Pattern Recognition Letters 34 (2013) 663–669

Contents lists available at SciVerse ScienceDirect

Pattern Recognition Letters

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

Gait recognition via optimally interpolated deformable contours

Chin Poo Lee [a,*], Alan W.C. Tan [b], Shing Chiang Tan [a]

[a] Faculty of Information Science and Technology, Multimedia University, Jalan Ayer Keroh Lama, 75450 Bukit Beruang, Melaka, Malaysia
[b] Faculty of Engineering and Technology, Multimedia University, Jalan Ayer Keroh Lama, Bukit Beruang, 75450 Melaka, Malaysia

ARTICLE INFO

Article history:
Received 24 April 2012
Available online 4 February 2013

Communicated by S. Sarkar

Keywords:
Gait recognition
Fourier descriptor
Shape interpolation

ABSTRACT

Gait as a biometric was inspired by the ability to recognize an acquaintance by his manner of walking even when seen at a distance. In this paper, we describe a novel Fourier descriptor based gait recognition method that models the periodic deformation of human contours. A new measure of similarity using the product of Fourier coefficients is proposed as a distance measure between closed curves. In order to maximize the similarity between subsequent closed curves, the assembly of contours in gait cycle is circularly shifted by a circular permutation matrix. Subsequently, an element-wise frame interpolation is correspondingly applied to produce length invariant gait signatures. The experiments on OU-ISIR gait database and CASIA gait database reveal promising recognition accuracy. The element-wise frame interpolation method is able to preserve temporal information even when the gait cycles change, and therefore offers a better robustness to slight variation in walking speed.

1. Introduction

Walking is a bipedal type of locomotion that seems simple, but requires great neural control. The mastering of the erect bipedal type of locomotion appears to be a combination of instinct and learning (Rose and Gamble, 2006). If walking is a learned activity, it is not surprising that each of us displays certain personal peculiarities despite the basic pattern of bipedal locomotion. One can often recognize an acquaintance by his manner of walking even when seen at a distance. Much attention has been devoted to the use of human gait patterns as a biometric and to the analysis of human motion in general. Studying this motion, however, is a difficult task. The aim of computer vision-based analysis of human gait is to automatically discover and describe human gait accurately with minimal human intervention.

Many approaches have been proposed for the description of the human gait. Sundaresan et al. (2003) deployed Hidden Markov Models (HMMs) for recognition of individuals from their gait. Wang et al. (2003a,b) obtained an eigenshape as gait signature using Procrustes shape analysis method. Bobick and Johnson (2001) introduced four static body parameters and then calculated the ratio between the volume of the individual variation density to the overall population by computing the Maximum Likelihood estimate Gaussian density. Zhang et al. (2004) constructed a 2D five-link biped locomotion model to represent human body in the image sequences when the person is walking in lateral view.

Some other studies proposed appearance-based gait features to describe human motion. Liu and Sarkar (2004) engaged an average silhouette method which represents human motion in a single image while preserving temporal information. Han and Bhanu (2006) developed a similar representation called the Gait Energy Image (GEI) by combining the real and synthetic templates for gait recognition. Wang et al. (2010) preserved the temporal information among gait frames via color mapping to generate a chrono-gait image (CGI). Zhang et al. (2010) proposed an active energy image (AEI) method by accumulating image difference between subsequent silhouette images. Lam et al. (2011) generated a gait flow image (GFI) by using the optical flow field from the gait image sequence.

