Neurocomputing 175 (2016) 324–335

Contents lists available at ScienceDirect

Neurocomputing

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

View-invariant gait recognition via deterministic learning

Wei Zeng [a], Cong Wang [b,c,*]

[a] School of Mechanical & Electrical Engineering, Longyan University, Longyan 364012, China
[b] School of Automation Science and Engineering, South China University of Technology, Guangzhou 510640, China
[c] Key Laboratory of Biomedical Engineering in Guangdong, Guangzhou 510006, China

ARTICLE INFO

Article history:
Received 3 December 2014
Received in revised form
20 September 2015
Accepted 19 October 2015
Communicated by Huaping Liu
Available online 29 October 2015

Keywords:
Gait recognition
Silhouette features
Deterministic learning
Gait dynamics
View angle variation

ABSTRACT

Performance of gait recognition can be affected by many factors, especially by the variation of view angle which will significantly change the available visual features for matching. In this paper, we present a new method to eliminate the effect of view angle for efficient gait recognition via deterministic learning theory. The width of the binarized silhouette models the periodic deformation of human gait shape and is selected as the gait feature. It captures the spatio-temporal characteristics of each individual, represents the dynamics of gait motion, and sensitively reflects the variance between gait patterns across various views. The gait recognition approach consists of two phases: a training phase and a recognition phase. In the training phase, gait dynamics underlying different individuals' gaits observed from different view angles are locally accurately approximated by radial basis function (RBF) neural networks. The obtained knowledge of approximated gait dynamics is stored in constant RBF networks. In order to address the problem of view change no matter the variation is small or significantly large, the training patterns from different views constitute a uniform training dataset containing all kinds of gait dynamics of each individual observed across various views. In the recognition phase, a bank of dynamical estimators is constructed for all the training gait patterns. Prior knowledge of human gait dynamics represented by the constant RBF networks is embedded in the estimators. By comparing the set of estimators with a test gait pattern whose view pattern contained in the prior training dataset, a set of recognition errors are generated. The average L_1 norms of the errors are taken as the similarity measure between the dynamics of the training gait patterns and the dynamics of the test gait pattern. The test gait pattern similar to one of the training gait patterns can be recognized according to the smallest error principle. Finally, comprehensive experiments are carried out on the widely adopted multiview gait databases: CASIA-B and CMU MoBo to demonstrate the effectiveness of the proposed approach.

Gait as a biometric has recently gained considerable attention because of its unobtrusiveness and gait information can be captured at a distance from a camera. It has the potential to be applied in various areas of real-world applications such as surveillance, health care, entertainment, access control and border control. Comprehensive studies conducted in this field have proved its feasibility and have achieved significant progress in the performance. However, there are a number of covariate factors that affect gait recognition performance, such as lighting condition, clothes, carrying status, shoe type, walking speed, view angle and so on [1]. Hence, in gait recognition, one important requirement is of robustness to these variations. Among these factors, view angle is one of the most important factors which heavily affects recognition performance [2–4]. The difficulties lie in that gait appearance changes due to the variation of views or walking directions, and it is impossible to expect all the subjects to walk in a particular direction.

Most of the existing gait recognition approaches are either applicable only to the side view or at least view-dependent. To handle the variation of gait sequences caused by different view angles, many researchers have proposed several methods to address this problem. They can be roughly divided into the following categories: (1) extracting view-invariant gait features; (2) projecting gait feature from one view angle to the other by using view transformation; and (3) synthesizing view angle based on a three-dimensional (3D) model.

In the first category, Kale et al. [5] synthesized a side view from any other arbitrary view using a single camera. Two methods, namely the perspective projection model and the optical flow based structure from motion equations were adopted by working with calibrated single-camera system. However, this method

* Corresponding author. Tel.: +86 20 87114615; fax: +86 20 87114612.
E-mail addresses: zw0597@126.com (W. Zeng), wangcong@scut.edu.cn (C. Wang).

http://dx.doi.org/10.1016/j.neucom.2015.10.065
0925-2312/

assumed the subjects to be far away from the camera and could not cope with large variations in view angle. Goffredo et al. [6,7] proposed a method for viewpoint independent markerless gait analysis that did not require camera calibration and worked with a wide range of walking directions. It included two consecutive stages: markerless lower limb joints' estimation from the image sequence and viewpoint rectification. However, it is hard to guarantee the accuracy of lower limb joints' estimation.

In the second category, approaches aim to learn a mapping relationship between gait features of the same subject observed across views [8–10]. When matching gait sequences from different views, the gait features are mapped into the same view before a distance measure is computed for matching. Makihara et al. [8] selected reference views for the various-view gait identification using a view transformation model. Then, it extracted frequency-domain gait features from gait silhouette sequences, and obtained the various-view gait features by transforming a few reference features with the view transformation model. Jean et al. [9] used the 2D trajectories of both feet and the head extracted from the tracked silhouettes. A homography transformation was then computed for each walking plane to make it appear as if walking were observed from a fronto-parallel view. View normalization made head and feet trajectories appear as if seen from a fronto-parallel viewpoint, which was assumed to be optimal for gait modeling purposes. In [10], a view transformation model was constructed based on regression processes by adopting multi-layer perception as a regression tool. The view transformation model estimated gait feature from one view using a well selected region of interest on gait feature from another view. Then, trained view transformation models could normalize gait features from across views into the same view before gait similarity was measured.

