Detection of Gait Phases Using Orient Specks for Mobile Clinical Gait Analysis

R L Evans and D K Arvind
Centre for Speckled Computing
School of Informatics, University of Edinburgh
10 Crichton Street, Edinburgh EH8 9AB, Scotland, UK
robert.livingston.evans@gmail.com; dka@inf.ed.ac.uk

***Abstract*—This paper presents a hybrid method based on a feed-forward neural network (FNN) embedded in a hidden Markov model (HMM), for detecting phases in a gait cycle, based on data from inertial sensors attached to the lower body. The method was validated against the ground truth obtained concurrently from a Vicon optical motion capture system for five volunteers. The method was characterised using metrics such as sensitivity and specificity for sensor placements, and gait analysis. The results demonstrate that the proposed method is accurate within 23 milliseconds with the added advantages of mobility afforded by wireless sensors and the flexibility of the classification method.**

Keywords: mobile gait analysis, inertial sensors, feed-forward neural network, hidden Markov model.

I. Introduction

The 3-D motion capture of human gait outside the clinic has been a challenge. The use of wireless inertial sensors attached to the lower limbs is a promising approach for mobile clinical gait analysis [7]. This paper addresses the problem of detecting automatically the different phases in a normal gait based on inertial sensor data and the results are validated against the ground truth provided concurrently by a Vicon optical motion capture system. In practice a method is desired which is not brittle or difficult to adapt, by not relying on careful sensor alignment or on rule-based thresholding, respectively.

This paper proposes a hybrid feed-forward neural network (FNN)/hidden Markov model (HMM) approach used successfully in other applications such as Automatic Speech Recognition [1], which combines the strengths of both techniques. The HMM transition probabilities encode the sequential ordering of human gait phases; this ordering is not encoded by the FNN, which processes windows of sensor values in isolation. Similarly, the HMM emission probabilities encode the error rates of the FNN classifications given the current state, which provides contextual information unavailable to the FNN on its own. Thus the HMM complements the FNN by providing context to its pattern recognition. In principle, any machine learning method can be used instead of FNN, such as Support Vector Machines (SVM), but the strength of FNN lies in its ability to learn non-linear combinations of the inputs automatically, whereas the kernel function in SVM has to be tailored for this purpose.

In the rest of this paper: Section II describes other methods proposed for gait phase detection; Section III describes the different phases in a human gait, and a hybrid method for detecting them; Section IV details the experimental setup for synchronised data capture from the inertial sensors and the Vicon optical motion capture system. The FNN/HMM method is characterised in terms of metrics such as sensitivity and specificity (see Section IV-C), and gait analysis results are presented for five subjects.

II. Related Work

Pappas *et al* [5] developed a rule-based method for gait phase detection which used threshold values from a gyroscope embedded in the insole of the shoe. HMMs have been applied in isolation, but generally for low dimensional inertial data [8]. Bae and Tomizuka [2] presented a HMM-based gait phase analysis method for data derived from a "smart shoe" with air-bladder sensors to measure ground contact forces. Mannini and Sabatini [3] reported a classifier based on the HMM for gait phase detection and walking/jogging discrimination using a uniaxial foot-mounted gyroscope. Guenterberg *et al* [4] introduced a generic method called hidden Markov *event* model based on HMM for extracting temporal parameters, which was applied to a walking dataset derived from inertial sensors. The features for the HMM were extracted manually and a genetic algorithm selected the optimal combination for the feature vectors.

The proposed hybrid approach is well suited to the particular constraints of classifying high dimensional sensor data, and the resulting gait phase detector is less dependent on brittle sensor configurations [2,3,5]. The hybrid approach can also be adapted to segment subjects' other activities where training data is available. The performance of the hybrid method compares favourably with the carefully aligned uniaxial gyroscope in the sagittal plane reported by Mannini and Sabatini [3], of an average sensitivity value of 94% and specificity value of 98% within 20ms for four phases (excluding the mid-swing), compared to 88% and 97%, respectively within 23ms for five phases (see Section IV-D). The capacity to integrate many sensor readings allows the proposed method to detect events defined by multiple body areas, *e.g.,* the mid swing is defined by the knee angles whereas the heel strike is defined by the foot. Previous work in this area invariably focused on foot-defined events [2,3,5].

