Pattern Recognition 48 (2015) 798–811

Contents lists available at ScienceDirect

Pattern Recognition

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

Robust view-invariant multiscale gait recognition

Sruti Das Choudhury, Tardi Tjahjadi *

School of Engineering, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, United Kingdom

ARTICLE INFO

Article history:
Received 10 September 2013
Received in revised form
15 August 2014
Accepted 24 September 2014
Available online 6 October 2014

Keywords:
Gait recognition
Entropy
Gaussian filter
Focus value
Weighted random subspace learning

ABSTRACT

The paper proposes a two-phase view-invariant multiscale gait recognition method (VI-MGR) which is robust to variation in clothing and presence of a carried item. In phase 1, VI-MGR uses the entropy of the limb region of a gait energy image (GEI) to determine the matching gallery view of the probe using 2-dimensional principal component analysis and Euclidean distance classifier. In phase 2, the probe subject is compared with the matching view of the gallery subjects using multiscale shape analysis. In this phase, VI-MGR applies Gaussian filter to a GEI to generate a multiscale gait image for gradually highlighting the subject's inner shape characteristics to achieve insensitiveness to boundary shape alterations due to carrying conditions and clothing variation. A weighted random subspace learning based classification is used to exploit the high dimensionality of the feature space for improved identification by avoiding overlearning. Experimental analyses on public datasets demonstrate the efficacy of VI-MGR.

1. Introduction

Compared with physiological biometrics, e.g., fingerprint, face, iris and earlobe geometry, the behavioural biometrics of gait has the advantage of being able to identify a human subject unobtrusively using low resolution video sequences [1]. The markerless gait recognition methods can be classified into appearance-based and model-based. Appearance-based methods (e.g., [2–7]) analyse the spatio-temporal shape and dynamic motion characteristics of silhouettes in a gait sequence without using a human body model. Model-based methods (e.g., [8–11]) characterise a human subject using a structural model to measure time-varying gait parameters, e.g., gait period, stance width and stride length, and a motion model to analyse the kinematic and dynamical motion parameters of the subject, e.g., rotation patterns of hip and thigh, joint angle trajectories and orientation change of limbs.

The main challenges to successful gait recognition are variation in view, variation in subject's clothing and presence of a carried item. Motivated by the unavailability of a gait recognition method that addresses all these challenges, this paper proposes the view-invariant multiscale gait recognition method (VI-MGR) which is robust to variation in clothing and presence of a carried item. VI-MGR is based on integrative scientific principles with a consideration of low computational complexity that enable it to robustly identify a human subject in the presence of numerous challenging factors of realistic scenario for effective visual surveillance.

A gait energy image (GEI) [3], which is formed by averaging the silhouettes of a gait period, is widely used in appearance-based methods as it facilitates noise-resilient robust gait feature extraction with reduced storage space and computation time. The proposed VI-MGR is an appearance-based method based on GEIs. It comprises two phases: (1) to determine the matching gallery view with the probe and (2) to identify the probe subject. Compared to other parts of a subject's body, the limb region of a GEI better captures the discriminative information due to variation in view, and is least affected by most carrying conditions and clothing variation. Thus, the phase 1 of VI-MGR computes entropy of the limb region of the GEIs based on anthropometry to determine the matching view of the probe in the gallery using 2-dimensional principal component analysis (2D PCA) and 2D Euclidean distance classifier.

The Gaussian filter is the only filter which can generate a scale-space representation of an image parameterised by the size of a smoothing kernel. Since the Gaussian filter suppresses the fine details in an image by attenuating high frequency components, it blurs an image. As the scale of the Gaussian filter increases, the blurriness increases which results in a gradual loss of boundary and exterior region, thus highlighting the inner shape characteristics. The method in [12] demonstrated the superiority of multiscale shape analysis using the Gaussian filter for shape classification compared to conventional shape classification methods, e.g., elliptic Fourier descriptor, Zernike moments, and wavelet transform. Multiscale shape analysis using the Gaussian filter is more effective in gait recognition as it results in the loss of boundary characteristics of a subject which are usually affected by the variation in clothing and carrying conditions. Although successful attempts have been made, e.g., [5–7,13–15] in using GEI to outperform the original GEI method [3], multiscale

* Corresponding author. Tel.: +44 24 76523126; fax: +44 24 76418922.
E-mail address: t.tjahjadi@warwick.ac.uk (T. Tjahjadi).

http://dx.doi.org/10.1016/j.patcog.2014.09.022
0031-3203/

analysis of a GEI has not been exploited despite its high discriminatory power. Hence, the main motivation for VI-MGR is to demonstrate the efficacy of multiscale shape analysis in gait recognition to achieve high identification rate in the presence of clothing variation and carrying conditions. Thus, in phase 2, VI-MGR analyses shape characteristics of GEIs in the image space using the Gaussian filter to exploit the discriminatory shape characteristics of the subject at three scales. The blurred GEIs obtained by the application of the Gaussian filter at different scales are combined to generate a multiscale gait image (MGI) for use as the gait signature. The probe gait signatures are matched with the gait signatures of the same view of the gallery subjects for identification. Since only dominant features persist across scales, the method is noise-resilient.