Many Fourier descriptor-based techniques have also been proposed in the literature. A major advantage of Fourier descriptors is that, when representing a shape in the Fourier domain, one can readily access its frequency components. This can be useful, as macroscopic features are found in the lower frequencies, whereas microscopic features are found in the higher frequencies (Mowbray and Nixon, 2003). Mowbray and Nixon (2003) adopted a Fourier series to represent the shape boundary, with the coefficients of the series being the Fourier descriptors of the shape. Tian et al. (2004) described global and local features of shape contour using Fourier descriptors. Dynamic time warping is then applied to align gait sequences of different lengths. This, however, introduces significant computational overhead and, as a result, the approach suffers from heavy computational cost. Lu et al. (2008) represented every gait cycle as four key frames, denoted as differential coefficient, closing up, heel and striding. These frames are then processed with the Fourier transform to obtain the Fourier

* Corresponding author. Tel.: +60 6 2523172.
E-mail addresses: cplee@mmu.edu.my (C.P. Lee), wctan@mmu.edu.my (A.W.C. Tan), sctan@mmu.edu.my (S.C. Tan).

0167-8655/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.patrec.2013.01.013

descriptors. The method, however, tends to lose the temporal deformation information of gait sequences since substantial gait information may be contained in the discarded frames.

Appearance-based approaches are basically region-based in nature while Fourier descriptors are boundary-based approaches. Region-based approaches rely on information in the region of interest, such as texture, or intensity homogeneity. Boundary-based techniques, on the other hand, rely on information provided by the object boundaries or the shape properties (Zhang et al., 2012). In most of the appearance-based methods, transient information is not preserved because all frames in a gait cycle are usually accumulated into a composite frame from which the gait descriptor is determined. In contrast, the boundary-based approach is able to preserve the transient information of a gait cycle where each frame is extracted and coded in a vector. Besides that, boundary-based approaches focus on the shape contour that contains more discriminating features between subjects. To this end, instead of analyzing the whole region of human body, only its boundary is considered.

In this paper, we propose a Fourier descriptor-based gait recognition algorithm. In the proposed approach, we adopt Fourier descriptors to represent deforming human contours in gait cycles. The boundary coordinates are vectorized into one-dimensional complex coordinate system. To further ease processing, every video is divided into segments, each consisting of boundary vectors in half gait cycles. Fourier transform is then applied to the complex boundary vectors in each segment to obtain the Fourier descriptors. We define a new measure of similarity for closed curves inspired by the maximization of the product of Fourier coefficients. This measure of similarity is then applied in shape alignment to minimize the difference across the sequence of deformable contours. We also propose a frame interpolation method to normalize the segments of optimally aligned contours into some desired number of frames. The main contributions of the proposed algorithm are the circular alignment of consecutive contours, and frame interpolation that lends a certain property of invariance to gait cycle length and, in doing so, reduces the overall computational complexity in actual implementation.

2. Definition: gait recognition

To bring into proper perspective the key issues considered in this paper, it is useful to digress for a moment to the following definition of the gait recognition problem.

Let $\mathcal{X}$ denotes the gait signature space that contains all gait signatures and $\mathcal{C}$ be the set of class labels of the known classes. The central task of gait recognition is the assignment of gait signatures to classes with common characteristics, or equivalently, the mapping $\mathcal{X} \to \mathcal{C}$. The gait signature space, however, is usually overly redundant and as a preliminary step to the actual classification task, a feature extractor $\mathcal{F} : \mathcal{X} \to \mathcal{Y}$ is deployed to reduce the dimensionality of the gait signature space. The feature space $\mathcal{Y}$ should essentially contain the underlying motion elements of the gait under consideration. The mapping $\mathcal{G} : \mathcal{Y} \to \mathcal{C}$ completes the recognition process.

3. Gait signature extraction

This section outlines the feature extraction phase for gait recognition in this paper.

3.1. Boundary representation

The silhouette boundary edge is obtained using a boundary tracing algorithm. Let $(x_n, y_n), n = 0, 1, \ldots, N-1$, be the coordinate point of the nth pixel on the boundary, and N be the total number of boundary pixels. These coordinate points are organized into a vector of complex numbers $\mathbf{v} \in \mathbb{C}^N$ where $[\mathbf{v}]_n = x_n + jy_n$.[1] Without the loss of generality, we assume that these points are centered around the centroid, i.e., $\mathbf{1}^T\mathbf{v} = 0$, and the shape boundary vector $\mathbf{v}$ is normalized to unity, i.e., $\|\mathbf{v}\|^2 = 1$.

