant information, hence, the model is widely acknowledged by researchers. Based on VTM, in this paper, we propose an optimized algorithm, which converts all other view angles to 90° thus to obtain more gait features.

The rest of the article is organized as follows. In Sect. 2, we make a summary of the solutions designed for multi-view human gait recognition, and analyze the main challenges of them. In Sect. 3, we introduce the concept involved in View Transform Model (VTM) and demonstrate its application possibility for video-based human gait recognition. In Sect. 4, after demonstrating the experiment result, we discuss the preview of VTM based multi-view gait recognition, which could be utilized under different scenarios. Finally, we conclude the article and discuss the future directions of this research theme in Sect. 5.

2 Technical Background and Related Works

In general, gait recognition has three main steps, (1) gait detection and tracking, (2) gait feature extraction, (3) gait feature matching, after the three main steps, the recognition result can be obtained. As shown in Fig. 1, the captured video sequences containing probe gait are imported into the gait recognition system, and the human silhouettes are extracted from each frames. After that, we can extract the gait features according to the human silhouettes information. Finally, the extracted features are matched with the gallery gait and then recognized. In this case, the key topic of gait recognition is focusing on how to obtain more distinctive features. However, when it comes to multi-view gait recognition, the key topic can be specifically described as matching extracted features from different view angles properly.

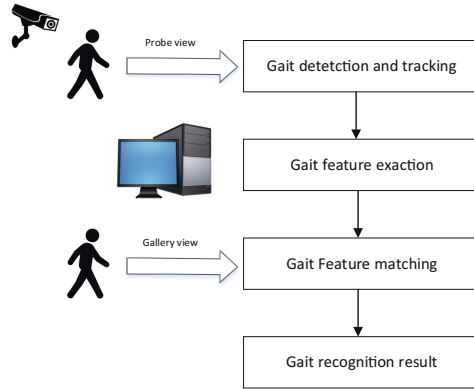


Fig. 1. A typical flow of gait recognition.

After several decades of development, there have been many resolutions proposed in allusion to multi-view gait features extraction, (1) seeking view-invariant gait characteristics; (2) constructing 3D gait model with couples of cameras; (3) establishing view transform model.

For the first method, Liu [4] represented samples from different views as a linear combination of these prototypes in the corresponding views, and extracted the coefficients of feature representation for classification. Kusakunniran [5] extracted a novel view-invariant gait feature based on Procrustes Mean Shape (PMS) and consecutively measure a gait similarity based on Procrustes distance. In general, the common features in different view angles can be extracted in this method, thus, it is possible to realize gait recognition when the view angle span is wide. However, the common feature is often insufficient, which lead that the recognition performance is relatively poor.

For the second method, Kwolek [6] identified a person by motion data obtained by their unmarked 3D motion-tracking algorithm. Wolf [7] presented a deep convolutional neural network using 3D convolutions for multi-view gait recognition capturing spatial-temporal features. Comparably, this kind of methods can achieve higher recognition accuracy, however, constructing 3D model requires a heterogeneous layer learning architecture with additional calibration, as a result, the computational complexity constructing 3D model is significantly higher than other methods.

For the third method, Makihara [8] proposed a method of gait recognition from various view directions using frequency-domain features and a view transformation model. Kusakunniran [9] applied Linear Discriminant Analysis (LDA) for further simplifying computing. Compared with the two methods above, constructing VTM can achieve state-of-the-art accuracy with less cost, in this case, VTM based approaches become the mainstream approach of the multi-view gait recognition. Considering its' advantage, in this paper, our proposal is based on VTM.

Building VTM, however, the recognition effect depends largely on the selection of which angle we transform other angles to. Meanwhile, the computational complexity when building the VTM model is relatively high. Considering the challenges mentioned above, we propose an optimized algorithm, which select a better gallery view

angle for exacting more gait features and reduce the computational complexity. In this algorithm, we first get humans' Gait Energy Image (GEI) from videos containing walking sequences. After that, we transform all other view angles to 90° to exact more gait features, and then apply Principal Component Analysis (PCA) to reduce computational complexity. Finally, we measure the Euclidean distance between gallery GEI and transformed GEI for recognition.

3 Related Concept and Proposed Algorithm

3.1 Extracting Gait Energy Image (GEI)

The concept of Gait Energy Image (GEI) was first put forward by Han [10], which presents gait energy by normalizing gait silhouette in periods. As shown in Fig. 2, the higher the brightness, the higher the probability of silhouette appearing in the position.

