34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Markerless Identification of Key Events in Gait Cycle Using Image Flow

Nalini Vishnoi[1], Zoran Duric[2] and Naomi Lynn Gerber[3]

Abstract— **Gait analysis has been an interesting area of research for several decades. In this paper, we propose image-flow-based methods to compute the motion and velocities of different body segments automatically, using a single inexpensive video camera. We then identify and extract different events of the gait cycle (double-support, mid-swing, toe-off and heel-strike) from video images. Experiments were conducted in which four walking subjects were captured from the sagittal plane. Automatic segmentation was performed to isolate the moving body from the background. The head excursion and the shank motion were then computed to identify the key frames corresponding to different events in the gait cycle. Our approach does not require calibrated cameras or special markers to capture movement. We have also compared our method with the Optotrak 3D motion capture system and found our results in good agreement with the Optotrak results. The development of our method has potential use in the markerless and unencumbered video capture of human locomotion. Monitoring gait in homes and communities provides a useful application for the aged and the disabled. Our method could potentially be used as an assessment tool to determine gait symmetry or to establish the normal gait pattern of an individual.**

I. INTRODUCTION

Gait is defined as an individual's characteristic way of moving on foot and is described in terms of three planes of movement: sagittal plane, frontal plane and transverse plane [1]. It is considered one of the most important measures of a person's functional independence. A gait cycle is measured from the heel-strike of one foot to the next heel-strike of the same foot [2]. It is composed of several phases and events which together account for approximately 40% swing and 60% stance [3]. Biomechanics studies performed by [2], [4], [5] used 3D markers to study the limb movement pattern during normal gait for men, women and the elderly. The primary purpose of these studies was to establish a baseline of a normal gait. The kinematic variables (velocities, accelerations and joint angles) were extracted from the recorded 3D data. However, this quantitative study of gait parameters required setting up laboratories with expensive equipment. Cumbersome markers were placed on the subject's limb segments and were tracked using 3D motion capture systems (optical or magnetic trackers). This approach has limited the ability to obtain reliable data in non-laboratory settings.

Recently, the focus has been to develop markerless vision based motion capture systems to analyze human movement for different applications e.g visual surveillance, clinical analysis, computer animation/games, robotics and biometrics. A review of the early work on approaches in understanding human motion and related applications can be found in [6]. In biometrics, human gait has also been used for gender and ethnicity recognition [7]. Since the majority of the movement in the knee, ankle and hip occurs in the sagittal plane, markerless gait analysis is generally 2D oriented. Images have also been used for measuring body sway and identifying footfalls from the gait data [8]. In [9] video data was used to study foot-ankle complex during stance. Few studies [8], [10] have focussed on methods based on the estimation of Center of Mass (CoM) of the body. The major drawback is that the computation of CoM is very sensitive to the accuracy of the segmentation of moving body from the background.

We propose a new method based on image-flow to compute the velocities of different body segments and to identify the corresponding key image-frames. We base our method on the fact that human gait is highly periodic, constrained by human anatomy and reproducible. There are defined phases of gait whose sequence describes the temporal events leading to locomotion. While walking, an individual's head follows a repeatable sinusoidal pattern (characteristic of moving CoM) [15]. This may be altered by stride length, degree of joint range of motion and other limb segment characteristics. The head/body excursion is maximum at the mid-swing and minimum at double-support. These invariant transition points can be detected by extracting the change in the direction of head velocity. We also use the anthropomorphic measures from [11] to consider the lower limb segments and apply image motion models to find the change in the velocity direction corresponding to toe-off and heel-strike events. Our method handles the noise in the segmentation process since we combine the motion of the person together with the segmented foreground. In the end, we compare a gait sequence with the Optotrak motion capture system to establish the validity of our methods.

Section II describes the methodology for finding different gait events. Section III presents results and validation experiment. Section IV provides the conclusion and future work.

