

Detection Human Motion with Heel Strikes for Surveillance Analysis

Sung-Uk Jung and Mark S. Nixon

ISIS, School of Electronics and Computer Science
University of Southampton, SO17 1BJ, UK
{suj08r,msn}@ecs.soton.ac.uk

Abstract. Heel strike detection is an important cue for human gait recognition and detection in visual surveillance since the heel strike position can be used to derive the gait periodicity, stride and step length. We propose a novel method for heel strike detection using a gait trajectory model, which is robust to occlusion, camera view and to low resolution which can generalize to a variety of surveillance imagery. When a person walks, the movement of the head is conspicuous and sinusoidal. The highest point of the trajectory of the head occurs when the feet cross. Our gait trajectory model is constructed from trajectory data using non-linear optimization. Then, the key frames in which the heel strike takes place are extracted. A Region Of Interest (ROI) is extracted using the silhouette image of the key frame as a filter. Finally, gradient descent is applied to detect maxima which are considered to be the time of the heel strikes. The experimental results show a detection rate of 95% on two databases. The contribution of this research is the first use of the gait trajectory in the heel strike position estimation process and we contend that the approach is a new approach for basic analysis in surveillance imagery.

Keywords: Heel strike detection, gait trajectory model, gradient descent, gait.

1 Introduction

Heel strike detection is a basic and important process in non-invasive analysis of people at a distance, especially, for human motion analysis, and recognition by gait, in a visual surveillance environment. Since the gait periodicity, stride and step length can be calculated directly from the position of the heel strikes; this information can be used to represent the individual characteristics of a human, the walking direction, and the basic 3D position information (given a calibrated camera).

There are two central observations concerning heel strike detection. During the strike phase, the foot of the striking leg stays at the same position for half a gait cycle (when the foot is in contact with the floor), whilst the rest of the human body moves forward [1]. Another is that when the left foot and right feet cross, the head is at its highest position in a gait cycle. We develop our new heel strike detection method based on these observations.

Generally, heel strike detection is a preliminary step in gait recognition or model-based human body analysis and visualization. There are two major previous

approaches to heel strike detection. The first is a model-free approach which uses low level data such as silhouettes and edges to detect gait motion. Bobick and Johnson [2] recovered static body and stride parameters of subjects using the action of walking to extract relative body parameters. BenAbdelkader et al. [3] identified people from low resolution video by estimating the height and stride parameters of their gait. Jean et al. [4] proposed an automatic method of detecting body parts using a human silhouette image. A five point human model was detected and tracked. They solved for self-occlusion of the feet by using optical flow and motion correspondence. Bouchrika and Nixon [1] built the accumulator map of all of corner points using Harris corner detector during an image sequence. Then, the heel strike position was estimated using the density of proximity of the corner points.

An alternative approach is a model-based approach which uses prior information such as 3D shape, position and trajectory of body motion. The heel strike position is sub-result of these researches. Vignola et al. [5] fitted a skeleton model to a silhouette image of person. Each limb (two arms, two legs) was fitted independently to speed-up the fitting process. Zhou et al. [6] extracted full-body motion of walking people from video sequences. They proposed a Bayesian framework to introduce prior knowledge into system for extracting human gait. Sigal and Black [7] estimated human pose using occlusion-sensitive local image likelihoods method. Zhang et al. [8] presented a 3-level hierarchical model for localizing human bodies in still images from arbitrary viewpoints. They handled self-occlusion and large viewpoint changes using Sequential Monte Carlo (SMC) optimization. Sundaresan et al. [9] proposed a graphical model of the human body to segment 3D voxel data of human into different articulated chains.

However, many approaches consider the fronto-parallel view of a walking subject where the subject walks in a direction normal to the camera's plane of view. Also, in the model-based approaches there is much computational load in initialization and tracking. Moreover, in the visual surveillance environment the image quality could be low and the image sequences derived from a single camera only are available. Therefore, an alternative heel strike detection method is needed which is robust to low resolution, foot self-occlusion, camera view point and suitable for single camera.

In this paper, to overcome the above constraints, we propose a novel method of heel strike detection using the gait trajectory model. As mentioned before, the frame in which the heel strike takes place can be extracted through the gait motion even when the foot is hidden by another leg and the image quality is low. In this research, the gait trajectory model [10] is deployed. This model is applied to detect the key frame in which the heel strikes happen. The silhouette image at the key frame is used to remove background data.

