J. Vis. Commun. Image R. 33 (2015) 69–77

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Short Communication

Gait recognition with Transient Binary Patterns ☆

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 Melaka, Malaysia
[b] Faculty of Engineering and Technology, Multimedia University, Jalan Ayer Keroh Lama, 75450 Melaka, Malaysia

ARTICLE INFO

Article history:
Received 2 February 2015
Accepted 9 September 2015
Available online 15 September 2015

Keywords:
Gait recognition
Gait
Human walking
Transient Binary Patterns
Binary Patterns
Texture
Texture descriptor
Histograms

ABSTRACT

In this work, we present a combination of spatiotemporal approach and texture descriptors to extract the temporal patterns in gait cycles. Unlike most conventional methods that focus on spatial information while limiting temporal information captured, spatiotemporal methods preserve both spatial and temporal information. Inspired by the success of texture descriptors in face recognition, the proposed method likewise constructs texture descriptors of gait motion over time. For each gait cycle, the pixel-wise binary patterns along the temporal axis, referred to as the Transient Binary Patterns (TBP), is analyzed. These pixel-wise TBPs are then grouped into regional blocks from which we construct regional TBP histograms. These regional TBP histograms collectively form the global TBP histogram that represents both the distribution of temporal patterns and spatial location. Experimental results clearly show the superiority of the proposed approach over other considered methods.

1. Introduction

When person identification is attempted in natural settings, established biometric modalities that require close contact or cooperative subjects are less applicable. In many crime scene forensics, the situation is aggravated by low resolution video data or by poor illumination. Human walking, on the other hand, is often readily perceivable in the video sequence. Henceforth, gait is the biometric that can address some of these drawbacks [1].

The term 'gait recognition' refers to the automatic extraction of visual cues from a video sequence that characterizes the motion of a walking person for identification purposes [2]. Gait can be beneficial over other biometric modalities in the following ways. Firstly, gait is difficult to disguise. Human walking is a complex action of locomotion involving interaction among different body parts and joints. These variations among the properties of body structures and the motion patterns contribute to the unique cues for identity recognition [2]. Secondly, gait can be utilized for recognition at a distance, when other biometrics, e.g., fingerprints, iris, and etc., are of too low resolution to be perceived. The established biometric modalities such as fingerprint and face usually require sensing the users at close range. Hence, gait is a very attractive modality for recognition in surveillance applications.

Two main theories exist for gait recognition in the computer vision community [3]. The first suggests to recover the three-dimensional structure of the moving subject in video sequence, and subsequently use the structure for recognition; this theory leads to the model-based approaches [2,4–8] to representing gait. Though model-based approaches are view and scale invariant, they are difficult to accurately recover the structures of human body due to the high flexibility and to self occlusion [9,10].

The second theory suggests to use the motion information directly without structure recovery; this leads to the model-free approaches [11–18] for gait representation. Many model-free approaches constructed cumulative energy image from the gait sequences. The higher the energy at the position in the image, the more frequent the motion occurs at the position. Bobick and Davis [19] obtained motion energy image (MEI) and motion history image (MHI) from gait sequences. MEI is an image representing where the motion occurs in an image sequence. MHI, on the other hand, is an image representing the recency of motion. Later, Liu and Sarkar [20] proposed an averaged silhouette approach. They aligned and averaged the silhouettes to describe the normalized accumulative energy in every gait cycle. Like the averaged silhouette approach, Han and Bhanu [21] proposed a gait energy image (GEI) descriptor to denote the normalized cumulative energy image of a gait sequence. Similarly, Zhang et al. [22] constructed an active energy image (AEI) by accumulating the frame difference between two consecutive images. Subsequently, each AEI was projected onto a feature subspace via two-dimensional locality preserving projections (2DLPP) method. Chen et al. [23] constructed

☆ This paper has been recommended for acceptance by Yehoshua Zeevi.

* Corresponding author.
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).

http://dx.doi.org/10.1016/j.jvcir.2015.09.006
1047-3203/

the frame difference energy image (FDEI) of a frame by adding the dominant energy image (DEI) of the corresponding cluster in a gait cycle and the positive portion of the frame difference between consecutive frames. In a more recent development, Roy et al. [24] introduced a Pose Energy Image (PEI), where they averaged the silhouettes of all key poses in a gait cycle. The cumulative energy approaches however represent a complete gait cycle as a single composite image, limiting the transient information captured.

In order to preserve more transient information, other approaches characterize motion via spatiotemporal data volume spanned by the moving person in the image. Niyogi and Adelson [25] suggested that analyzing a spatiotemporal volume of walking people reveals some distinctive patterns. Following that, Liu et al. [26] used Frieze patterns to represent gait along the time axis. Yu et al. [27] applied three dimensional Fourier transform to the spatiotemporal volume to obtain a unique frequency for each person's walking pattern. Lee et al. [28] proposed an efficient spatiotemporal probabilistic gait representation by obtaining the binomial distribution of every pixel in a gait cycle. The advantage of spatiotemporal approach is that it encodes both spatial information in the spatial domain and temporal information along the time axis.