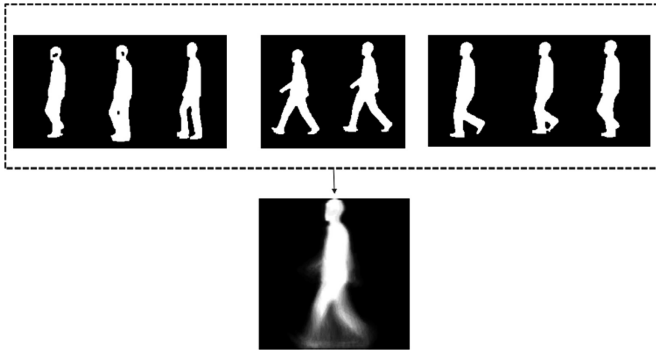


Fig. 2. A sample of gait energy image (GEI)

To get the GEI, as a basis, we need to estimate the number of gait periods contained in each gait sequence. As mentioned in [11], we can determine the gait period by calculating the ratio of height (H) to width (W) of a person's silhouette. Here, we use N to represent the number of periods contained in one gait sequence, and use T to represent the number of frames contained in one gait period. Then, we normalize all silhouettes by rescaling them along both horizontal direction and vertical direction to the same Width (W) and the same Height (H). The GEI can be obtained by:

$$GEI(x, y) = \sum_{n=1}^N \sum_{t=1}^T S_{n,t}(x, y) / (N + T) \quad (1)$$

Where, $S_{n,t}(x, y)$ represents a particular pixel locating in position (x, y) of t -th ($t = 1, 2, \dots, T$) image from n -th ($n = 1, 2, \dots, N$) gait period of a gait sequence.

3.2 Constructing an Optimized View Transform Model (VTM)

GEI contains several significant features, including gait silhouette, gait phase, gait frequency etc. However, the richness of GEI features in different view angles is different. As shown in Fig. 3, GEI in 90° contains most gait information. Therefore, we propose to transform other different angles to 90° for more distinctive features.

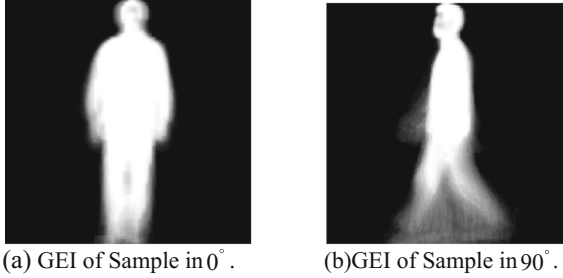


Fig. 3. Comparison between GEI of one sample in 0° and 90°

Makihara [8] first put forward the concept of VTM in his paper. In order to construct VTM, we apply the method of Singular Value Decomposition (SVD). SVD is an effective way to extract eigenvalue. For any matrix, it can be represented as the following form:

$$A = U \Sigma V^T \quad (2)$$

If the size of A is $M \times N$, U is a orthogonal matrix of $M \times M$, V is a orthogonal matrix of $N \times N$. Σ is a $M \times N$ matrix, in addition to the diagonal elements are 0, elements on diagonal called singular value.

We create a matrix G_K^M with K rows and M columns, representing gait data containing K angles and M individuals.

$$G_K^M = \begin{bmatrix} g_1^1 & \cdots & g_1^m \\ \vdots & \ddots & \vdots \\ g_k^1 & \cdots & g_k^m \end{bmatrix} = USV^T = \begin{bmatrix} P_1 \\ \vdots \\ P_K \end{bmatrix} [v^1 \quad \cdots \quad v^M] \quad (3)$$

In formula 3, g_k^m is a column vector of N_g , representing the GEI characteristics of the m -th individual at the k -th angle. The size of U is $KN_g \times M$, the size of V and S are $M \times M$. $P = [P_1 \quad \cdots \quad P_K]^T = US$. After the singular value decomposition is completed, we can calculate the g_k^m with this formulation

$$g_k^m = P_k \times v^m \quad (4)$$

Then, suppose G_θ^m representing probe GEI feature with view θ and $\hat{G}_{(\varphi \leftarrow \theta)}^m$ representing the transformed GEI feature with view φ . Firstly, using the probe GEI feature G_θ^m estimate a point on the joint subspace $\hat{G}_{(\leftarrow \theta)}^m$ by

$$\hat{G}_{(\leftarrow \theta)}^m = P(\theta)^+ G_\theta^m \quad (5)$$

$$\text{Where } P(\theta)^+ = ((P(\theta))^T P(\theta))^{-1} P(\theta)^T \quad (6)$$

Where $\|\cdot\|_2$ denotes the L2 norm.

