# Motion based Markerless Gait Analysis using Standard Events of Gait and Ensemble Kalman Filtering

Nalini Vishnoi[1], Anish Mitra[2], Zoran Duric[3] and Naomi Lynn Gerber[4]

***Abstract*— We present a novel approach to gait analysis using ensemble Kalman filtering which permits markerless determination of segmental movement. We use image flow analysis to reliably compute temporal and kinematic measures including the translational velocity of the torso and rotational velocities of the lower leg segments. Detecting the instances where velocity changes direction also determines the standard events of a gait cycle (double-support, toe-off, mid-swing and heel-strike). In order to determine the kinematics of lower limbs, we model the synergies between the lower limb motions (thigh-shank, shank-foot) by building a nonlinear dynamical system using CMUs 3D motion capture database [1]. This information is fed into the ensemble Kalman Filter framework to estimate the unobserved limb (upper leg and foot) motion from the measured lower leg rotational velocity. Our approach does not require calibrated cameras or special markers to capture movement. We have tested our method on different gait sequences collected from the sagttal plane and presented the estimated kinematics overlaid on the original image frames. We have also validated our approach by manually labeling the videos and comparing our results against them.**

## I. INTRODUCTION

Gait analysis has been an essential part of rehabilitation science and clinical practice, helping to establish evaluation and treatment planning for people with disability. Studies using 3D markers to establish normal gait parameters [6], [7] have helped distinguish normal and abnormal gait using data collected from limb kinetics and kinematics. The kinematic variables (velocities, accelerations and joint angles) were extracted from the recorded 3D data. Limitations for the use of these methods for gait analysis include expense, need of capture in lab settings and use of markers that may restrict or inhibit movement. Recently, the focus has been on developing 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. Since most of the movement in the knee, ankle and hip occurs in the sagittal plane, markerless gait analysis is generally 2D oriented. In [15] Center of Mass (CoM) of the body was used to determine the phases of the gait cycle; this is very sensitive to the accuracy of the extracted silhouettes of the moving body. Recent work [3] focusses on the spatiotemporal segmentation of moving body parts and using snakes [9] in order to trace the outline of the body.

We will be focussing mostly on the lower limb segments of the human body. We take advantage of highly repetitive normal gait and model the relationship between upper leg-lower leg and lower-leg-foot motions using non linear models and use Kalman filters to estimate the upper leg and foot velocities given the lower leg data. Kalman filter has been used previously for constant velocity constraint [8]. To the best of our knowledge, the research described in this paper is the first to model a relationship between two segments and use this non linear dynamic model with Kalman filter to estimate limb motion. We extend the idea of gait events (double-support, mid-swing, toe-off and heel-strike) introduced in [11]. We segment lower legs and apply image motion models to find the change in the velocity direction corresponding to toe-off and heel-strike events. We build motion models from a relatively small 3D data set (attributed to the restricted range of motion for normal gait). We then use Unscented Kalman filter to estimate the upper leg and foot motion.

Section II describes the methodology for finding different gait events and estimating the kinematics of lower limb segments. Section III presents results and validation. Section IV provides the conclusion and discussion.

## II. METHODOLOGY

We first extract the image frames from the video data. The silhouette is extracted using background subtraction and normal flow [13] is computed from pairs of image frames (see Fig. 1a). Using the upper $30\%$ of the body height, we estimate the instantaneous up-down and forward torso velocity using the method in [11]. The zero crossings in the torso up-down velocity correspond to double support and mid swing events of the gait cycle. Toe-off and heel-strike events can be identified by observing the instantaneous rotational velocity of the lower leg segments (shanks). These events occur at zero values of the shank's rotational velocity. For finding these events, we consider the lower 28.5% of the body height H as suggested in [14]. Once we have obtained the keyframes corresponding to double-support and mid-swing, we divide the whole gait cycle into two different sets depending on their distance from the keyframes. A frame closer to mid-swing gets assigned to a different set than the frame closer to double support. We use spatial and motion segmentation in these two sets to compute the lower leg rotational velocities. Frames that are closer to double support are segmented using spatial information whereas the frames