II. METHODOLOGY

We first extract the image frames from the video data. The background is then subtracted from each image to obtain the foreground [14]. Morphological operations (erosion and dilation) are performed on the segmented image to eliminate noise. The image flow is then computed from the segmented foreground using the algorithm described in [12]. Image flow

[1]Nalini Vishnoi is in the Department of Computer Science, George Mason University, Fairfax, USA, nvishnoi at cs.gmu.edu

[2]Zoran Duric is with the Faculty of Computer Science, George Mason University, Fairfax, USA, zduric at cs.gmu.edu

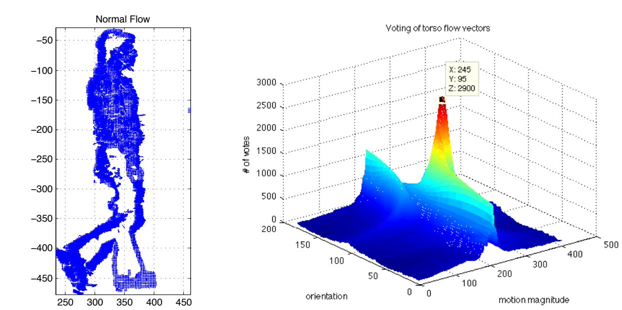
[3]Naomi Lynn Gerber is with the Faculty of the College of Health and Human Services, George Mason University, Fairfax, USA, ngerber1 at gmu.edu

is the instantaneous velocity vector field for an image of a moving environment. It corresponds to apparent image motion between a pair of frames. An example of estimated normal flow field is shown in Fig. 1a. Flow vectors in the background are all close to zero, and non-zero value corresponds to the estimation noise.

A. *Double-support and mid-swing extraction*

It is known that the height of a human torso approximately corresponds to 0.52H, where H is the body height [11]. During gait, the instantaneous head and torso velocity can be approximated, with a high degree of accuracy, by pure translational velocity. This motion comprises forward translation of the body, t_x, and up/down movement/excursion, t_y [15]. We aim to find this motion, $M_t = (t_x, t_y, 0)$, by virtue of "voting" over a range of possible motion values using image flow vectors.

We use the top 52% of the whole body for voting. It counts the number of flow vectors in agreement with the estimated motion corresponding to a bin. This is done for a number of bins (representing possible motions, bin size was 1 degree by 0.1 pixels) and the bin having the maximum number of flow vectors gives the instantaneous translational velocity of the torso. Note that double-support and mid-swing events occur at the zero value of the instantaneous velocity, i.e. the minimum excursion occurs when the velocity changes sign from negative to positive, and the maximum excursion occurs as the velocity changes sign from positive to negative. An example of voting technique is shown in the Fig. 1b. The red peak represents the bin with the highest number of votes.



(a) Normal flow (b) Voting

Fig. 1: (a) Normal flow estimated for a detected foreground. (b) Voting using flow vectors in the top 52% of body height. The maximum is shown in red.

Fig. 2 shows the vertical instantaneous velocity profile of a subject's torso for a sequence of 90 frames. The zero crossings are computed automatically to identify the consecutive double-support and mid-swing phases of the gait cycle.

B. *Toe-off and heel-strike extraction*

Toe-off and heel-strike events can be identified by observing the instantaneous rotational velocity of the lower limb segments (shanks). These events occur at zero values of shank's rotational velocity. For finding these events, we consider the lower 28.5% of the body height H as suggested

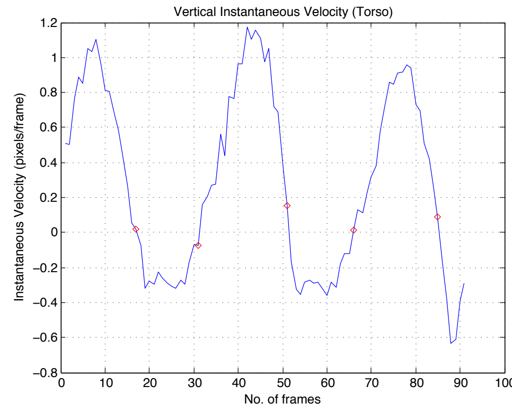


Fig. 2: Instantaneous up and down velocities computed for the torso using voting method. The zero crossings have been marked by red diamonds.

