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

Gait Cycle Spectrogram Analysis using a Torso-attached Inertial Sensor

Mitchell Yuwono, Steven W. Su, Bruce D. Moulton, and Hung T. Nguyen

Abstract—**Measurement of gait parameters can provide important information about a person's health and safety. Automatic analysis of gait using kinematic sensors is a newly emerging area of research. We describe a new way to detect walking, and measure gait cadence, by using time-frequency signal processing together with spectrogram analysis of signals from a chest-worn inertial measurement unit (IMU). A pilot study of 11 participants suggests that this method is able to distinguish between walk and non-walk activities with up to 88.70% sensitivity and 97.70% specificity. Limitations of the method include instability associated with manual fine-tuning of local and global threshold levels.**

I. Introduction

The ability to measure walking activities can be important for monitoring certain medical conditions and assessing treatment efficacy [1]. For example, in a clinical study involving 488 chronic heart failure patients, Passantino reported strong correlation between the chance of survival and changes in the distance walked [2]. Walking and balance involves complex coordination of the limbs and torso [3]. Gait cadence (step rate) can be determined from small swings in torso angle in the sagittal plane due to the periodic shifting of moment of inertia that occurs on each phase [4]. Previous attempts to measure gait using spectra from torso-attached accelerometers include Barralon's estimation of spectral power magnitude using a Short-time Fourier transform (STFT) and Discrete Wavelet Transform (DWT) [4]. Barralon reported detection sensitivity of 78% and specificity of 68.7%. Another approach using accelerometer and wavelet decomposition was reported by Bidargaddi to distinguish walking from other high impact activities with 89.14% sensitivity and 89.97% specificity [1]. This research aims to develop more-accurate gait cycle analysis for ambulatory monitoring systems, such as those we are working on at University of Technology Sydney Centre for Health Technologies [5].

This paper describes a new approach for processing and analysing signals from a chest-mounted IMU. Section II gives an overview of the hardware. Section III describes the method. Section IV presents the data collection. Section V presents results and analysis, and Section VI provides conclusions.

II. Overview

This pilot study used a Shimmer MEMS kinematic module with a 9DoF daughterboard. The 9DoF board has a Freescale MMA7361 tri-axial accelerometer, a Honeywell HMC5843 magnetometer, and an InvenSense500 gyroscope.

Data collection and an Attitude Heading Reference System (AHRS) is run externally in Java 2 Standard Edition (J2SE) running a custom driver adapted from original Shimmer driver, with 3-D visualization under jMonkeyEngine 3.0 [6]. The IMU samples at 50 Hz [7]. We do algorithmic prototyping in MATLAB. The device is strapped on the participant's chest in a way that the torso angle can be observed directly from the pitch measurement. During walking, the frequency of torso-swing ranges from 0.6 to 2.5 Hz, [4, 8], so we sample data for processing at 20Hz.

M. Yuwono, S.W. Su, B. Moulton, and H. Nguyen are with the Faculty of Engineering and Information Technology, University of Technology, Sydney, Ultimo, 2007, NSW, Australia. (e-mail: mitchellyuwono@gmail.com).

III. Method

The method includes three processes: torso angle estimation, time-frequency signal processing, and spectrogram image processing. First, the orientation quaternion q of the sensor is estimated using the explicit complementary filter (ECF) [9] applied to measurements of angular velocity ω and acceleration a. Pitch information θ is calculated from the sensor orientation. The signal is then convolved with a digital band pass filter (bpf) with cutoff frequencies of 0.5Hz and 5Hz to yield θ_{bpf}. Autocorrelation is used to minimize noisy signals on θ_{bpf}. The Discrete Fourier Transform (DFT) is then be applied to the autocorrelated signal using Bartlett's method to extract the spectrogram $S(f,t)$. Overlapped spectrograms are averaged. Spectrogram image processing techniques are then used, starting with smoothing using a Gaussian Filter. A Gabor filter and median-C thresholding are then used. This is done by applying a median filter to the Gabor filtered image and subtracting the result from the normalized image to get the logical mask $M_2(f,t)$. A morphological filter performing smoothing and erosion is then applied to the logical mask. Gait cycles such as walk/not-walk and cadence are then easily extracted. A block diagram is given in Figure 1.