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

[2]Anish Mitra is in the Department of Electrical and Computer Engineering, George Mason University, Fairfax, USA, amitra1 at gmu.edu

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

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

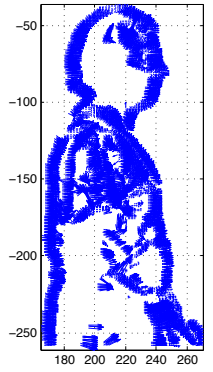
978-1-4244-7929-0/14/$26.00 ©2014 IEEE

that are closer to mid swing are segmented using motion information. In the process, we also introduce histograms of normal flow.

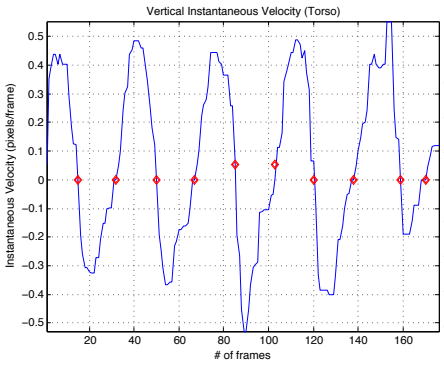
We formulate a nonlinear model to capture the relationship between upper leg - lower leg and lower leg - foot motions. The model is used together with the Unscented Kalman filter to estimate the unknown upper leg and foot motion based on the lower leg rotational velocity.

### A. Computing lower leg velocities from images

We use the approach described in [11] to compute the instantaneous translational velocity of the torso. Fig. 1a shows an example of the torso normal flow and Fig. 1b shows the instantaneous vertical (up-down) velocity profile of the subject's torso. The zero crossings are computed automatically to identify the consecutive double-support and mid-swing phases of the gait cycle. The instances of torso velocity crossing from -ve to +ve value correspond to double support events whereas the frames where velocity changes from +ve to -ve determine mid-swing events.



(a) Normal flow




(b) VerticalExcursion

Fig. 1: (a) An example showing normal flow computed using [13] for the torso. (b) Instantaneous up and down velocities computed for the torso The zero crossings are marked by red diamonds.

*1) Spatial Segmentation:* For frames that are closer to double support, the lower legs are not occluded from one another. This makes the segmentation process easier. We first project all the pixels belonging to the lower $28.5\%$ of body height on the horizontal axis and find the separation between two limbs based on the local minimum between two local maxima in the projected number of pixels. We then label the pixels on different sides of the dividing point as belonging to two different lower leg segments. Fig. 2a illustrates the segmentation.

*2) Motion Segmentation using Histograms of Normal Flow:* We compare the motion of the lower leg segments with the torso motion in frames closer to mid swing. In these frames, *the swinging limb has a velocity greater than the torso velocity whereas the stance (standing) limb's velocity is lower compared to the velocity of the torso*. The lower legs in these selected frames are occluded by different amounts, therefore we use superpixels: first to compute a coarse motion model and then use it to refine the model at pixel level eliminating the noise and interference with the correct motion estimation. Hence, we divide the region in overlapping superpixels of size $20 \times 20$. We introduce the concept of histograms of normal flow. For each superpixel, we assign the pixels into 16 different bins based on their orientations. The count of pixels, the median value and the standard deviation for each bin is computed. Let $A_{16\times 2}$ be a collection of unit vectors representing 16 bins. We have used weighted least squares [2] for solving the histogram values. The weight matrix, $W_{16\times 16}$ is a diagonal matrix in which the terms correspond to the number of flow vectors in each bin. $B_{16\times 1}$ is a vector containing the median magnitude of the vectors stored in the particular bin of interest. The resultant flow vector for the superpixel, $\mathbf{v}$, is given by,

$$(A^T W A)\mathbf{v} = A^T W B \tag{1}$$

Fig. 2c shows the normal flow values for the superpixels computed using our method.