The contributions of this paper are: the first instance of a hybrid FNN/HMM model for detecting gait phases in high dimensional data derived from inertial sensors; the validation of this method against ground truth data provided by a Vicon optical motion capture system for all the five phases in the stance and swing periods of normal human gait; characterisation of the hybrid method and derivation of gait analysis results potentially of use in mobile clinical gait analysis.

III. The Method

A. Hybrid FNN/HMM Classifier

A hybrid feed-forward neural network (FNN)/hidden Markov model (HMM) classifier is proposed for detecting gait phases in high dimensional inertial sensor data.

FNNs are well suited for recognising patterns in high dimensional sensor input data, although the input is required to have a fixed size. This presents difficulties when dealing with continuous streams of motion capture data from inertial sensors, which have events of variable time spans. In addition, there is no explicit treatment by the FNN of the input as a sequence.

HMMs, by contrast, model sequences explicitly. Observed sequences are modelled as emissions from a set of unobservable states. An HMM can be characterised as follows: a set of transition probabilities between hidden states; a set of emission probabilities for each state; and the initial state probability distribution. Given a sequence of observed emissions and an HMM, the Viterbi algorithm can infer the most likely sequence of hidden states to have produced the emissions. HMMs are less suited for high dimensional data as the number of possible states and emissions grows exponentially.

The proposed hybrid model embeds a multi-layer FNN within the HMM. The five gait phases of interest are the hidden states of the HMM, and the classifications of sensor data by the neural network are the emissions from the hidden states. The hybrid classifier is described in more detail next.

The FNN uses a sliding window to classify the continuous stream of sensor data, *i.e.*, each frame contains all the sensor values for a given instant, and the values from all frames in the window are concatenated to form the input vector for the neural network. The FNN assigns a class label to each vector. Class labels are learned from annotated training data, where each window's class is the phase of gait at the time of the central frame (*e.g.,* the swing phase).

The FNN sliding window classifications represent the emissions of the HMM. The Viterbi algorithm is used to infer the most likely sequence of hidden states (gait phases) given the observed sequence of emissions (FNN classifications) and a trained HMM. The final output of the classifier is the most likely sequence inferred. To train the HMM, the FNN classifications on a training set are compared with the true labels for the training set. This provides the emission probability of each FNN output given the true state. The transition probabilities between true states are calculated by counting the transitions between true window classes in the training data.

B. Gait phases

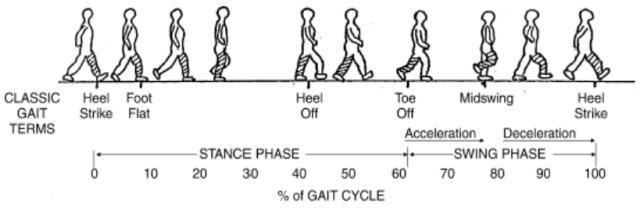

Figure 1. Phases in a normal human gait [6].

The human gait cycle consists of two main stages: the swing and the stance, which are defined independently for each leg. For our analysis, the swing and stance stages are sub-divided into five phases demarcated by the following gait events: stage – STANCE: phases - heel strike, foot flat, heel off and toe off; stage – SWING: phases - toe off, mid-swing and heel strike.

The mid-swing is defined as the moment of maximum knee flexion of the swinging leg. The other gait events are illustrated in Figure 1.

IV. Data Capture and Results

A. Synchronised data capture

The Orient-3 specks [9] are body-worn wireless inertial-magnetic units for full-body, 3-D motion capture in real-time developed by the Centre for Speckled Computing at the University of Edinburgh. A single unit measures 36x28x11 mm, weighing 23 grams (including the battery), and consists of three-axis accelerometer, gyroscope and magnetometer sensors. These sensors are sampled at a rate of up to 512 Hz, and the orientation estimations are performed locally on a 16-bit dsPIC processor at a rate of 256 Hz. Using a simple Time Division Multiple Access (TDMA) protocol, fifteen devices transmit wirelessly to a base station a frame of full-body motion data at a rate of up to 64 Hz. However, this study used the Orient-3 specks in a different mode where the raw data from the three sensors (although only the accelerometer and gyroscope data are used in this study) were transmitted wirelessly to the base station attached to a laptop for storage and subsequent analysis.