A. Torso Angle Estimation

Torso angle is estimated using ECF applied to the information from the gyroscope and accelerometers [9]. Our initial tests suggest that the information provided by these two sensors is sufficient to estimate torso angle.

The output of the rate gyroscope ω and normalized accelerometer reading $\hat{a}$ can be represented in vector form as in (1) and (2). The symbol ^ denotes a (1-norm) unit vector.

$$\omega = \begin{bmatrix} 0 & \omega_x & \omega_y & \omega_z \end{bmatrix}^T \quad (1)$$

$$\hat{a} = \begin{bmatrix} 0 & \hat{a}_x & \hat{a}_y & \hat{a}_z \end{bmatrix}^T \quad (2)$$

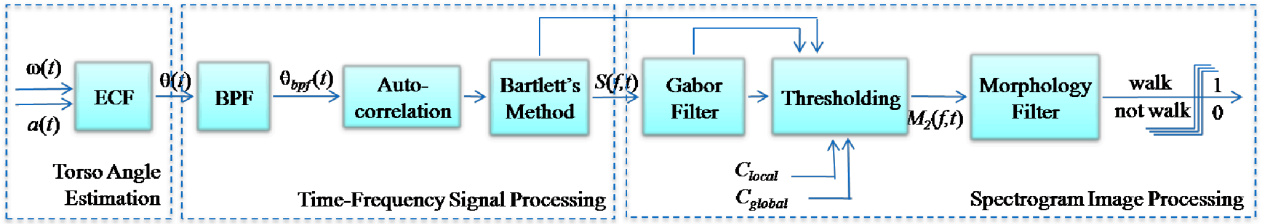

Figure 1. Block diagram describing overall gait parameter measurement - in this case measuring the parameter [walk|not-walk].

System attitude $\hat{q}$ is estimated by integrating $\dot{\hat{q}}$ (3). Rotation $\dot{\hat{q}}$ is evaluated using a simple quaternion product of the current estimate and the compensated gyroscope measurement (4). Innovation δ is updated using proportional-integral (PI) compensation (5). The proportional gain K_P corrects the attitude information by referring to gravity. The integral gain K_I corrects gyroscope bias. The error is the relative rotational discrepancy between the $\hat{q}$ estimate of the z-axis of the inertial frame and the gravitational reference from the accelerometer $\hat{a}$ (6-7).

$$\hat{q} = \int \dot{\hat{q}} dt \tag{3}$$

$$\dot{\hat{q}} = \frac{1}{2}\hat{q} \otimes (\omega + \delta) \tag{4}$$

$$\delta = K_P e + K_I \int e dt \tag{5}$$

$$e = \hat{a} \times \hat{v} \tag{6}$$

$$\hat{v} = \begin{bmatrix} 2(\hat{q}_1\hat{q}_3 + \hat{q}_0\hat{q}_2) \\ 2(\hat{q}_2\hat{q}_3 + \hat{q}_0\hat{q}_1) \\ \hat{q}_0^2 - \hat{q}_1^2 - \hat{q}_2^2 + \hat{q}_3^2 \end{bmatrix} \tag{7}$$

The orientation of the torso is then calculated using a quaternion to Euler angle transformation (8),

$$\theta = \tan^{-1}\left(\frac{2(\hat{q}_0\hat{q}_1 + \hat{q}_2\hat{q}_3)}{1 - 2(\hat{q}_1^{\,2} + \hat{q}_2^{\,2})}\right) \tag{8}$$

In (8), θ represents the pitch angle equivalent of the quaternion that corresponds to the torso angle as per the installed orientation of the sensor.

A block diagram describing the data flow of the torso orientation estimation algorithm is given in Figure 2.