Secondly, the GEI feature $\hat{G}_{(\varphi \leftarrow \theta)}^m$ of transformed view can be generated by projecting the estimated point on the joint subspace.

$$\hat{G}_{(\varphi \leftarrow \theta)}^m = P(\varphi) \hat{G}_{(\leftarrow \theta)}^m \quad (7)$$

3.3 Applying Principal Component Analysis (PCA)

PCA [12] replace the original n features with less number of m for further reducing computational redundancy. New features is a linear combination of the characteristics of the old. These linear combinations maximize the sample variance, and make the new m features mutually related. Mapping from old features to new features captures the inherent variability in the data.

The main processes of PCA are:

Supposing there are N gait samples, and the grayscale value of each sample can be expressed as a column vector x_i with a size of $M \times 1$, the sample set can be expressed as $[x_1, x_2 \cdots x_N]$.

The average vector of the sample set:

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N x_i \quad (8)$$

The covariance matrix of the sample set is:

$$\Sigma = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(x_i - \bar{x})^T \quad (9)$$

Then, calculating the eigenvectors and eigenvalues of Σ , the eigenvalues of X can be arranged in the following order $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_N$. We take the eigenvectors corresponding to the important eigenvalues to get a new dataset.

3.4 Contrasting Gait Similarity

In this paper, Euclidean distance is adopted to measure the similarity between gallery GEI (G^i) and transformed GEI (G^j). The Euclidean distance can be obtained by:

$$d(G^i, G^j) = \sum_x \sum_y |G^i(x, y) - G^j(x, y)| \quad (10)$$

Where, $d(G^i, G^j)$ refers to Euclidean distance between G^i and G^j , (x, y) refers to the location of one specific pixel in GEI. The smaller the Euclidean distance is, the more likely the two GEIs belong to the same person.

4 Experiment

CASIA - B [13] is an extensively used database, which was collected in 2005. It contains a total of 124 objects, each object separately contains 11 view-angles, which take 18° as ladder, range from 0° to 180° , and meanwhile, there are 6 video sequences for each person in each angle. The sketch map of the database is showed as Fig. 4. In this paper, we construct VTM with 100 objects in one of the video sequences, and use other 24 objects to evaluate the performance of our proposal.

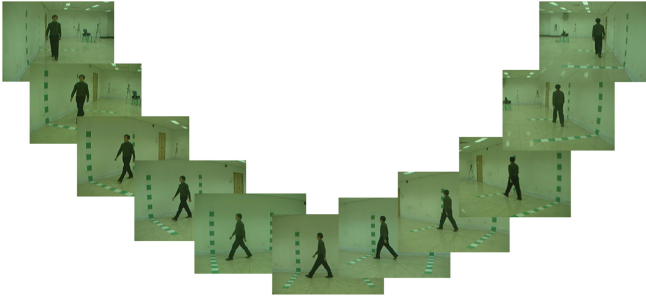


Fig. 4. The sketch map of the database

The overview framework of our proposal is shown in Fig. 5.

Figure 6 illustrates one object transforming GEI view angle from 36° to 90° as example, which reflect the performance of VTM. Table 1 shows the performance of our proposal on the CASIA-B dataset. The first column indicates the probe view angle, and the second column indicates the accuracy of our proposal. Our proposal can significantly increase the gait feature information, in particular, the recognition accuracy is relatively higher when the angles are close to 90° .