(a) (b) (c) (d)

Fig. 2: (a) The segmented lower legs are shown in red and green based on spatial information. (b) Histogram of normal flow using 16 bins for edge orientations for a $20 \times 20$ superpixel. (c) Histogram solved using weighted least squares. (d) Segmentation using histograms of normal flow.

Once we compute the resultant flow values of superpixels, we compare its motion magnitude against that of the torso. The superpixels whose motion is lower than that of the torso are classified as belonging to the standing (stance) limb. The superpixels whose magnitude are greater than the torso motion are labeled as belonging to the swinging leg.

After the segmentation of lower segments, we describe the motion of the lower leg segments by translation in the sagittal plane, $(t_x, t_y)$, and rotation, $\omega$, around an axis orthogonal to the sagittal plane. It can be formalized using a reduced version of an affine motion model as described in [13]. We use displacement as an estimate of velocity. The model defines $(\dot{x}, \dot{y})$, i.e. the instantaneous velocity at point $(x, y)$,

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 0 & -\omega \\ \omega & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \tag{2}$$

Using multiple edge points and the corresponding normal flow vectors, the parameters $(\omega, t_x, t_y)$ are computed as the linear least squares (LS) solution [13].

After spatial segmentation, two different motion models are fitted in each frame corresponding to the motion of two segmented lower legs. While performing motion segmentation, we first compute two different motion models according to the labeling of the superpixels. The models are fitted to the entire lower leg segment region (at the pixel level) and residuals are computed. Based on the value of residuals, two finer motion models are fitted to the lower leg segments.

*3) Lower Leg Rotational Velocities and Angular Positions:* The parameter $\omega$ estimated in Sec. II-A.1 and Sec. II-A.2 represents the instantaneous rotational velocity of lower legs at each frame. The residual error in model estimation in certain frames can contribute to an inaccurate velocity profile. This error is accumulated over frames. In order to account for this inaccuracy, we find the orientation of lower legs in frames corresponding to double support and mid swing. To achieve this, we fit Radon transforms [9] to the boundaries of the segmented lower legs, which gives the desired orientation. The residual error is then distributed over the intervals between instances of double support and mid swing.

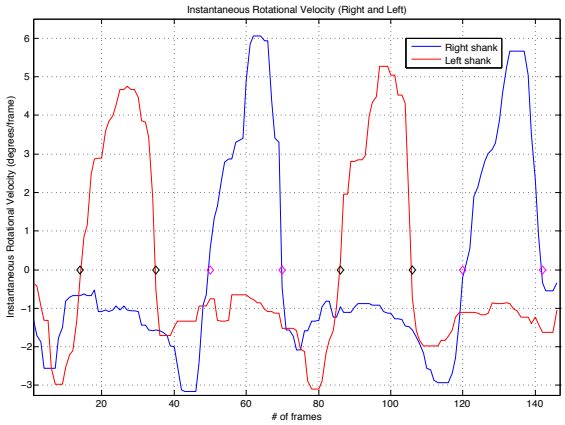



Fig. 3: Instantaneous rotational velocities computed for the two lower leg segments using our method. The zero crossings are marked using black and magenta diamonds.

Fig. 3 shows the instantaneous rotational velocity profiles of both left and right lower leg 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. The initial value for the angular position is automatically computed at one of the double support or mid swing frames. Integrating the velocity profiles gives us the angular positions of lower legs for the entire sequence. This information is then used in Sec. II-C to estimate the angular positions of upper leg and foot.

### B. Modeling Nonlinear Dynamic Systems

In this section we derive a nonlinear state space system [5] to mathematically model the relationship between a pair of joint angles through differential equations. We use 3D motion capture data with angular positions of the upper leg (hip to knee), lower leg (knee to ankle) and foot (ankle to toe) to optimize the parameters of the model.