The remainder of this paper is organized as follows: Section 2 explains the key frame calculation. Section 3 describes the heel strike candidate detection and verification method using the key frame and gradient descent. Section 4 shows the experimental results based on visual surveillance databases. In Section 5, we conclude our work.

2 Key Frame Calculation

To find the moment of a heel strike the basic characteristics of gait is used. When a person walks the movement of head is conspicuous and sinusoidal [14]. The highest point of a human's trajectory in one gait cycle is the moment when the feet cross. This fact is deployed to implement the heel strike detection system.

2.1 Gait Trajectory Model Construction

The gait trajectory model [10] is described in Eq. (1). The modified version of this model is used and assumes that a subject walks at the same speed. This model only considers the vertical position of the object. The gait trajectory can be divided into two parts: a periodic factor and a scaling factor. The periodic factor is proportional to walking position; the scaling factor depends on imaging geometry. Therefore, the gait trajectory model is defined in the following way, where the vertical position y is a function of gait frequency ω .

First, the general case is

$$y = f(t) * \sin(\omega t + \theta) + g(t) \quad (1)$$

$$\text{periodic factor: } f(x) = C_1 \ln|\alpha(1 - t/\lambda)| + C_2 \quad (2)$$

$$\text{scaling factor: } g(x) = C_3 \ln|\alpha(1 - t/\lambda)| + C_4 + I_{gait} \quad (3)$$

where ω is a gait frequency, θ is the initial phase, I_{gait} is the initial position of gait, λ is a total time, α is a total length of walking, and C_n are constants.

Around the image center, the scaling factor can be modeled as

$$\text{scaling factor: } g_c(t) = C_0 t + I_{gait} \quad (4)$$

where C_0 is constant

To clarify the model the exact human trajectory using some manually chosen points are extracted. We use 34 samples (24 males, 9 female, with around 40 images in each sequence) chosen at random from a gait biometric database [11] and extract the corresponding points for all frames. The extracted gait trajectories are normalized in order to express the error as a percentage. After fitting using the Levenberg-Marquadt algorithm R-squared and Sum of Square Error (SSE) is applied. Table 1 shows the numerical result of model fitting. The value of R-squared for all samples is over 98% and the sum of squared errors is less than 4%, both reflecting a good fit.

Table 1. The numerical data of each model fitting

Measure	Value (%)
SSE	3.87
R-Squared	98.0

2.2 Key Frame Selection

The key frame can be calculated after construction of the model. The highest points of trajectory are calculated from gait frequency f . ($x = ((2n+3/2)\pi - \theta) / 2\pi f$, where n is any integer). Figure 1 shows a sample of the key frame extraction process. In fig. 1(a) the highest position of y is when the left foot and the right foot cross (here, the key frame number is 97). Fig. 1(b) and (c) is the original and silhouette image at the key frame. As shown in fig. 1(b) it indicates that the head is in the highest position is when the feet cross. In the next stage, the silhouette image is used for filtering the accumulator map to extract the ROI since the accumulator map is constructed using all of the images in a sequence and the filtered accumulator map must contain at least one heel strike.

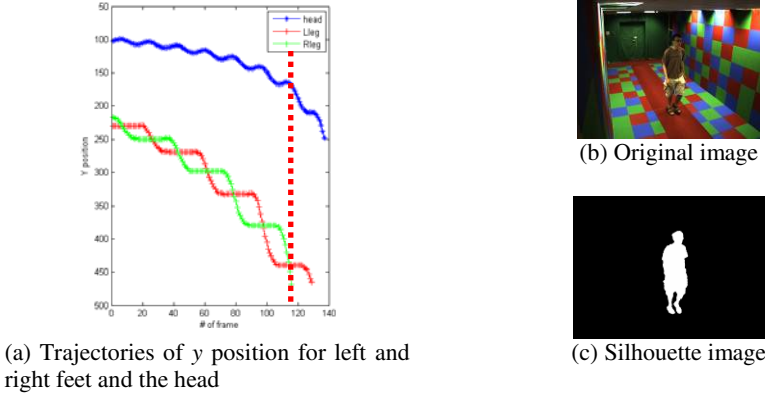


Fig. 1. Key frame extraction

3 Heel Strike Detection

This section shows the process of detecting heel strike position using the pre-calculated key frame information. An accumulator map is used to be derived by adding samples of the walking subject's silhouette to determine which parts of the body remained longest at same position. Generally, during the strike phase, the foot of the striking leg stays at the same position for half a gait cycle, whilst the rest of the human body moves forward.

