Computer Vision and Image Understanding 126 (2014) 38–52

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

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

Gait recognition by fluctuations ☆

Muhammad Rasyid Aqmar *, Yusuke Fujihara, Yasushi Makihara, Yasushi Yagi

Department of Intelligent Media, The Institute of Scientific and Industrial Research, Osaka, Japan

ARTICLE INFO

Article history:
Received 31 August 2013
Accepted 15 May 2014
Available online 28 May 2014

Keywords:
Gait recognition
Temporal fluctuations
Spatial fluctuations
Phase estimation
Score-level fusion
Quality measure

ABSTRACT

This paper describes a method of gait recognition by suppressing and using gait fluctuations. Inconsistent phasing between a matching pair of gait image sequences because of temporal fluctuations degrades the performance of gait recognition. We remove the temporal fluctuations by generating a phase-normalized gait image sequence with equal phase intervals. If inter-period gait fluctuations within a gait image sequence are repeatedly observed for the same subject, they can be regarded as a useful distinguishing gait feature. We extract phase fluctuations as temporal fluctuations as well as gait fluctuation image and trajectory fluctuations as spatial fluctuations. We combine them with the matching score using the phase-normalized image sequence as additional matching scores in the score-level fusion framework or as quality measures in the score-normalization framework. We evaluated the methods in experiments using large-scale publicly available databases and showed the effectiveness of the proposed methods.

1. Introduction

Gait recognition has attracted growing attention in the past decade as a promising biometric cue for identification and verification. A main advantage of gait from a biometric point of view lies in its ability to be captured even at a distance by video camera, which is a desirable property for security, surveillance, and forensic applications.

Many video-based gait recognition approaches have been developed in recent years. The approaches for video-based gait recognition are divided into two categories: model-based [1–4] and appearance-based [5–11] approach. Some recent developments in model-based approach include arm-leg movements extraction [12], the usage of Kinect camera for skeleton tracking [13], and 3D gait model construction [14], while silhouette-based approach include extracting frequency-domain feature from gait entropy image [15] (another variant than that from gait energy image [8]), optical flow from silhouette [16], shape contour and Fourier descriptor [17], and extracting silhouette from infrared thermal camera [18]. Currently, silhouette-based approach is a more popular gait recognition method due to its simplicity and inexpensive computational cost, and it performs well under less-variations condition. We may refer the readers to the book [19] for more wider coverage of gait recognition approaches.

However, gait recognition performance is significantly affected by intra-subject variations such as clothing [20], view [21,22], speed [23,24], elapsed time [25], surface [26], and carrying items [27]. These problems are represented in several public databases that include some of the above covariates. USF database includes the variations in view, shoes, surface, carrying load, and time with low quality silhouette. CASIA database provides view and clothing covariates with mid-level quality silhouette. SOTON database includes footwear, bodywear, carrying load and speed covariates with high quality silhouette.

Recently, a couple of gait recognition research groups tried to handle more realistic challenges, such as uncontrolled outdoor environment and occlusion. DeCann et al. [28] created a gait database recorded outdoor where the database contains illumination variations, background artifacts, multiple walking paths, and spatial resolution differences. Hofmann et al. [29] created a gait database recorded outdoor that include static and dynamic occlusion.

In addition to the above problems, it has been reported that inconsistency in gait between a probe and gallery sequence because of *gait fluctuations* also degrades gait recognition performance [30]. We refer this inconsistency as inter-period gait fluctuations.

Inter-period gait fluctuations fall into two types: (1) temporal fluctuations, and (2) spatial fluctuations. The temporal fluctuations are derived from unstable phase (gait stance) evolution and result in non-uniform phase intervals between adjacent frames, which

☆ This paper has been recommended for acceptance by Yaser Sheikh.

* Corresponding author. Address: Department of Intelligent Media, The Institute of Scientific and Industrial Research, Osaka University, 8-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan.

E-mail address: aqmar@am.sanken.osaka-u.ac.jp (M.R. Aqmar).

http://dx.doi.org/10.1016/j.cviu.2014.05.004
1077-3142/

induces misalignment of the frames even after a frame synchronization process [26]. The spatial fluctuations are derived from the differences in pose (e.g., variations in arm swinging, leg motion, and degree of stooping posture) and vertical positions (e.g., head top and feet bottom) within the same phase.

Temporal fluctuations can usually be solved by non-linear time warping approaches such as dynamic time warping (DTW) and its extension in a probabilistic framework, the hidden Markov model (HMM). Both approaches are based on finding the optimum path of non-linear time-varying patterns between a probe and gallery sequence. However, the accuracy of the optimal path in the DTW and HMM is constrained by the intervals of frames and states, respectively, and hence phase misalignment at the inter-frame or inter-states level cannot be solved by the conventional non-linear time warping approaches.

The first contribution of this study is to suppress such phase misalignment at inter-frame levels by generating a gait image sequence with equal phase intervals (see the left side of Fig. 1), called a phase-normalized image sequence, based on sub-frame order phase estimation using Self-DTW [31] and shape morphing [32].