Capitalizing on the advantage of spatiotemporal approach, and inspired by the promising results of local binary patterns (LBP) operator [29] as texture descriptors, a spatiotemporal texture representation is proposed. The LBP operator produces a binary code that describes the local texture pattern in an image by thresholding the intensity values of neighboring pixels by its center pixel. The LBP operator was extended to a Local Binary Patterns from Three Orthogonal Planes (LBP-TOP) operator [30] by calculating the LBP features of each frame and thereafter concatenating the histograms to form a dynamic texture representation. Unlike the LBP and LBP-TOP, the silhouette images are binarized prior to application of the proposed method. The proposed Transient Binary Patterns (TBP) operator represents the motion between frames in a pixel-wise manner along the temporal axis, independent of its other neighboring pixels. These pixel-wise TBP are then grouped into block regions to reflect the regional motion patterns. For each region, a regional TBP histogram is used to characterize the TBP distribution of all the pixels in the region. All the regional TBP histograms are then concatenated into a single global TBP histogram representing the spatiotemporal binary patterns of the gait cycle.

The main contributions of the proposed approach are:

1. The TBP approach describes the temporal motion of gait sequence using binary patterns.
2. The TBP approach captures the spatiotemporal patterns at different levels of granularity: pixel-wise, regional, and global.
3. The TBP approach inherently encodes both spatial and temporal information.

2. Gait signature extraction

This section outlines the gait signature extraction procedures of the proposed TBP method.

2.1. Gait cycle estimation

Assume that binary silhouettes have been extracted from the gait sequences. A preprocessing procedure is then applied on the extracted silhouette sequences. Let $I(x,y,t) \in \{0,1\}$ where $x = 1,2,\ldots,C$, $y = 1,2,\ldots,R$, $t = b, b+1,\ldots,T$ be the pixel at the (x,y) coordinate of the $R \times C$ binary image at time t. Similarly, let $\mathcal{P}_t = \{(x,y) : I(x,y,t) = 1\}$ be the set of foreground pixel coordinate points in the image at time t. Each silhouette is proportionally resized so that all silhouettes have the same desired height y_d, and the aspect ratio of each silhouette is maintained. The silhouette is thereafter centered by aligning its horizontal centroid $\bar{x}_t$ with the desired image center x_d. Formally, this means

$$\begin{aligned}&\forall (x,y) \in \mathcal{P}_t : \\ &y' = y \times \frac{y_d}{\max_y\{\mathcal{P}_t\} - \min_y\{\mathcal{P}_t\}} \\ &x' = x + (x_d - \bar{x}_t)\end{aligned} \tag{1}$$

where

$$\bar{x}_t = \frac{\sum_{(x,y)\in\mathcal{P}_t} x}{|\mathcal{P}_t|} \tag{2}$$

and $|\cdot|$ is the cardinality of the set. Fig. 1 displays some sample images before and after silhouette preprocessing procedure. Considering that regular human walking is a periodic motion, it is necessary to detect the period in the gait sequence to correctly preserve the temporal information. The silhouette width of the image at time t is given by

$$w_t = \max_x\{\mathcal{P}_t\} - \min_x\{\mathcal{P}_t\} \tag{3}$$

Accumulating all the silhouette widths in the gait sequence with length T_0, the time series signal of silhouette width $W = \{w_t, t = 1,2,\ldots,T_0\}$ is constructed. A digital filter is then used

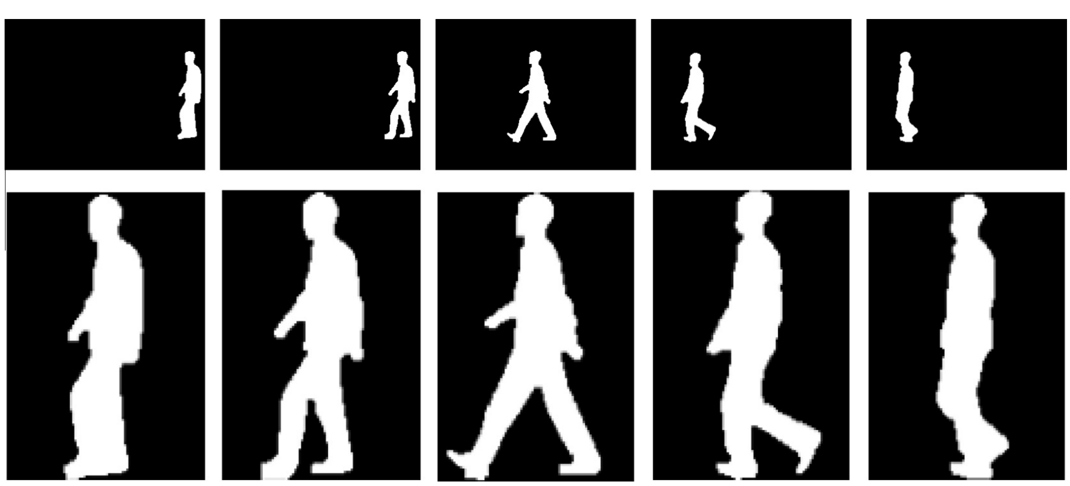

Fig. 1. Top row shows the images before silhouette preprocessing process. Bottom row shows the corresponding images after silhouette preprocessing process.