in [11]. We use a slightly modified version of RANSAC algorithm [13] to find motion models for the lower limb segments in each frame. We first divide the foreground into small, equally spaced overlapping patches to reduce the overhead of dealing with each flow vector thereby saving time and reducing complexity. The motion of the lower limb segments can be described by translation in the sagittal plane, (t_x, t_y), and rotation, ω, around an axis orthogonal to the sagittal plane. The model can be approximated as [12],

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 0 & -\omega \\ \omega & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \qquad (1)$$

where $(\dot{x}, \dot{y})$ denote the instantaneous velocity of (x, y). We compute a motion model, M_{tr}, for each patch using Eq. (1) and depending on its error being less than a threshold value, the patch is marked as good or bad. We automatically initialize the left and right lower limb boundaries in the frame close to double support by using their spatial separation and compute their motion.

We then predict the new position of each pixel in the next frame (based on the motion models) and estimate new boundaries. To tackle the noise introduced by noisy measurements of flow and motion estimation, the boundaries are adjusted by applying Principal Component Analysis (PCA) [14] on the pixels marked as inliers. PCA estimates the orientation of the leg segments using those pixels. These boundaries are then used to find new models using the approach given in Algorithm 1 . Once the models are found, the whole process is repeated for the next frame. Our approach allows us to segment and classify the lower body pixels belonging to either the left or the right limb segment (see Fig. 3b and 3c). This computation gives us the estimation of instantaneous rotational velocities of lower limb segments for each frame of the image sequence.

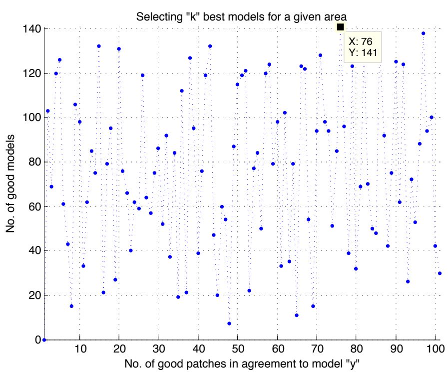
Fig. 4 shows the instantaneous rotational velocity profiles of both left and right lower limb segments of a subject for a single sequence. The zero crossings are computed automatically and correspond to the toe-off and heel-strike events of the gait cycle.

Algorithm 1 RANSAC for estimating motion models

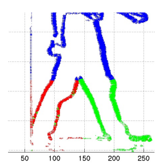
INPUT: Boundary B of current region, Patches P_i, and their respective motion models m_i, $i = 1, 2......n$, n is the number of total patches.

OUTPUT: Model M, best describing the motion of all patches within B.

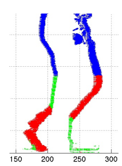
repeat
  Randomly sample two patches p_i and p_j (within B) for which the models m_i and m_j have error less than a threshold value, t.
  Compute a model m_c for the combined patch p_c.
  if m_c has error value $\leq t$ **then**
    Apply m_c to all patches inside B and compute the error for each patch.
    Count the total number of patches that have error $\leq t$.
  end if
until number of iterations $\leq N$
Find the model, M that has the largest number of votes/patches (see Fig. 3a).
return M



(a) Selecting Model



(b)



(c)

Fig. 3: Figures illustrating the working of RANSAC (a) Selection of a model which is in agreement with majority of patches. (b) & (c) Two images showing the segmentation done using RANSAC at different stages in gait cycle. The first model is shown in green and the second model is shown in red.