B. *Time-Frequency Signal Processing*

On detection period T_d, which is set to happen every 20 samples, the time domain signal of torso angle is first filtered with a band pass filter. Autocorrelation is applied to the filtered signal to extract its fundamental frequency and reduce unwanted noise. The signal is transformed using Bartlett's method to estimate the power spectra.

A 4th order Butterworth filter with cutoff frequencies of 1 Hz and 5 Hz is used because of its perfectly flat frequency response within its pass band. Autocorrelation is used to extract fundamental frequency of a single tone signal based on its periodicity. This property is highly useful for gait parameter analysis where the gait-cadence frequency is of primary interest [10]. The waveforms before and after the time frequency signal processing are shown in Figure 3.

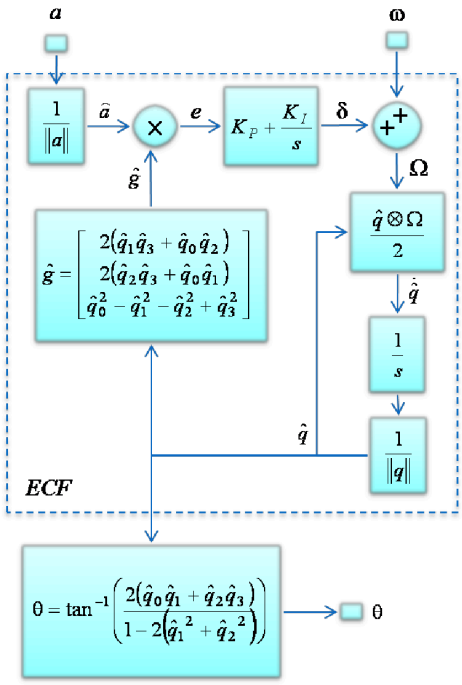

Figure 2. Block Diagram of the Torso Orientation Estimation Algorithm

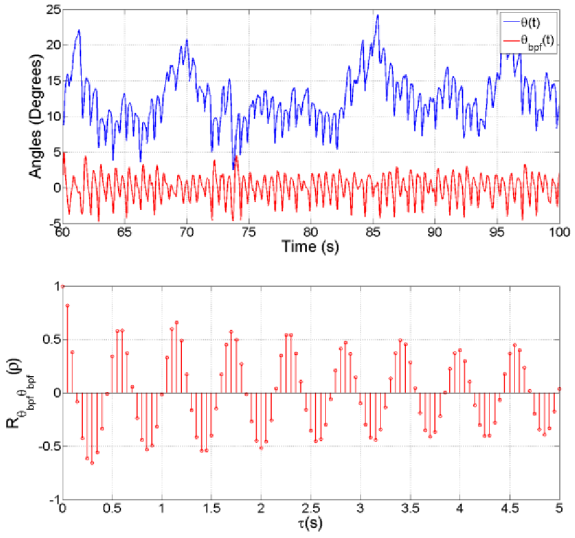

Figure 3. Torso angle θ and and its band passed filtered signal with respect to time during walk (top) and autocorrelation (bottom)

Spectrogram overlaps are averaged as more signals are collected over time. Blackman-Harris windows are used for each DFT transform. In this situation, Bartlett's method is preferable to that of conventional STFT because the spectrograms it produces are denser, more defined, and less noisy. One of the important features of walking is its relatively constant frequency, thus to detect walking, and measure cadence, it is necessary to process the signal in a way that preserves information relevant to gait cadence.

C. *Spectrogram Image Processing*

Spectrogram image processing is exploited as a powerful signal processing method in many fields including speech processing [11] and biomedical engineering [12, 13]. Recent research by El-Gohary demonstrated the capability of spectrogram analysis using kinematic sensors to track tremors in Parkinson Disease patients in free-living conditions [13]. We were interested in exploring whether spectrogram image processing could be used to extract gait cycle information. To do this, we decided to try to extract characteristic "blobs" associated with walking activity. The time-frequency signal processing had filtered out most of the unwanted noise from the spectrogram. However some significant information remained, which, for the purposes of gait parameter measurement, could still be considered noise. This includes non-gait "noise" from large movements such as standing up, sitting, lying down, coughing, looking around, and other high magnitude torso orientation changes that occur during ordinary daily activities. A further challenge we found is that some gait cadences are not sufficiently regular to allow detection solely by identifying periodicity.

