J. Vis. Commun. Image R. 25 (2014) 822–826

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Time-sliced averaged motion history image for gait recognition

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 3 September 2013
Accepted 17 January 2014
Available online 30 January 2014

Keywords:
Gait
Gait recognition
Gait analysis
Motion
Motion history image
Gait energy image
Histograms of oriented gradients
Histograms

ABSTRACT

In this paper, we propose a time-sliced averaged motion history image (TAMHI) alongside the histograms of oriented gradients (HOG) to generate gait signatures in a gait recognition problem. Building on the motion history image (MHI), TAMHI divides the gait cycle into several regular time windows to generate the same number of TAMHI composite images. HOG descriptors are then calculated on these composite images to obtain the gait signature. The time-slicing procedure to produce multi-composite images preserve more detailed transient information of gait cycles. Additionally, time-normalization also introduces gait length invariancy into the representation, hence, offering a better recognition rate to slight changes in walking speed.

1. Introduction

Most of the computer vision-based human recognition methods, such as iris, fingerprints, face, or palmprint biometric modalities, generally require a cooperative subject, specific views, physical contact or close proximity. Gait, which concerns recognizing individuals by the way they walk, is an alternative biometric without these disadvantages. However, gait also has some limitations. Gait can be affected by clothing, shoes, physical and emotional conditions or environmental context. The gait variation of the same person under different conditions reduces the discriminating power of gait as a biometric, but the inherent gait characteristic of an individual still makes it irreplaceable in visual surveillance applications [1].

Previous work on extracting gait features can be mostly divided into two categories: model-based approaches and appearance-based approaches. Model-based approaches [2–7] aim to explicitly model the human body or motion. These approaches usually measure parameters on the model such as trajectories, limb lengths, and angular speeds. Though model-based approaches are view and scale invariant and reflect the kinematic characteristics of walking manner, they are hard to accurately locate the joints' position due to the highly flexible structure of non-rigid human body and to self-occlusion [8,9]. Therefore, current literature focuses on appearance-based approaches [10–17].

Appearance-based approaches directly operate on the gait sequences without assuming any specific model. Among the appearance-based approaches, there is an effective method which engages gait energy image (GEI) [1,18] as gait signature. The GEI is a time-normalized accumulative energy image of human walking in a complete gait cycle. Yang et al. [19] used variation analysis to construct a dynamics weight mask to enhance the dynamic region in GEI and suppress the noises on the unimportant regions. Chen et al. [20] constructed 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. Zhang et al. [21] proposed an active energy image (AEI) method by accumulating image difference between subsequent silhouette images. Lam et al. [22] generated a gait flow image (GFI) by using the optical flow field from the gait image sequence. These accumulative energy images however cannot indicate in what order the pixel experienced the motion and therefore cannot encapsulate timing patterns in a motion.

Other approaches address the lack of timing patterns problem by indicating how recently motion occurred at each pixel. Bobick and Davis [23] explored two temporal templates: motion energy image (MEI) to indicate the presence of motion in a gait cycle, and a motion history image (MHI) as a function of the recency of motion in a gait cycle. Wang et al. [24,25] preserved the temporal information in a gait cycle via color mapping to generate a chrono-gait image (CGI). These approaches however represent a complete gait cycle as a single composite image, limiting the transient information captured.

* 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.2014.01.012
1047-3203/

Nixon et al. [26] defines a gait cycle as the time interval between successive instances of initial foot-to-floor contact (heel strike) for the same foot. Each leg has two distinct periods: a stance phase, when the foot is in contact with the floor, and a swing phase, when the foot is off the floor moving forward to the next step. Therefore, a comprehensive gait signature should be able to capture these distinct gait components in a complete gait cycle [27].

To this end, we propose a time-sliced averaged motion history image (TAMHI). The TAMHI representation essentially divides the gait cycle into several regular time windows to generate the same number of TAMHI composite images. Unlike the single composite image representation, more detailed temporal deformation information is captured in this multi-composite image representation. In addition, time-normalization is performed on the TAMHI composite image to eliminate the template variations caused by different gait lengths.

For each TAMHI composite image, the histograms of oriented gradients (HOG) are then computed. The HOG representation has several advantages. It captures edge or gradient structure that characterizes the local shape. Since it operates on localized cells, it shows invariance to geometric and photometric transformations [28]. Recent works in action recognition had also adopted these appearance-based methods alongside HOG [29,30]. Early results by these studies confirm the potential of HOG for appearance-based recognition. The TAMHI-HOG descriptors of these time windows in a half cycle collectively form the gait signature of the half cycle. The main contributions of the proposed algorithms are:

1. Time-slicing of the gait cycle to preserve more transient information of the gait motion components.
2. Time-normalization of the cumulative energy image to reduce the effect of diverse gait lengths.
3. Generalization of some state-of-the-art gait recognition approaches, i.e. key frame-based approach [31] and GEI approach.

