Neurocomputing ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Gait recognition method for arbitrary straight walking paths using appearance conversion machine

Xiaohui Zhao [a,b], Yicheng Jiang [a,*], Tania Stathaki [b,*], Huisheng Zhang [c]

[a] Research Institute of Electronic Engineering Technology, Harbin Institute of Technology, Harbin 150001, China
[b] Department of Electrical and Electronic Engineering, Imperial College, London SW7 2AZ, UK
[c] Department of Mathematics, Dalian Maritime University, Dalian 116026, China

ARTICLE INFO

Article history:
Received 25 October 2014
Received in revised form
19 May 2015
Accepted 4 July 2015
Communicated by Xiaoqin Zhang

Keywords:
Multi-view gait recognition
Human identification
Appearance conversion machine
Extreme learning machine
Viewing angle estimation
Transformation invariant low-rank texture

ABSTRACT

We investigate the problem of multi-view human gait recognition along any straight walking paths. It is observed that the gait appearance changes as the view changes while certain amount of correlated information exists among different views. Taking advantage of that type of correlation, a multi-view gait recognition method is proposed in this paper. First, we estimate the viewing angle of the monitor equipment in terms of the probe subject. To this end, our method considers this as a classification problem, where the classification signals are the viewing angles, and the classification features are the elements of the transformation matrix that is estimated by the Transformation Invariant Low-Rank Texture (TILT) algorithm. Then, the gallery gait appearances are converted to the view of the probe subject using the proposed Appearance Conversion Machine (ACM), where the gait features of the spatially neighbouring pixels of the gait feature are considered as the correlated information of the two views. In the end, a similarity measurement is applied on the converted gait appearance and the testing gait appearance. Experiments on the CASIA-B multi-view gait database show that the proposed gait recognition method outperforms the state-of-the-art under most views.

1. Introduction

Gait has been widely accepted as a biometric feature for human recognition. One of the proposed techniques for gait recognition is the Gait Energy Image (GEI) [21]. This technique has inspired many researchers to conduct human recognition based on the GEI, owing to its high discriminative power, simplicity, and time-efficient calculation [7,12,2]. In [25], the performance of the GEI is measured and it is shown that the best recognition result is achieved when recognitions are performed under the same view especially for profiles. However, in the real world, the non-cooperate individuals usually walk in random paths (in this paper, we assume subjects walk in random straight paths). It is observed in [25] that gait recognition is sensitive to varying views based on the fact that the visual features will change as the viewing angle changes. Therefore, it is not easy to achieve acceptable recognition accuracy when probe subjects are walking in random directions [22].

Several methods have been proposed for gait recognition from various different perspectives [1,5,15,18–20,23]. The state-of-the-art methods mainly fall into 3 categories, namely, view-invariant models estimation, viewing angle rectification, and view transformation model methods. One example that estimates the view-invariant model is the Joint Subspace Learning (JSL) method proposed in [20], where the view-invariant model is represented by a weighted sum of sufficiently small number of prototypes of the same view. Similarly, in [23], a distance metric is learned with a good discrimination ability based on the clustered and averaged GEI. However, this kind of methods can achieve good recognition results only for similar views. When the probe gait sequences are significantly different from the gallery sequences, they usually have poor performances. For viewing angle rectification methods, in [18], the Transform Invariant Low rank Textures (TILT) algorithm is used to rectify the deformed human silhouette into profile silhouette, and then a similarity measurement process is applied on the rectified profile silhouettes. Regarding the TILT, a generalized framework for low-rank recovery is proposed in [27] where the optimization problem is solved by adopting a proximal gradient based altering direction method. However, this kind of methods relies on the rectified result which is not stable in all circumstances. The methods proposed in [15–17] are recently published View Transformation Model (VTM) methods. All VTM-based methods construct the gait features of the target view using the information of its corresponding view, which require a technique that can find the best corresponding information among different views with stable performance. In [19], a technique is

* Corresponding authors.
E-mail addresses: jiangyc@hit.edu.cn (Y. Jiang), t.stathaki@imperial.ac.uk (T. Stathaki).

http://dx.doi.org/10.1016/j.neucom.2015.07.012
0925-2312/© 2015 Published by Elsevier B.V.

proposed for the extraction of the optimal correlation type of information. This method involves a motion co-clustering process that partitions gaits from different views into multiple groups. After that, the two most correlated gait features of two views are applied with a mapping operation to map the gait feature from one view into another view using the trained Canonical Correlation Analysis (CCA) subspaces. However, this method is quite complicated and difficult to implement. Furthermore, the performance of the VTM-based methods under distant views still requires improvement. Overall, most of present methods can achieve good recognition results for similar views but achieve poor recognition accuracy under distant views.

This paper proposes a multi-view gait recognition method that aims at achieving improved recognition accuracy compared to existing methodologies. The widely researched GEI is employed as the gait feature and the following processes will be applied based on the GEIs. Our method includes two parts, namely, the viewing angle estimation and the multi-view gait recognition.