3.1 Heel Strike Candidate Extraction

As a preprocessing step, we calculate the silhouette image [12] from the intensity and the color difference (between the background image and foreground image) at each pixel. Then, the accumulator map of a silhouette is the number of silhouette pixels in $(i,j)^{th}$ position. Low pass filtering is deployed to smooth the accumulator surface.

$$Accumulator(i, j) = \sum_{\# \text{ of images}} Silhouette(i, j) \quad (5)$$

Figure 2 shows an accumulator map and filtering result of a key frame silhouette image. The colour in the figure indicates the number of silhouette pixels from blue (few) to red (many). As shown in fig. 2(a), the heel strike region can clearly be distinguished from the other body parts.

The filtered accumulator map shows the distribution of the number of silhouette pixels. It reveals that the position of heel strike has a relatively higher distribution than other regions. Using the characteristic, we extract a Region of Interest (ROI) and make it smoother to apply Gradient descent algorithm. Figure 3 shows the process for finding the heel strike positions from the accumulator map. The accumulator map shown in Fig. 2(a) is filtered by Gaussian function (filter size 12×12 , $\sigma = 2.0$). Then, the approximate heel region, which is one eighth of person's silhouette height from the bottom of silhouette, is extracted (fig. 3(a)). Accordingly, the heel strike position can be extracted by gradient descent. Fig. 3(b) shows the three dimensional view of the extracted ROI. Fig. 3(c) shows the result of analyzing Fig. 3(b) using the Gradient Descent algorithm. The small arrow in the figure is the point where the orientation has changed. Fig. 3(c) shows the trace convergence to the local maximum.

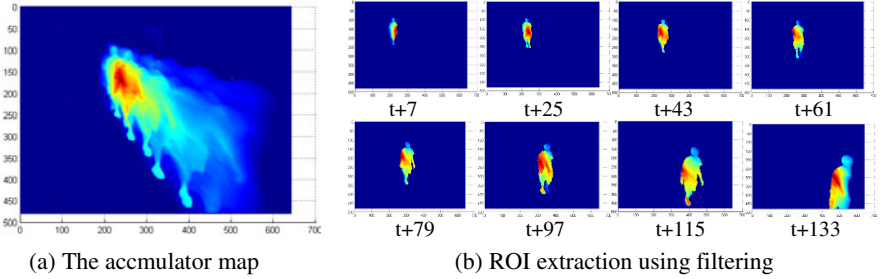


Fig. 2. Silhouette accumulator map

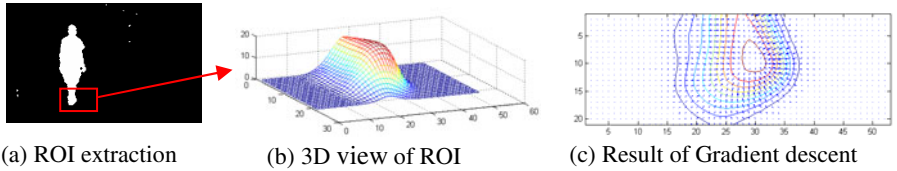


Fig. 3. The procedure of heel strike detection

3.2 Heel Strike Position Verification

In previous section, the process for extracting the heel strike candidates is described. This section describes the heel strike verification process. In our method the silhouette image is used when the feet cross, so it is possible to extract the candidates from foot of heel strike and also another foot. For instance, in the second heel strike of fig. 4(a), the candidates are detected from both feet. To reduce the invalid candidates, the key