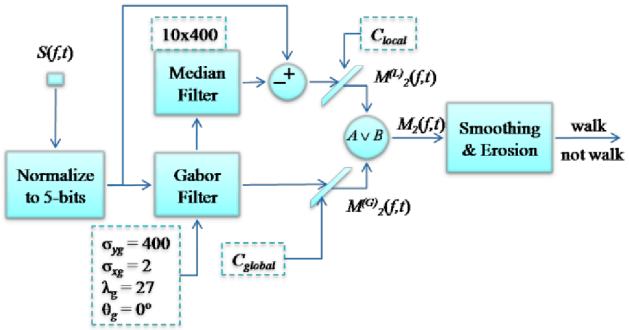

Figure 4. Block diagram of Spectrogram Image Processing for walk detection

Our approach is summarized in the block diagram in Figure 4. After normalization to a 5 bit (32 brightness level) image (Figure 7 top left), the normalized spectrogram image is filtered using a Gabor filter. Gabor filters are proposed to be powerful for texture segmentation [14]. Given our sampling frequency of 20 Hz and 128 DFT bins we defined a Gabor kernel with the following characteristics. The orientation of the filter is set at 0 degrees, to accentuate horizontal lines. The horizontal variance is set to 2 (frequency bandwidth of 0.0781 Hz). The vertical variance is set to 400 (time smoothing for 20 seconds) to magnify and connect disjoint frequencies in the time axis. The wavelength of the sinusoid is set to 1 pixel, which is approximately the frequency bandwidth of 0.0391Hz. An example of a Gabor-filtered image is seen in Figure 5 (top right).

We use both local and global thresholding approaches to preserve important information in a similar way to an approach described by Steinberg for speech spectrogram analysis [11]. To get the local image, a median filter is applied to the Gabor filtered image (Figure 5 bottom left). The result is subtracted from the normalized image. The global image is simply the Gabor filtered image. The local and global thresholds C_{local} and C_{global} are then applied to both images. Both local and global binary masks $M^{(L)}{}_2$ and $M^{(G)}{}_2$ are combined to create the final binary mask M_2. This mask contains logical information of how the blobs are shaped. We then perform smoothing and erosion to filter out most small noises and join closely separated blobs. We identify blobs that satisfy predefined shape parameters. The parameters include a centroid frequency of between 0.6Hz and 2.5Hz, frequency fluctuation of lower than 1Hz and continuity of higher than 15 seconds. The result of the algorithm can be seen in Figure 5, bottom right.

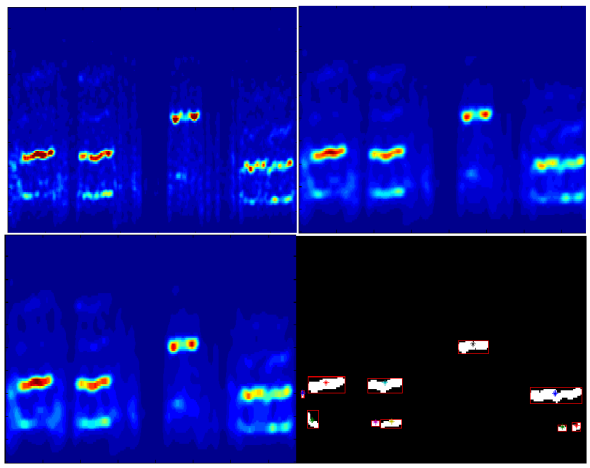

Figure 5. Normalized image (top left), Gabor filtered image (top right), median filtered Gabor Image (bottom left), detection of walk from binary mask after morphology filter (bottom right).