The upper leg-lower leg and lower leg-foot sets of angular positions are each modeled as a two-state nonlinear dynamic model. It has been shown that in applications where the time-varying variables are coupled, the individual velocities can be defined as a nonlinear function of the positions [10]. The function chosen to model the angular velocities in gait is a sum of $N^{th}$ order polynomials of the angular positions as well as a coupling function that is a product of both angles.

$$\begin{aligned}\theta_X(k+1) =& f(\theta_X(k), \boldsymbol{a}_X) + f(\theta_Y(k), \boldsymbol{b}_X) \\ &+ c_X \theta_X(k)\theta_Y(k) \\ \theta_Y(k+1) =& f(\theta_X(k), \boldsymbol{a}_Y) + f(\theta_Y(k), \boldsymbol{b}_Y) \\ &+ c_Y \theta_X(k)\theta_Y(k)\end{aligned} \tag{3}$$

where $\{X, Y\}$ represent $\{\text{upperleg}, \text{lowerleg}\}$ and $\{\text{foot}, \text{lowerleg}\}$ for the two models respectively. $\theta_X, \theta_Y$ are the angular positions. $f(x, \boldsymbol{a})$ is an $N^{th}$ order polynomial function of the variable $x$. $\boldsymbol{a}$ is an $N$-dimension vector representing the coefficients of the polynomial. $c$ is the coefficient of the coupling function.

$$f(x, \mathbf{a}) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{N-1} x^{N-1} \tag{4}$$

The two equations in Eq. 3 represent a coupled nonlinear dynamic system with the state variables representing the motion of the lower limbs. The parameter set is represented by $\Phi_X = \{\mathbf{a}_X, \mathbf{b}_X, c_X\}$, where $X = H/K/A$ represents the upper leg / lower leg / foot. We choose a $5^{th}$ order polynomial for our model ($N = 5$).

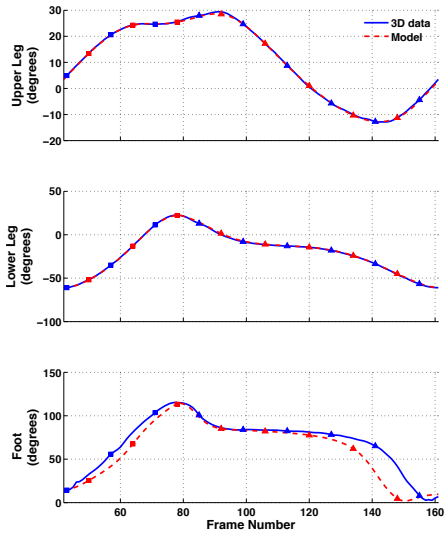



Fig. 4: Observed (3D data) and computed model

Data obtained from the CMU Graphics Lab Motion Capture Database is used to estimate the optimal parameter set by minimizing the least squares error [2] between the angular postions calculated from the model and observed from the

3D data.

$$\Phi_X = \arg \min_{\Phi_X} [\theta_{X,model} - \theta_{X,obs}]^2$$

To further improve the accuracy of the computed angular positions, one gait cycle is divided into two phases based on the angular position of the shank - (i) Toe Off (TO) to Heel Strike (HS) and (ii) HS to TO. Each phase is represented by a different set of parameters ($\Phi_{X,1}, \Phi_{X,2}$) in the model. The onset of each of the phases can be detected using image flow techniques described in Sec. II-A.3 (Fig. 3). The model parameters switch appropriately at the time of transition from one phase to another. The dynamics of the model is independent of the initial angular positions. It only depends on the correct switching between the different parameter sets.

Fig. 4 shows the observed angular positions with the model angular positions for 2 gait cycles. The upper leg flexion, lower leg flexion and foot flexion is shown. Phase 1 data points are represented by squares and phase 2 by triangles. The range of motion for each individual varies slightly within an established interval [6], [7]. We develop different models based on the range of motion of the shank, and choose the appropriate parameter set after comparing the results of Sec. II-A.3, with the existing database.