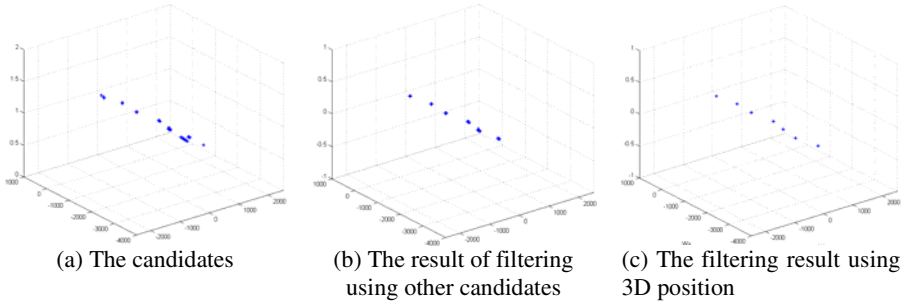


Fig. 4. The verification process

frame is calculated when the position of y is lowest in one gait cycle (fig. 1(a)) and the same procedure is executed in Section 3.1 to find other heel strike candidates. Since the moment at the lowest y is that the gait stride is largest, the feet still stay on the floor, and positions of the feet are separated, so the candidates from other key frames are considered as the potential heel strikes. These candidates are deployed to remove the invalid candidates. Simply, the distance between these two groups of candidates (at the highest and lowest y coordinate) is calculated. Then, the candidates in the fixed distance (here, 5 pixels) are selected from the group of candidates of lowest values for y . As shown in fig. 4(b), after this filtering process the invalid candidates from another foot are removed.

The accumulator map depends on camera view and once the camera is calibrated the invalid candidates could be removed using the back projection from a 2D image plane into a 3D world space. Using the 3D projection the candidates which are the closest from the camera are selected. Since a single camera is used in our approach, we assume a ground floor is known, i.e. that $z=0$ (the z axis is vertical position). This enables calculation of the intersecting points between the projection ray from 2D image points and the ground floor. The closest heel strike to the floor is considered as the final heel strike position, thereby filtering the invalid positions. Figure 4 shows a result of the filtering process and the invalid points in fig. 4(a) are removed to give the final result in fig. 4(c).

4 Experimental Results

To evaluate the proposal heel detection system, we use two visual surveillance databases: a biometric gait database [11] and PETS 2006 database [13]. The biometric tunnel data consists of 25 samples (18 males, 7 female, with around 130 images in each sequence) and each sample has two views of image sequences which are different from the data used to verify the gait trajectory model in Section 2. Moreover, we choose 10 samples from an image sequence of PETS2006 dataset which is recorded at a train station. In the dataset, each sample has a different walking direction with around 80 images in each sample sequence.

Figure 5(a) and (b) show the detection result with different environments; the biometric tunnel and a train station from the PETS data. The white crosses in the figure represent the points detected as the heel strike positions.

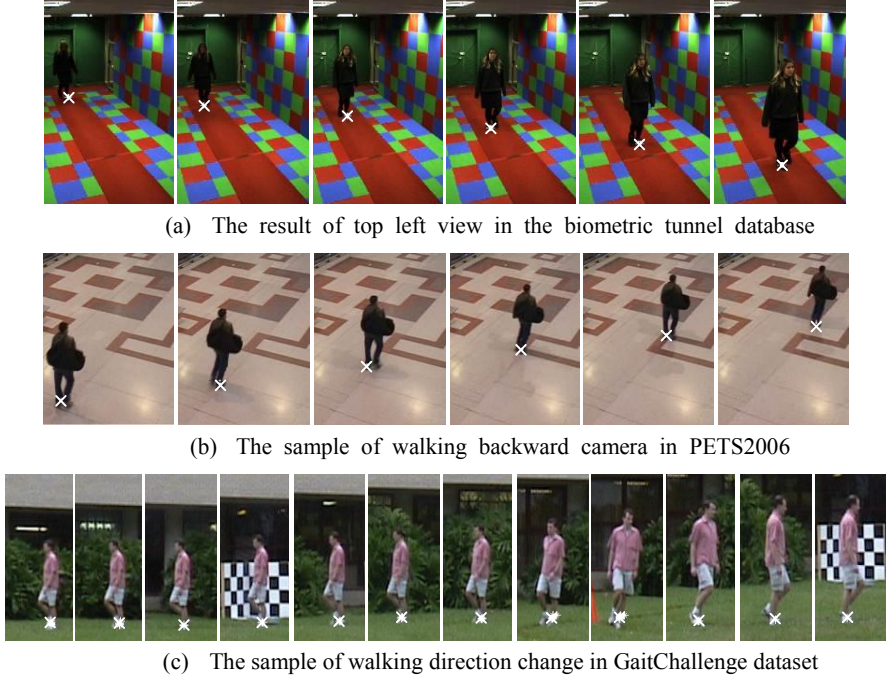


Fig. 5. Detection result with different environments

As shown in fig. 5 the proposed method can detect the heel strike position regardless the camera view since the method uses the basic characteristic of gait: its periodic factor. In the cases of above samples the subjects walk in series of straight lines. To confirm that our method works in the case of walking direction change a single image sequence is tested in Gait Challenge database [15] where the subject is walking around an elliptical path. In this situation, Equation 1 is not suited to finding the gait frequency. So, the gait frequency in the frequency domain is calculated since the walking speed is almost constant. Then, the same procedure is applied. Figure 5(c) shows the result of detection. Even the walking direction changes the proposed method still follows the position of heel strike.