IV. Experimental Settings and Data Collection

The algorithm was tested using data recorded from 11 participants. Three of the participants are elderly people aged over 55. The data was collected in an office environment. The data collection scheme was devised to resemble movements associated with ordinary daily living activities. Each data collection period consisted of ten to fifteen minutes of data collection per participant. During each data collection period, the subjects typically alternated between walking and not walking for a few minutes at a time.

The participants were encouraged to relax and move however they wanted to. The activities of the participants during non-walk periods included talking, browsing the internet, standing up/sitting down, lying down, making coffee, writing, drawing, playing a musical instrument, singing, playing computer games, stretching, and reading books. The participants were also encouraged to walk at a comfortable pace for a period of around two minutes to five minutes. Casual activity such as talking, eating and drinking was encouraged throughout the whole data collection period.

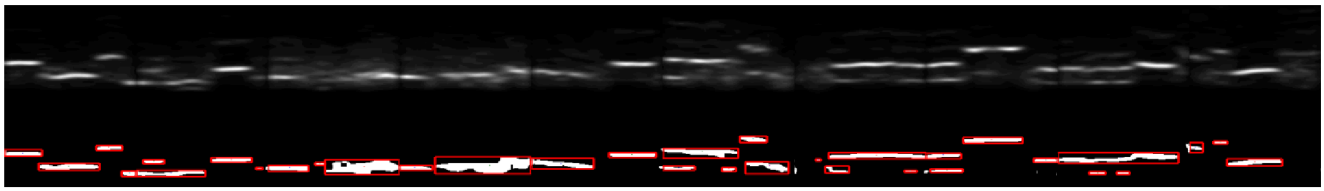

Figure 6. Processed data relating to walking

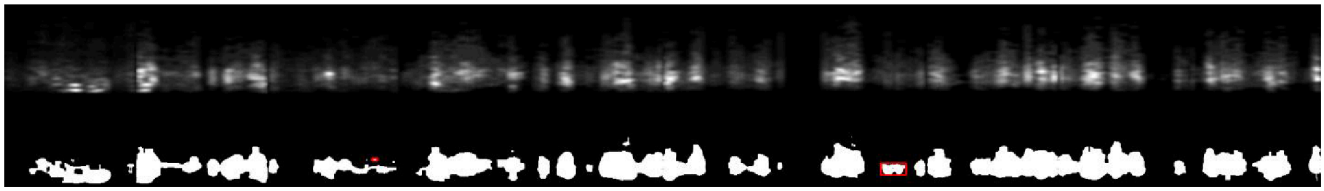

Figure 7. Processed data relating to not-walking.

V. Results and Discussions

A total of 3617.6 seconds (72352 samples) of walking activity and 5511.2 seconds (110223 samples) of not-walking activity were recorded. Processed data relating to walking is shown in Figure 6, and processed data relating to not-walking is shown in Figure 7. Comparing Figure 6 with Figure 7, it can be seen that the morphology of walking activity appears quite different from that of non-walking activity. Using simple extraction of binary morphology properties such as minor axis length, major axis length, and centroid locations, walking activity is distinguishable from not-walking activity.

Table I shows the performance using different local and global threshold levels. The threshold levels are estimated by trial and error based on pixel values in the produced local and global images. The method achieved 88.70% sensitivity and 97.70% specificity using local threshold C_{local}=0.05 and global threshold C_{global}=0.2. These threshold values are manually fine-tuned to properly capture the walking morphology in the spectrogram. These results suggest a significant improvement on specificity compared with prior research [1, 4, 8]. Rejecting noises based on spectrogram analysis is straightforward as the time-frequency morphology of walking is visually distinguishable to other activities. This quality, however, makes it harder to increase detection sensitivity.

Table 1. Performance at various thresholds

Threshold Level		**Performance**	
C_{local}	C_{global}	***Sensitivity***	***Specificity***
0.03	0.1	45.59%	99.18%
0.03	0.4	43.37%	90.72%
0.03	0.3	60.65%	98.17%
0.05	0.15	83.29%	97.70%
0.03	0.2	85.83%	97.70%
0.05	0.2	88.70%	97.70%

VI. Conclusions and Future Directions

Using an accelerometer and a gyroscope, the results of the pilot study suggest that spectrum analysis may be useful for measuring certain gait parameters, including the ability to distinguish walking activity from other activity. We have shown that greater specificity can be achieved from investigating morphology features of the spectrogram image. A limitation of the method is instability due to the need for (manually performed) fine-tuning of the local and global threshold levels. Our plans for the future include further work to implement automatic tuning of local and global threshold.