III. RESULTS

We have captured gait data from four subjects, $S1$-$S4$, with different types of shoes and clothing and processed the data to detect different events of the gait cycle. The data were recorded in the sagittal plane at 60 frames/second at resolution 640×480 pixels per frame using Point Grey Dragonfly®2 color camera. Each subject walked for 18 feet before being recorded to enable them to reach a consistent walking pattern. Fig. 5 shows the image frames from sequences $S1$ and $S3$, extracted by our algorithm, corresponding to double-support, mid-swing, toe-off and heel-strike events of the gait cycle. By visual inspection of the extracted frames, we can say that our results are reasonably

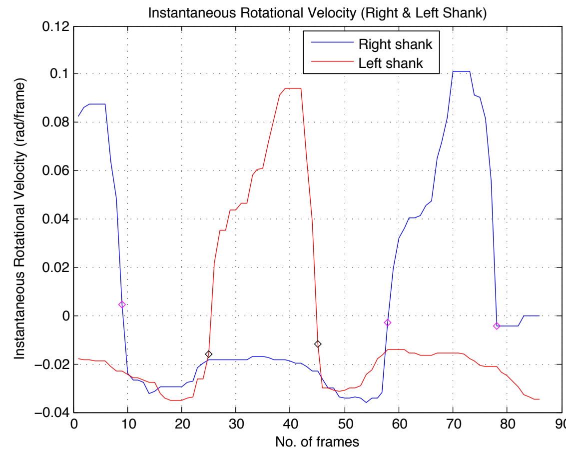


Fig. 4: Instantaneous rotational velocities computed for the two lower limb segments using RANSAC method. The zero crossings have been marked using black and magenta diamonds.

accurate. We have presented results from sequences $S1$ and $S3$ to establish the robustness of our method under varying conditions such as gender, clothing (trousers vs skirt), and shoe type (sneakers vs high heel sandals). We further quantify the accuracy by comparing our results against a well established 3D marker-based motion capture system.

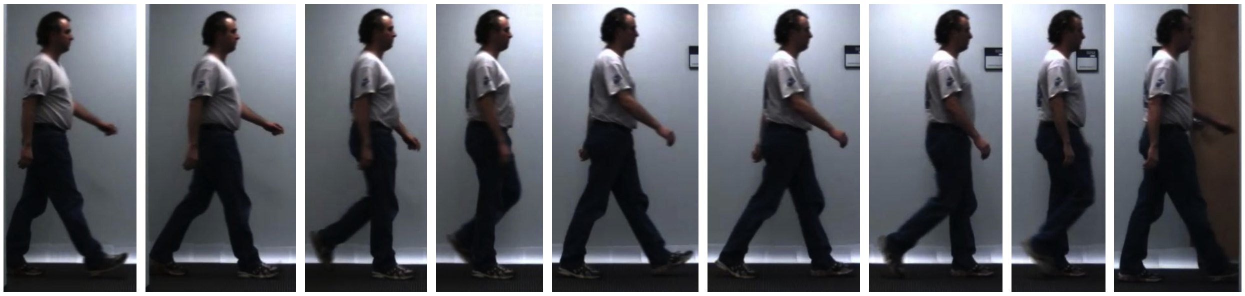
A. Validation

To establish the validity of our method, we have compared our results with the Optotrak motion capture system. We captured a subject's gait data sequences using video camera (sequence S_c) and Optotrak markers (sequence S_o) simultaneously. The markers were fixed on the subject's head and lower limb segments. The sampling rate for the Optotrak system was set as $100Hz$ in the experiment. Different gait events were found using our technique and Optotrak data. Tables I and II compare the time of occurrences of all the identified gait events. The numbers reported in the tables are in seconds where DS stands for double-support, MS for mid-swing, TO for toe-off and HS for heel-strike. We have used linear interpolation to find exact time occurrences of zero instantaneous velocities for various events.

TABLE I: Table comparing the time of double-support and mid-swing occurrences (Optotrak, S_o, vs Video camera, S_c)

	DS1	MS1	DS2	MS2
Optotrak	4.254s	4.540s	4.835s	5.125s
Camera	4.261s	4.526s	4.883s	5.122s
Abs. Diff	0.007s	0.014s	0.048s	0.003s

The average error computed over all the gait events is $\mathbf{0.0164 \pm 0.0196}$ seconds. Taking into consideration the frame rate of the video camera used, the error values are very small showing that our results are in agreement with the Optotrak data. This reinforces our approach of using image-flow-based technique as a markerless system for obtaining the key events of gait cycle accurately and reliably.