### C. Unscented Kalman Filter

We now implement an ensemble Kalman filter [4] with the derived nonlinear model to estimate the upper leg and foot angular positions from observations of the lower leg angular position obtained from images (Sec. II-A.3). The nonlinear dynamic model can be represented in terms of a state and output equation

$$\begin{aligned} x(k+1) &= g(x(k), \Phi) + w(k) \\ y(k) &= h(x(k)) + v(k) \end{aligned} \tag{5}$$

The dynamics of the state vector, $x = [\theta_X, \theta_Y]^T$, is described by Eq. 3. The output, $y$, is the observed variable. In our case, $y$ is the lower leg angular position, $\theta_K$, computed in Sec. II-A.3. $w$ and $v$ are the process and measurement noises respectively and are modeled by a Gaussian distribution.

The Kalman filter is implemented to estimate the unobserved state ($\theta_H$ and $\theta_A$) using observations obtained from image flow analysis. The unscented transform [12] is applied to generate an ensemble of data points from the *a priori* state estimate, $x^-(k)$. The covariance matrix of this ensemble is represented by $P^-_{xx}(k)$. The ensemble is propagated forward to the next time step according to Eq. 5 and the predicted output, $y^-_k$ is calculated. The Kalman gain, *a posteriori* covariance matrix, $P^+_{xx}(k)$ and state estimate, $x^+(k)$ are computed according to the following equations.

$$\begin{aligned} K(k) &= P_{xy}(k) P_{yy}(k)^{-1} \\ P^+_{xx}(k) &= P^-_{xx}(k) - P_{xy}(k) P_{yy}(k)^{-1} P_{yx}(k) \\ x^+(k) &= x^-(k) + K(k)[y(k) - y^-(k)] \end{aligned} \tag{6}$$

The estimation begins at the toe off event. An identity matrix of the appropriate dimension is used to initialize the covariance of the ensemble. $\theta_H$ is initialized at zero degrees and requires 3-4 observations to converge to the true value. The advantage of using the ensemble with the unscented transform is that it eliminates the need to linearize the nonlinear dynamic model.

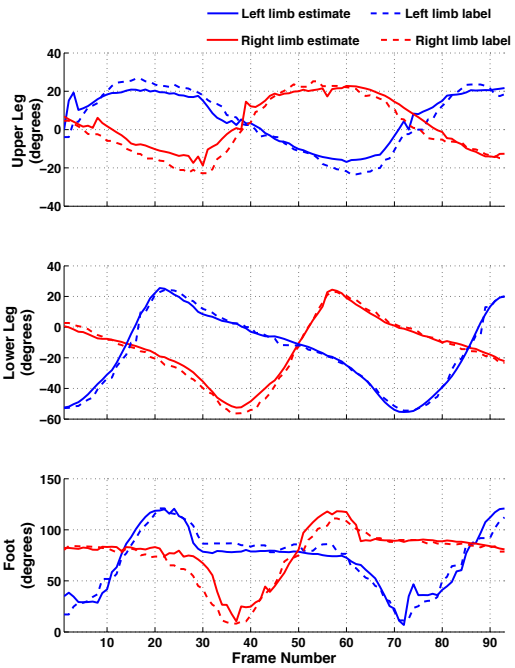



Fig. 5: Estimation of left (blue) and right (red) upper leg, lower leg and foot angular positions. The dotted lines represent manually labeled angular positions of limbs.

Fig. 5 shows the estimated angular positions of the lower limbs for a complete gait cycle. To compare the estimates with the true values, the angular positions were labeled manually and are represented by dotted lines.

## III. Results

We have captured five sequences of gait data with different types of shoes and clothing and processed the data to detect different events of the gait cycle and estimate the kinematics of the lower limbs. The data were recorded in the sagittal plane at 60 frames/second at resolution $640 \times 480$ 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.

### A. Kinematics of lower limbs

Fig. 6 shows the image frames from two sequences, overlaid with our results of tracking the lower limb segments. We have projected the angular positions (hip-knee) and (knee-foot) estimated in Sec. II-C on the corresponding image frames. It can be seen that the motion of the lower limbs is computed accurately using our technique.