For viewing angle estimation, one common sense shared among multi-view gait recognition researchers is that two distant views share much less correlated information than that of similar views [15–17,19]. In that case, recognition performed among two distant views leads to poor results. Therefore, it is reasonable to perform similarity measurement on the probe subject and its nearest-neighbouring view in the gallery. To this end, we start with estimating the viewing angle of the monitor equipment in terms of the probe subject. This is considered as a classification problem in this paper, where the classification signals are the viewing angles. Since the Transformation Invariant Low-Rank Texture (TILT) algorithm can describe the degree of difficulty of transforming an image into a low-rank image [26], the low-rank image can be considered as a standard and the TILT algorithm can be used to describe this transformation. Therefore, we propose to consider the elements of the transformation matrices estimated by the TILT algorithm as the classification features, and then employ the Extreme Learning Machine (ELM) classification method to solve this classification problem. For multi-view gait recognition, it is already known that the appearances related to different views share certain amount of correlated information [15–17,19]. Therefore, it is reasonable to assume that the gait appearances referring to two different views are able to convert between each other with tolerable error by using the correlated information that can be considered as the connection of two views. As can be seen from Fig. 1 that one subject exhibits similar appearance among different views, especially for similar views. Here, we propose the Appearance Conversion Machine (ACM) to convert the gait appearances across views (from one view to another view), where the view is estimated using the above-mentioned estimation method. The proposed ACM attempts to find the correlation function across views and then convert the subject by exploring the information related to this type of correlation. In contrast with the correlation coefficient method used in [16,17] or the complicated co-clustering method in [19], we assume that one pixel in the target view is only highly correlated with the spatial neighbouring pixels of the source view. This enables the correlation information to be extracted in a consistent fashion among different views, and thus leads to a stable conversion result. The ELM is employed as it achieves good generalization performance and solves the regression problem in the proposed ACM. The ELM was originally developed from feed forward networks and later extended to kernel

Fig. 1. The appearances of the same subject under different viewing angles (from left to right 0°, 18°, 36°, 54°, 72°, 90°, 108°, 126°, 144°, 162°, 180°).

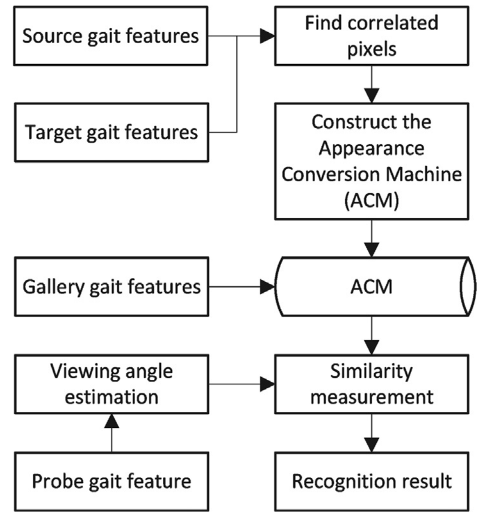

Fig. 2. The framework of the proposed recognition method.

learning [13]. This extension enables the ELM to achieve higherscalability with less computational complexity [12]. Having acquired the converted gallery (training) appearance, a similarity measurement is applied on the converted appearance and the probe appearance to identify the subject. The flowchart of the proposed gait recognition method is shown in Fig. 2. For scenarios with more than 1 camera, our method can be extended using the extended ACM by employing the gait information captured from two cameras. We explore the proposed method on the CASIA-B multi-view gait database (one of the most widely used multi-view gait databases with 124 subjects captured from 11 angles). As will be shown in the relative section, experimental results show the effectiveness of the proposed method. Furthermore, the recognition performance of the extended ACM with two cameras is complementary to the ACM with single camera. The encouraging experimental results assure the possibility of monitoring the entire scene (from 0° to 180°) with relatively high accuracy using only two cameras.

The main contributions of this work are as follows:

- We have proposed a robust method for estimating the viewing angle of the camera in terms of the probe subject.
- We have proposed to construct the Appearance Conversion Machine (ACM) for converting the gait appearance from one view to another view and thus increase the recognition performance of multi-view gait recognition by performing a similarity measurement based on the converted appearances.
- We have proposed a method for finding the correlated pixels between two views for the construction of the ACM.
- We have conducted a number of experiments on the widely used CASIA-B gait database and achieved significantly better recognition results than existing, widely used multi-view gait recognition methods.

The rest of this paper is organized as follows. The construction of the GEI is introduced in Section 2. The viewing angle estimation method and the ACM are discussed in Sections 3 and 4, respectively. Experiments and performance evaluations are discussed in Section 5. Discussions and conclusions are provided in Section 6.

2. Feature construction

In this paper, the GEI is employed as the gait feature. To construct a GEI for each subject, image patches of walking human are cropped from the sequential images of the CASIA-B gait

Please cite this article as: X. Zhao, et al., Gait recognition method for arbitrary straight walking paths using appearance conversion machine, Neurocomputing (2015), http://dx.doi.org/10.1016/j.neucom.2015.07.012