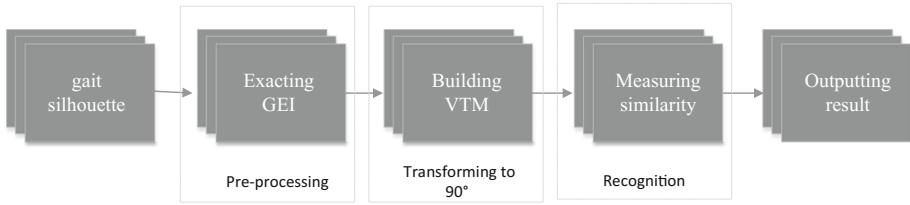


Fig. 5. The structure of the algorithm

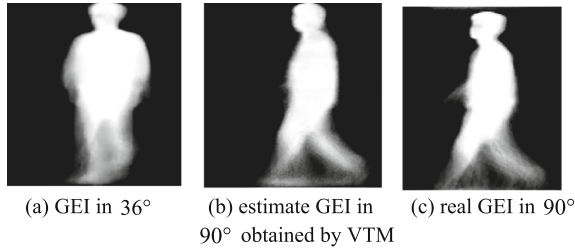


Fig. 6. Comparison between estimate GEI obtained by VTM with the real GEI in 90°

Table 1. Recognition accuracies with VTM

Probe view angle	Transformed to 90°
0°	81.3%
18°	90.6%
36°	91.5%
54°	95.2%
72°	98.1%
108°	97.6%
126°	93.1%
144°	90.7%
162°	89.4%
180°	80.9%

5 Conclusion and Future Work

The gait identification has gradually stepped into one of the mainstream approaches of biometric identification, considering its' limitation caused by multi-view in practice, we choose 90° as galley gait view to get more gait features. To transform other view angles to 90°, we first exact GEI from gait silhouettes, then, we construct a view transform model with SVD, and finally, we adopt PCA to further reduce the computational complexity. It can be drawn from Table 1 that our method can significantly improve the multi-view recognition performance.

The algorithm of multi-view gait recognition is still in the stage of continuous improvement, based on the researches we have done, our further study works mainly contains following aspect. Firstly, our proposal is suitable only for specific angles in the database, we will working for constructing a more ubiquitous view transformation model, which can realize the mutual transformation between arbitrary angles. Secondly, real-time recognition is not considered in this proposal, and how to establish a real-time multi-view gait recognition system is a tough task to be solved.

Acknowledgements. This work was supported by the Fundamental Research Funds for the Central Universities on the grant ZYGX2015Z009, and also supported by Applied Basic Research Key Programs of Science and Technology Department of Sichuan Province under the grant 2018JY0023.

References

1. Galar, M., et al.: A survey of fingerprint classification part I: taxonomies on feature extraction methods and learning models. *Knowl. Based Syst.* **81**(C), 76–97 (2015)
2. Jayakumari, V.V.: Face recognition techniques: a survey. *World J. Comput. Appl. Technol.* **1**(2), 41–50 (2013)
3. De Marsico, M., Petrosino, A., Ricciardi, S.: Iris recognition through machine learning techniques: a survey. *Pattern Recognit. Lett.* **82**(2), 106–115 (2016)
4. Liu, N., Lu, J., Tan, Y.-P.: Joint subspace learning for view-invariant gait recognition. *IEEE Signal Process. Lett.* **18**(7), 431–434 (2011)
5. Kusakunniran, W., et al.: A new view-invariant feature for cross-view gait recognition. *IEEE Trans. Inf. Forensics Secur.* **8**(10), 1642–1653 (2013)
6. Kwolek, B., Krzeszowski, T., Michalczuk, A., Josinski, H.: 3D gait recognition using spatio-temporal motion descriptor. In: *Asian Conference on Intelligent Information and Database Systems*, pp. 595–604. Springer, Cham (2014)
7. Wolf, T., Babae, M., Rigoll, G.: Multi-view gait recognition using 3d convolutional neural networks. In: *IEEE International Conference on Image Processing*, pp. 4165–4169. IEEE, USA (2016)
8. Makihara, Y., et al.: Gait recognition using a view transformation model in the frequency domain. In: *European Conference on Computer Vision*, pp. 151–163. Springer, Austria (2006)
9. Kusakunniran, W., et al.: Multiple views gait recognition using view transformation model based on optimized gait energy image. In: *IEEE International Conference on Computer Vision Workshops*, pp. 1058–1064. IEEE, Japan (2010)
10. Han, J., Bhanu, B.: Individual recognition using gait energy image. *IEEE Trans. Pattern Anal. Mach. Intell.* **28**(2), 316–322 (2005)
11. Gu, J., Ding, X., Wang, S., et al.: Action and gait recognition from recovered 3-D human joints. *IEEE Trans. Syst. Man Cybern. Part B* **40**(4), 1021–1033 (2010)
12. Qing-Jiang, W.U.: Gait Recognition Based on PCA and SVM. *Comput. Sci.* (2006)
13. Yu, S., et al.: View invariant gait recognition using only one uniform model. In: *International Conference on Pattern Recognition*, pp. 889–894. IEEE, Mexico (2017)