References

[1] N. Bidargaddi, A. Sarela, L. Klingbeil, and M. Karunanithi, "Detecting walking activity in cardiac rehabilitation by using accelerometer," in *Proc 3rd International Conference on Intelligent Sensors, Sensor Networks and Information,* Melbourne, 2007, pp.555-560.

[2] A. Passantino, R. Lagioia, F. Mastropasqua, and D Scrutinio, "Short-term change in distance walked in 6 min is an indicator of outcome in patients with chronic heart failure in clinical practice," *Journal of the American College of Cardiology*, vol. 48, no. 1, Jul, 2006 pp. 99-105.

[3] C.L. Vaughan, B. L. Davis, J. C. O'Connor, *Dynamics of Human Gait 2nd Edition*. South Africa: Kiboho Publishers, 1999, ch2.

[4] P. Barralon, N. Vuillerme, N. Noury, "Walk Detection With Kinematic Sensor: Frequency and Wavelet Comparison," in *Proc 28th Annual International Conference of the IEEE EMBS*, New York, 2006, pp.1711-1714.

[5] M. Yuwono, B.D. Moulton, S.W. Su, B. G. Celler, and H.T. Nguyen, "Unsupervised machine learning method for improving the performance of ambulatory fall detection systems," *BioMedical Engineering OnLine*, vol. 11, no. 9, 2012. [Online] Available: http://www.biomedical-engineering-online.com/content/11/1/9

[6] jMonkeyEngine.org (2012, May 26). JMonkeyEngine 3.0. [Online] Available: http://jmonkeyengine.com/

[7] S.O.H. Madgwick, *An efficient orientation filter for inertial and inertial/magnetic sensor arrays*. April 2010, ch 5. [Online] Available: http://www.x-io.co.uk/res/doc/madgwick_internal_report.pdf

[8] M. Henriksen, H. Lund, R. Moe-Nilssen, H. Bliddal, and B. Danneskiod-Samse, "Test-retest reliability of trunk accelerometric gait analysis." *Gait Posture*, vol. 19, no. 3, Jun 2004, pp. 288–97.

[9] M. Euston, P Coote, R. Mahony, J. Kim, T. Hamel, "A Complementary Filter For Attitude Estimation of a Fixed-Wing UAV," in *Proc IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Nice, 2008, pp. 340-345.

[10] C.C. Yang, Y.L. Hsu, K.S. Shih, and J.M. Lu, "Real-Time Gait Cycle Parameter Recognition Using a Wearable Accelerometry System," *Sensors,* vol. 11, no. 8, 2011, pp.7314-7326.

[11] R. Steinberg, D. O'Shaughnessy, "Segmentation of a Speech Spectrogram using Mathematical Morphology," in *Proc IEEE International Conference on Acoustics, Speech and Signal Processing*, Las Vegas, 2008, pp. 1637-1640.

[12] A. M. Gavrovska, M.P. Paskaš, and I.S. Reljin, "Determination of Morphologically Characteristic PCG Segments," *Telfor Journal*, vol. 2, no. 2, 2010, pp. 74-77.

[13] M. El-Gohary, J. McNames, K. Chung, M. Aboy, A. Salarian, and F. Horak, "Continuous At-Home Monitoring of Tremor in Patients with Parkinson Disease", in *Proc Biosignal 2010: Analysis of Biomedical Signals and Images*, vol. 20, 2010, pp. 420-424.

[14] D. Dunn, W.E. Higgins, "Optimal Gabor Filters for texture segmentation," *IEEE Trans. Image Processing*, vol. 4, no. 7, Jul 1995.