In the third category, a 3D model or visual hull of the walking body is usually generated for recognition under multi-camera system. In [11], principal component analysis constructed a 3D linear model from a set of Fourier represented examples. A set of coefficients derived from projecting 2D motion sequences onto the 3D model were used as a signature of a walker. The calculation of gait signatures used prior information to reduce ambiguity in a Bayesian framework. In [12], a 3D human model was set up from video sequences captured by multiple cameras. The motion trajectories of lower limbs extracted from the 3D model were used as dynamic features, and linear time normalization was exploited for matching and recognition. Bodor et al. [13] employed image-based rendering of a 3D model to adapt the input to meet the needs of the classifier by automatically constructing the proper view (image), which matched the training view, from a combination of arbitrary views taken from several cameras. Ribnick and Papanikolopoulos [14] reconstructed a periodic trajectory in 3D given only its appearance in image coordinates from a single camera view.

Over the past few years, significant progress has been made in terms of the diversity of gait recognition algorithms. Particularly, the techniques relying on the information feature contained in binary silhouettes of walking humans are of much interest since they do not presume the availability of any further information. Gait is a dynamic shape model such that it varies in pose and size throughout a walking cycle [15]. In practice, the gait shape of an individual can be easily altered by many factors, particularly by the change of views. Individuals can walk at various directions to the camera in any real-world situations, and gait shape changes nonlinearly according to views. This will lead to significant changes to walking patterns and generate difficulties for gait recognition. Understanding the above factors is crucial to develop robust and accurate gait recognition algorithms. In this paper, we will focus on proposing a new solution to deal with the effect of view variation on gait recognition no matter the variation is small or significantly large.

In our previous studies, gait dynamics represented by suitable periodic features, such as lower limb joint angles/angular velocities, were approximated by radial basis function (RBF) neural networks (NNs) [16,17]. The difference of gait dynamics between different individuals during walking can be used for gait recognition. Following this idea, we continue to search for periodic gait features that constitute gait dynamics and can reflect the variance between gait patterns across various views. Recent gait research revealed that silhouette cues, such as the width of the outer contour and silhouette area, play a primary role in gait recognition [18–25]. This finding enlightens us to extract silhouette features for gait identification. Since view variation can significantly change the gait shape description, suitable silhouette features representing gait dynamics across views will be found in this paper for gait recognition.

In this paper, we propose a new silhouette-based method for view-invariant gait recognition via deterministic learning theory. The width of the binarized silhouette models the periodic deformation of human gait silhouettes. It captures the spatio-temporal characteristics of each individual, represents the dynamics of gait motion, and sensitively reflects the variance between gait patterns across various views. It will be reliably computed, then be selected as gait feature to perform recognition. The gait recognition approach consists of two phases: a training phase and a recognition phase. In the training phase, gait dynamics underlying different individuals' gaits observed from different view angles are locally accurately approximated by RBF neural networks. The obtained knowledge of approximated gait dynamics is stored in constant RBF networks. In order to address the problem of view change no matter the variation is small or significantly large, the training patterns extracted from different views constitute a uniform training dataset containing all kinds of gait dynamics of each individual observed across various views. In the recognition phase, a bank of dynamical estimators is constructed for all the training gait patterns. Prior knowledge of human gait dynamics represented by the constant RBF networks is embedded in the estimators. By comparing the set of estimators with a test gait pattern whose view pattern contained in the prior training dataset, a set of recognition errors are generated. The average L_1 norms of the errors are taken as the similarity measure between the dynamics of the training gait patterns and the dynamics of the test gait pattern. In contrast to other existing approaches whose gallery and probe views are the same or for cross-view recognition, we construct a uniform training dataset consisting of gait patterns from different views to achieve view-invariant gait recognition.

The rest of the paper is structured as follows. Section 1 introduces preliminary knowledge about deterministic learning theory and problem formulation. Silhouette features extraction and gait signature derivation are presented in Section 2. Training and learning mechanism based on silhouette features are given in Section 3. In Section 4, gait recognition scheme is presented. The experiments of gait recognition are included in Section 5 to verify the effectiveness of our approach. Section 6 contains the conclusion.

1. Preliminaries and problem formulation

1.1. Deterministic learning theory

In deterministic learning theory, identification of system dynamics of general nonlinear systems is achieved according to the following elements: (i) employment of localized RBF networks; (ii) satisfaction of a partial persistence of excitation (PE) condition; (iii) exponential stability of the adaptive system along