Selecting an appropriate N is important in characterizing the boundary edge. The larger N is, the more detail of the shape is preserved. However, this also increases computational complexity and introduces noise around the shape boundary. On the other hand, choosing a smaller N compromises on the amount of detail, and possibly leading to poor performance in shape analysis. In practice, N is usually determined empirically; in our experiments, we chose $N = 100$.

3.2. Measure of similarity

Recall that $\mathbf{v} \in \mathbb{C}^N$ is a complex-valued shape boundary vector. For notational simplicity, we assume that $[\mathbf{v}]_{n-kN} = [\mathbf{v}]_n, \forall k \in \mathbb{Z}$. This assumption is intuitively sound as shape boundaries are essentially closed curves. We define the *measure of dissimilarity*[2] $d(\mathbf{v}, \mathbf{u})$ of $\mathbf{v}$ and another shape boundary vector $\mathbf{u}$ as

$$d(\mathbf{v}, \mathbf{u}) = \min_m \|P_m\mathbf{v} - \mathbf{u}\|^2 \tag{1}$$

where P_m is a circular permutation matrix, and $P_m\mathbf{v}$ has the effect of circularly shifting the elements of $\mathbf{v}$ by amount m, i.e., $[P_m\mathbf{v}]_n = [\mathbf{v}]_{n+m}$. The assumption $\|\mathbf{v}\| = \|\mathbf{u}\| = 1$ in the preceding section enables us to rewrite (1) as

$$d(\mathbf{v}, \mathbf{u}) = -\max_m \Re(r_m) \tag{2}$$

where

$$r_m = \sum_{n=0}^{N-1} [\mathbf{v}]_{n+m}[\mathbf{u}]_n^*$$

is the cross-correlation of $\mathbf{v}$ and $\mathbf{u}$. A corresponding *measure of similarity* $s(\mathbf{v}, \mathbf{u})$ may readily be established simply by changing the sign of (2), or

$$s(\mathbf{v}, \mathbf{u}) = -d(\mathbf{v}, \mathbf{u}) = \max_m \Re(r_m) \tag{3}$$

By the properties of the discrete Fourier transform, we recognize the cross-correlation as being equivalent to $\mathcal{F}^{-1}(\mathcal{F}(\mathbf{v}) \odot \mathcal{F}^*(\mathbf{u}))$,[3] where $\mathcal{F}(\cdot)$ denotes the Fourier transform. In this case, Fourier transform of the shape boundary vector $\mathbf{v}$ is given by

$$[\mathcal{F}(\mathbf{v})]_k = \sum_{n=0}^{N-1} [\mathbf{v}]_n e^{-j2\pi kn/N} \tag{4}$$

In the literature of image processing, the Fourier transform of the complex-valued boundary edge is referred to as the Fourier descriptor (Fig. 1).

3.3. Frame interpolation

Frame interpolation is concerned with normalizing the number of frame in the gait cycle when comparing gait cycles of different lengths. We use the ratio of human body height to width when analyzing the period of gait cycle. By identifying two consecutive local minima, we can approximately identify the start frame and end frame of half cycle.

[1] The notation $[\mathbf{v}]_n$ $(n = 0, 1, \ldots, N-1)$ denotes the nth element of the vector $\mathbf{v} \in \mathbb{C}^N$.

[2] The greater $d(\mathbf{v}, \mathbf{u})$ is, the less similar (or more dissimilar) the vectors $\mathbf{v}$ and $\mathbf{u}$ are.

[3] The operation $\mathbf{v} \odot \mathbf{u}$ denotes the element-wise product of vectors $\mathbf{v}$ and $\mathbf{u}$, i.e., $[\mathbf{v} \odot \mathbf{u}]_n = [\mathbf{v}]_n[\mathbf{u}]_n$.