Looked at from another perspective, if inter-period gait fluctuations within a gait image sequence are repeatedly observed for the same subject, they can be regarded as a type of useful gait feature for distinguishing the subjects from one another. For example, assuming the degree of arm swinging and stooping posture of a subject always fluctuate between periods (see the right side of Fig. 1) and that those of another subject are quite stable between periods, we can distinguish them based on the fluctuations of the degree of arm swinging and stooping posture even if the average degree of arm swinging and stooping posture are similar to each other.

The second contribution of this study is to use such gait fluctuations to improve gait recognition accuracy. More specifically, we extract phase fluctuations as temporal fluctuations, and gait fluctuation images and trajectory fluctuations as spatial fluctuations. We then combine them with a matching score using the phase-normalized image sequence as additional matching scores in the score-level fusion framework or as quality measures in the score-normalization framework (see Fig. 1).

This paper is a significant upgrade of our conference paper [33] and extensions from it are summarized by the following three points:

- Introduction of the phase and trajectory fluctuations in addition to the previously proposed gait fluctuation image [33].
- Using gait fluctuation as a *quality measure* in a score-normalization framework, as well as an additional matching score in a score-level fusion framework in [33].
- Evaluation using large-scale publicly available gait databases with several thousand subjects.

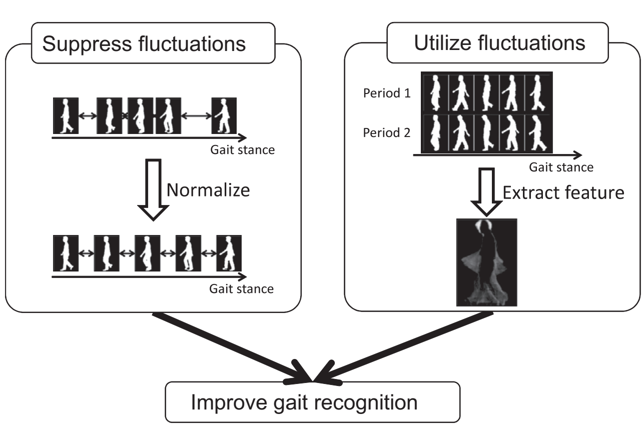

Fig. 1. Gait recognition by fluctuations.

2. Related work

2.1. Period-based gait recognition

Appearance and period-based representation are the current mainstream in gait recognition [19]. The detection of the period was usually based on heel strike information as well as by stance. The averaged silhouette [6], also known as the gait energy image (GEI) [7], represents shape and motion information obtained by averaging size-normalized and registered silhouettes over a gait period. The frequency-domain feature (FDF) [8] is extracted by applying a one-dimensional discrete Fourier transform along the temporal axis to each pixel in the silhouette sequence. The gait entropy image (GEnI) [34,10] emphasizes dynamic regions by measuring the uncertainty of the pixels of the GEI using Shannon entropy. The chrono-gait image (CGI) [11] embeds phase information into the silhouette contour by color encoding. The shape variance-based (SVB) frieze pattern [35] extracts variations between a key gait frame (double-support phase) and subsequent frames. The gait flow image (GFI) [16] is acquired by computing the optical flow based on the silhouette sequence and aggregating the magnitude of the optical flow over a gait period. The gait motion descriptor (GMD) [36] is computed by decomposing the optical flow field into dynamic regions of each of four orientations and static region.

All of these gait features are extracted independently from every gait period[1] of image sequences and hence none of them use the inter-period relationships such as gait fluctuations proposed in this study.

2.2. Non-linear time warping for gait recognition

To handle rate or speed-varying gait image sequences, non-linear time warping techniques such as DTW [37] and its extension in a probabilistic framework, HMM [38], have been widely used in the gait recognition community [5,39–43]. Both techniques require a template (reference) image sequence to be matched with a given test image sequence. The template length is limited by a finite number of frames and states in a DTW and HMM framework, respectively, based on the training data.

Cuntoor et al. [5] proposed DTW-based gait recognition using features derived from the width of the outer contour of the gait silhouette image. Veeraraghavan et al. in [39] used gait shape in a spherical manifold and used the Procrustes distance in their DTW implementation. In [40], Veeraraghavan et al. proposed a robust rate-invariant DTW by estimating the distribution of warping functions from the training sequence of human gait needed to robustly match the test sequence. Sundaresan et al. [41] proposed a generic HMM-based approach and employed several distance metrics to estimate the HMM state output probability. Liu et al. [42] proposed a so-called population HMM (pHMM) trained from the data of many gait subjects to be used as a reference for gait sequence normalization. Aqmar et al. [43] combined shape and motion-based features with a mixture of Gaussian distribution-HMMs to handle speed variations.

Although these non-linear time warping-based methods are typically robust against rate or speed variations, DTW and HMM, in particular, cannot handle phase misalignment at the inter-frame

[1] With the exception of every quarter gait period in the case of CGI.