If the gallery subjects for training are recorded under similar physical conditions, the learned features in the presence of covariates are likely to cause overfitting that decreases the subject identification rate. Thus, the methods in [3,5] manually compute synthetic gait templates following a distortion model based on anthropometry to take into account of lower body part distortions due to variation in walking surface, footwear and clothes. The use of the templates enables these methods to be insensitive to the lower body-part distortions, but not upper body-part due to clothing variation and carrying conditions. The dynamics normalisation based gait recognition (DNGR) method [4] uses eigenstance reconstruction model to improve the silhouettes by reducing the effect of shadows and segmentation errors. Since there are numerous covariates, it is challenging to create the appropriate gait template for robust gait recognition. The method in [16] applies part-based strategy to adaptively assign more weight to the unaffected body parts and less weight to the affected body parts to achieve insensitiveness to clothing variation. However, it is unrealistic to train the model with all known clothing types as attempted in [16]. VI-MGR avoids overfitting, and achieves invariance to clothing variation and carrying conditions by introducing MGI with weighted random subspace learning (WRSL) for classification. WRSL is an ensemble classifier which randomly selects multiple subspaces from the feature space, with a classifier for each subspace. It addresses the apparent dilemma of accuracy optimisation and over-adaptation by exploiting high dimensionality.

The novelties of VI-MGR are the following: (1) it achieves robustness to variation in view, clothing and presence of a carried item (i.e., the three main challenges of gait recognition), as well as several other covariate factors, e.g., segmentation noise, missing body parts, change in ground surface, shadows under feet and occlusions; (2) it introduces reflected GEI (R-GEI) to create variation of the reference gallery views to address any unknown probe view in the range $[0^\circ, 360^\circ]$; (3) it provides a new approach to achieve view invariance by comparing the probe with all the reference gallery views based on entropy of the limb region of a GEI; (4) it introduces multiscale shape analysis in gait recognition, and achieves invariance to clothing variation and carrying conditions by introducing MGI; (5) It uses focus value as a measure of blurriness of the filtered GEIs to determine the ideal range of scales. A minimum number of three scales are selected from this range to make a trade-off between computational complexity and identification rate; (5) by using WRSL, VI-MGR exploits high dimensionality to avoid overfitting and achieve high identification rate.

The rest of the paper is organised as follows. Section 2 discusses related works and Section 3 presents VI-MGR. Section 4 presents the experimental results and Section 5 concludes the paper.

2. Related work

Gait recognition methods that address variation in view either depend on (a) extraction of gait features which are view invariant; (b) learning mapping or projection relationship between the gait characteristics of one view and another based on view transformation; or (c) construction of a 3-dimensional (3D) model of a subject from 2D images captured from different views using multiple calibrated cameras.

A statistical feature extraction strategy is used in [17] to extract view-invariant features from parts of a GEI which overlap between different views of a probe and a gallery subject. However, the performance of the method is not satisfactory if there is little overlap due to extreme variation in view. The method based on joint's position estimation and viewpoint rectification (JPE-VR) in [18] determines motion of a subject's lower limb based on anatomical positions of hip, shin and ankle for view-invariant gait recognition using a viewpoint rectification. However, the ankle is most likely to be occluded by the presence of shadows under feet. Since it is impossible to estimate the positions of hip and shin in the case of a subject either wearing a skirt or a long coat, and carrying an item in an upright position, the method is also not robust against variation in clothing and carrying conditions. The method in [19] projects a gait texture image formed by averaging binary gait images of a gait period of a certain view onto the canonical view based on domain transformation using transform invariant low-rank textures. The method in [20] computes view-normalised trajectories of the subject's head and feet. The normalisation involves the decomposition of walking trajectory into piecewise linear segments to transform the head and feet trajectories from different views into fronto-parallel view based on homography. Since the feet trajectory is affected by self-occlusions, the method is applicable only to a limited variation in view.

View-invariant methods based on view transformation aim to learn a mapping relationship among gait features of a subject perceived across views. To identify a subject based on gait sequences of different views, the gait features in the probe view are transformed to that of the gallery view before a distance measure is computed [21]. The method in [22] uses discrete Fourier transform (DFT) to obtain gait features from a spatio-temporal gait silhouette volume, and applies a view transformation model on the extracted gait features. The method in [23] creates a view transformation model using support vector regression based on local dynamic feature extraction to transform gait characteristics of one view into the probe view. The method in [24] uses joint subspace to learn a subject's prototype of different views, and represents the subject as a linear combination of these prototypes. Although these methods can cope with large variation in view without relying on camera calibration, they suffer from degeneracies and singularities caused by gait features which are perceived in one view but not in the other view usually due to large view angle difference with the former. Similar to the methods based on view transformation, the method based on canonical correlation analysis (CCA) in [21] also captures the mapping relationship between gait features of different views. However, instead of reconstructing gait features to the required gallery view, this method uses CCA to project the gait sequences of two views onto two maximally correlated subspaces, and uses the correlation strength as a similarity measure between the two gait sequences to overcome the problems associated with the view transformation model. But the effect of variation in clothing and carrying conditions are not considered in this method. The methods based on mapping and projection relationships [21–24] rely on supervised learning, i.e., require the availability of the gait characteristics of all views to establish a relationship among them during training.

The method in [25] constructs the subject's 3D model from 2D images captured from multiple calibrated views. 2D gait features of the probe view are then obtained from the model for view-invariant gait recognition. The method uses a stick model to simulate a subject's gait. It combines static gait characteristics obtained by anthropometric measurements of different body parts with the dynamic gait characteristics obtained by analysing the joint angle trajectories of lower limbs for identifying a subject based on linear time normalisation. In addition to variation in view, the method is