Table 2 shows the result of detection rate. A total of 359 heel strikes were tested. We calculate the detection rate manually because the database does not provide ground truth of heel strike position. The detection rate for the biometric tunnel database is slightly higher than for the PETS 2006 database since the environment of the biometric tunnel is more controlled. The overall detection rate is 95.3%.

Table 2. The result of heel strike detection

Database	Value (%)
Biometric tunnel	95.6 (285/298)
PETS 2006	93.4 (57/61)

5 Conclusions

To deploy automatic gait recognition in unconstrained environments, we need to develop new techniques for analysis. This paper describes new techniques for heel strike estimation to be robust to feet self-occlusion and view of camera. The approach to heel strike estimation combines human walking analysis with characteristics of heel strike. The approach has been demonstrated on a visual surveillance database and on one for biometrics with a heel strike detection rate was over 95%. As such heel strike analysis can be used for basic gait analysis and derivation of walking direction estimation and this approach provides a new and more generalized approach for surveillance environments.

References

1. Bouchrika, I., Nixon, M.S.: Model-based feature extraction for gait analysis and recognition. In: Gagalowicz, A., Philips, W. (eds.) MIRAGE 2007. LNCS, vol. 4418, pp. 150–160. Springer, Heidelberg (2007)
2. Bobick, A.F., Johnson, A.Y.: Gait recognition using static, activity-specific parameters. In: Proc. CVPR, pp. 423–430 (2001)
3. BenAbdelkader, C., Cutler, R., Davis, L.: View-invariant estimation of height and stride for gait recognition. In: Tistarelli, M., Bigun, J., Jain, A.K. (eds.) ECCV 2002. LNCS, vol. 2359, pp. 155–167. Springer, Heidelberg (2002)
4. Jean, F., Bergevin, R., Albu, A.B.: Body tracking in human walk from monocular video sequences. In: Proc. of CRV, pp. 144–151 (2005)
5. Vignola, J., Lalonde, J.F., Bergevin, R.: Progressive human skeleton fitting. In: Proc. of ICVI, pp. 35–42 (2003)
6. Zhou, Z., Prugel-Bennett, A., Damper, R.I.: A Bayesian framework for extracting human gait using strong prior knowledge. IEEE TPAMI 28(11), 1738–1752 (2006)
7. Sigal, L., Black, M.J.: Measure locally, reason globally: occlusion-sensitive articulated pose estimation. In: Proc. CVPR, pp. 2041–2048 (2006)
8. Zhang, J., Luo, J., Collins, R., Liu, Y.: Body localization in still images using hierarchical models and hybrid search. In: Proc. CVPR, pp. 1536–1543 (2006)
9. Sundaresan, A., Chellappa, R.: Model-driven segmentation of articulating humans in laplacianeigenspace. IEEE TPAMI 30(10), 1771–1785 (2008)
10. Jung, S.U., Nixon, S.M.: On using gait biometrics to enhance face pose estimation. In: Proc. of IEEE BTAS, p. 6 (2010)
11. Seely, R.D., Samangoeei, S., Middleton, L., Carter, J.N., Nixon, M.S.: The university of southampton multi-biometric tunnel and introducing a novel 3D gait dataset. In: Proc. of IEEE BTAS, p. 6 (2008)
12. Cheung, G., Kanade, T., Bouquet, J., Holler, M.: A real time system for robust 3d voxel reconstruction of human motions. In: Proc. CVPR, pp. 714–720 (2000)
13. PETS: Performance Evaluation of Tracking and Surveillance, <https://www.cvg.cs.rdg.ac.uk/slides/pets.html>
14. Aristotle: On the Motion of Animals, B.C. 350
15. Phillips, P.J., Sarkar, S., Robledo, I., Grother, P., Bowyer, K.W.: The gait identification challenge problem: data sets and baseline algorithm. In: Proc. ICPR, pp. 385–388 (2002)