(a)

(b)

Fig. 5: Frames corresponding to zero crossings of vertical excursion and lower-limb rotational velocity in sequences S1 and S3. From left to right: heel-strike right, double-support, toe-off left, mid-swing, heel-strike left, double-support, toe-off right, mid-swing and heel-strike right.

TABLE II: Table comparing the time of toe-off and heel-strike occurrences (Optotrak, S_o, vs Video camera, S_c)

	TO1	HS1	TO2	HS2
Optotrak	4.989s	5.317s	4.751s	5.559s
Camera	4.942s	5.311s	4.753s	5.563s
Abs. Diff	0.047s	0.006s	0.002s	0.004s

IV. CONCLUSION AND FUTURE WORK

We have presented a new markerless image-flow-based technique for identifying different events of a gait cycle (double-support, mid-swing, toe-off and heel-strike). We have captured videos of four walking subjects under different conditions and were able to extract the key frames corresponding to various gait events. We have also compared our results with a marker-based motion capture system and found that our results were in agreement with velocities derived from motion capture data within reasonably high accuracy. In the future, we plan to work on computing joint angles and motions of body parts in videos obtained from different viewing directions. We also plan to extend our methods to the analysis of anomalous gait patterns. An inexpensive and universal markerless framework for human gait analysis will be very useful in applications such as tele-rehabilitation, monitoring and surveillance.

ACKNOWLEDGMENT

This material is based upon work supported by the National Science Foundation under Grant No. CNS-0722575.

REFERENCES

[1] M. Nordin and V. H. Frankel, Biomechanics of the knee, Basic Biomechanics of the Musculoskeletal System, pp. 115-134, 1989.

[2] M. P. Murray, A. B. Drought, and R. C.Kory, Walking patterns of normal men, Journal of Bone and Joint Surgery, vol. 46, no. 2, pp. 335-360, 1964.

[3] G. J. Sammarco, Biomechanics of the foot, Basic Biomechanics of the Musculoskeletal System, pp. 163-181, 1989.

[4] M. P. Murray, R. Kory, and B. Clarkson, Walking patterns in healthy old men, Journal of Gerontol, pp. 169-178, 1969.

[5] M. Murray, R. Kory, and S. Sepic, Walking patterns of normal women, Arch. Phys. Med Rehabil., pp. 637-650, 1970.

[6] D. M. Gavrila, The visual analysis of human movement: A survey, Computer Vision and Image Understanding: CVIU, vol. 73, no. 1, pp. 82-98, 1999.

[7] R. C. Mark S. Nixon, Tieniu N. Tan, Human Identification Based On Gait. Springer, 2006.

[8] E. E Stone and M. Skubic, "Evaluation of an Inexpensive Depth Camera for Passive In-Home Fall Risk Assessment", Proceedings, Pervasive Health Conference, Ireland, 2011, pp 71-77.

[9] E. Surer, and A. Cereatti, and E. Grosso, and U. D. Croce. (2011) A Markerless estimation of the anklefoot complex 2D kinematics during stance. Gait and Posture, Vol. 33 (4), p. 532-537.

[10] J. H. Yoo, and M. S.Nixon. (2003) Markerless Human Gait Analysis via Image Sequences. At International Society of Biomechanics XIXth Congress, Dunedin, New Zealand.

[11] D. A. Winter. (1990) The Biomechanics and Motor Control of Human Movement. John Wiley and Sons.

[12] H. Wechsler, Z. Duric, F. Li, and V. Cherkassky. "Motion Estimation Using Statistical Learning Theory". IEEE Transactions on Pattern Analysis and Machine Intelligence, 26:466-478, 2004.

[13] M. A. Fischler and R. C. Bolles (June 1981). "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography". Comm. of the ACM 24 (6): 381-395.

[14] M. Sonka, V. Hlavec, and R. Boyle. Image Processing, Analysis, and Machine Vision. Thomson.

[15] J. Perry. Gait Analysis: Normal and Pathological Function.