### B. Validation

To establish the validity of our method, we have compared our results with manually labeled angular positions of the lower limbs in all the video sequences. Table I gives the average error (in degrees) in the computation of lower limb

Fig. 6: Frames overlaid with our results of motion estimation of lower limb segments.

angular positions using our method. Usually, the normal and abnormal gait parameters differ by more than $7 - 8°$. An estimation error of less than $5°$ is acceptable for practical purposes.

TABLE I: Table listing the average error (in degrees) in the computation of angular positions of different lower limbs

| | Upper Leg | Lower leg | Foot | Average |
|---|---|---|---|---|
| Left | 3.66° | 2.11° | 5.76° | 3.84° |
| Right | 4.71° | 2.45° | 6.52° | 4.56° |
| Average | 4.18° | 2.28° | 6.14° | 4.20° |

## IV. Conclusions and Discussion

We have presented a markerless system to analyze human gait. We have used normal flow and motion models to compute the torso instantaneous velocity and rotational velocities of lower leg segments accurately. The key phases of a gait cycle (double-support, heel-strike, mid-swing, toe-off) were identified where the translational or rotational component of displacement changes its direction (instantaneous velocity is zero). We then fitted a non-linear dynamic model to 3D motion capture gait data defining the relationship between the upper leg - lower leg and lower leg -foot motions. We used Unscented Kalman filter to estimate the upper leg and foot rotation based on the observations of the lower leg rotational velocity. This novel approach, enables accurate identification of key phases of gait and the velocities of lower limb segments. Several examples were presented showing the overlay of our results on the original frames of the video sequences. We believe this method can be reliably used as an assessment tool to determine individual gait patterns, identify abnormalities in need of treatment and assess treatment response. In the future, we would like to model the relative relationships among other body segments and utilize Kalman filter for non linear motion analysis for the whole body. Also, we would test our approach with outdoor video sequences and for multiple views. Applications to pathological gait, and to activities in which the sampling is greater than 60 frames/sec, will require additional investigation.

## References


[1] http://mocap.cs.cmu.edu/.

[2] S. P. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[3] J. Courtney and A. M. De Paor. A monocular marker-free gait measurement system. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 18(4):453–460, 2010.

[4] S. J. Julier and J. K. Uhlmann. New extension of the kalman filter to nonlinear systems. In *AeroSense'97*, pages 182–193. International Society for Optics and Photonics, 1997.

[5] H. K. Khalil. *Nonlinear systems*, volume 3. Prentice hall Upper Saddle River, 2002.

[6] A. B. D. M P Murray and R. C. Kory. Walking patterns of normal men. *Journal of Bone and Joint Surgery*, 46(2):335–360, 1964.

[7] M. P. Murray, R. C. Kory, and S. B. Sepic. Walking patterns of normal women. *Archives of physical medicine and Rehabilitation*, 51(11):637, 1970.

[8] K. Rohr. Towards model-based recognition of human movements in image sequences. *CVGIP: Image understanding*, 59(1):94–115, 1994.

[9] M. Sonka, V. Hlavac, and R. Boyle. *Image processing, analysis, and machine vision*. Thomson-Engineering, 2007.

[10] R. Van Buskirk and C. Jeffries. Observation of chaotic dynamics of coupled nonlinear oscillators. *Physical Review A*, 31(5):3332, 1985.

[11] N. Vishnoi, Z. Duric, and N. L. Gerber. Markerless identification of key events in gait cycle using image flow. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 4839–4842. IEEE, 2012.

[12] E. A. Wan and R. Van Der Merwe. The unscented kalman filter for nonlinear estimation. In *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*, pages 153–158. IEEE, 2000.

[13] H. Wechsler, Z. Duric, F. Li, and V. Cherkassky. Motion estimation using statistical learning theory. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(4):466–478, 2004.

[14] D. A. Winter. *Biomechanics and motor control of human movement*. Wiley. com, 2009.

[15] J.-H. Yoo and M. S. Nixon. Markerless human gait analysis via image sequences. 2003.