2. Gait signature extraction

This section outlines the gait signature extraction procedures of the proposed method. The block diagram of the TAMHI-HOG method are shown in Fig. 1.

2.1. Time-slicing of gait cycle

We assume that silhouettes have been extracted from original human walking sequences. A silhouette preprocessing procedure [32] is then applied on the extracted silhouette sequences. Firstly, size normalization is performed where each silhouette image is proportionally resized so that all silhouettes have the same height. Then, horizontal alignment is performed where the silhouettes are centered with respect to its horizontal centroid.

Nixon et al. [26] defines a gait cycle as the time interval between successive instances of initial heel strike for the same foot. Each leg has two distinct periods, namely a stance phase and a swing phase. Considering that regular human walking is a periodical motion, it is necessary to detect the period in the gait sequence to correctly preserve the temporal information. Early studies

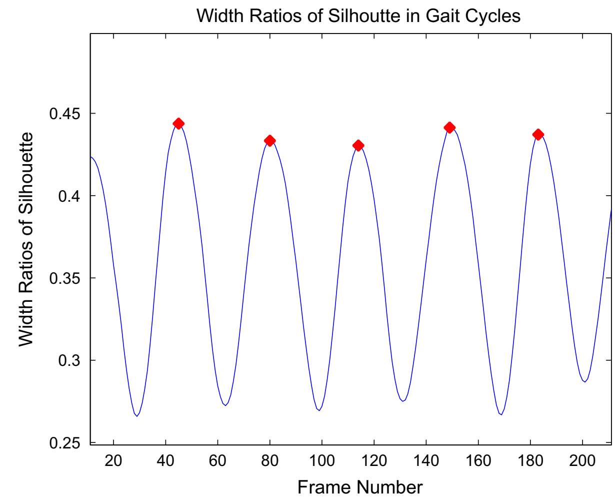

Fig. 2. The width of the silhouette throughout gait cycles (diamonds denote the key frames bounding each half gait cycle. Specifically, each half gait cycle is bounded by two diamonds).

proposed that the width of the silhouettes will have a local maximum when the two legs are farthest apart from each other and reach a local minimum when the two legs wholly overlap. Fig. 2 indicates that the method produces sharp peaks and valleys, and thus preserves the correct temporal order.

Each video is divided into a number of half cycles, where the key frames bounding each half cycle are decided based on the local maximum of the width signal of the silhouette. These half cycles are further time-sliced into L non-overlapping regular time windows.

2.2. Time-sliced averaged motion history image

To calculate the TAMHI composite image for a time window with length n, the following procedure applies. Let $I(x,y,t)$ be an image at time t. The TAMHI composite image H can be computed as:

$$H = \frac{1}{\sum_{t=1}^{n} \alpha^{n-t}} \sum_{t=1}^{n} \alpha^{n-t} I(x,y,t) \tag{1}$$

where α is the decay parameter ($0 \leqslant \alpha \leq 1$). The result of this computation is a scalar-valued composite image that implicitly represents the direction of motion. The brighter the region in TAMHI, the more recent the motion.

Selecting an appropriate α is important in characterizing the motion. The larger α is, the more detail of the gait sequence is preserved. However, this also suppresses the information on the recency and direction of motion. On the other hand, choosing a smaller α compromises on the amount of detail, and possibly leading to loss of earlier trail of the motion sequence.

On the extreme cases, the choice of $\alpha = 0$ will only capture the last image in the time window, producing a key frame type gait signature [31] for each time window. Meanwhile, $\alpha = 1$ will generate a GEI composite image. In this sense, the TAMHI representation

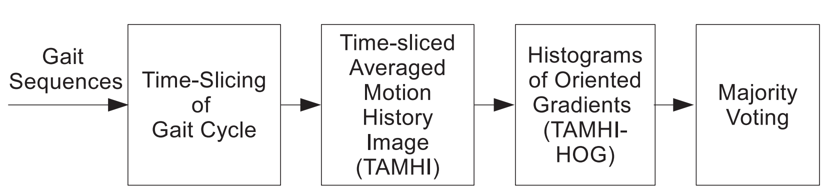

Fig. 1. The system flow of the TAMHI-HOG method.