A marker-based Vicon gait analysis system captured the motion of the human subjects during normal walking and provided the ground truth for the experimental validation of the techniques presented in this paper. The data capture on the Vicon system was triggered from the Orients which ensured that both systems were synchronised [7]. The two data frames were matched exactly for comparison in the temporal domain by interpolating the Orient data sampled at 42.67 Hz with the Vicon data sampled at 100 Hz.

Seven pairs of reflective markers and Orient specks were attached at the following locations on the subject's lower body: base of the spine, and one on each thigh, shank and foot. After a brief calibration period lasting 15 seconds, the subjects were invited to walk the length of a 30-meter level walkway in a clinical gait laboratory. Five subjects (two male, three female) with normal walk were chosen for the experiments and four trials per subject were recorded.

B. Data Analysis

The FNN was modelled in Matlab's Neural Network Toolbox and consisted of three hidden layers of ten units each, and an output layer consisting of five units corresponding to the five gait phases. The input vector to the FNN was obtained by using a sliding window of fifteen frames, with a step size of one. Each window had 630 data points obtained from seven Orients, each unit producing six dimensions of data from the three-axis accelerometer and

gyroscope (6 sensors x 7 Orients x 15 frames). The network was trained using a scaled conjugate gradient back propagation algorithm until it converged, using cross-entropy as the performance measure during training. A visual inspection of the Vicon data animation was used to detect the different gait phases for labelling the inertial data from the Orients which was used to train and test the FNN/HMM classifier.

Leave-one-subject-out cross validation was used to split the data into test and training sets to obtain the results in this section, *i.e.*, for each fold of the cross validation, the test set consisted of a single trial from one subject (~ 77 frames, on average), while the data from other subjects formed the training set (~ 1240 frames, on average). The test set was rotated with each fold so that all the data was eventually used for testing and training and the average performance of all tests are reported. The test and training sets were mutually exclusive in each fold. When reporting the average sensitivity and specificity over all folds, the relevant counts from each fold were summed independently and substituted into the respective equations (see Section IV-C).

The HMM and Viterbi decoding algorithms were modelled using the MATLAB Statistics toolbox. For training the HMM, the trained FNN was first used to classify all of the windows in the training set, and the correspondences between true states and classified states were used to estimate the emission matrix of the HMM. One pseudo count was added for each possible state emission to avoid zero probabilities, if FNN misclassifications occur that do not happen in the training data. A uniform distribution was used for the probability of the initial state. The transition matrix was estimated by counting the transitions between ground truth classes of each sliding window. Examples of transition and emission matrices for the HMM are shown in Tables 1 and 2, respectively (the phases in the tables are named after the most recent events, *e.g.*, the 'Heel strike' phase continues until the 'Foot flat' event).

The transition probabilities encode the sequential ordering of human gait phases. This ordering is not encoded by the FNN, which only considers the windows of sensor values in isolation. Similarly, the emission probabilities encode the error rates of the FNN classifications given the current state, which provides contextual information unavailable to the FNN in isolation. In summary, the HMM complements the FNN by providing context to its pattern recognition.

C. Performance measures

Useful measures for the performance of the gait phase detection method were defined to be *sensitivity* (true positive rate) and *specificity* (true negative rate) for each phase class.

Sensitivity = True positive count / (True positive count + False negative count)
Specificity = True negative count / (True negative count + False positive count)

The true, and false, positives and negatives were counted independently for each class, with one count for each frame. To obtain a performance measure over all classes, the counts of all classes were summed and used to calculate single sensitivity and specificity figures. More precisely, the true positive counts from each phase class were summed, as were the true negatives, false positives and false negatives. These summed counts were substituted in the sensitivity and specificity equations.

TABLE 1. MATRIX OF HIDDEN STATE TRANSITION PROBABILITIES

Phase transition		**TO**				
		Heel strike	Foot flat	Heel off	Toe off	Midswing
FROM	Heel strike	0.75	0.25	0	0	0
	Foot flat	0	0.95	0.05	0	0
	Heel off	0	0	0.86	0.14	0
	Toe off	0	0	0	0.89	0.11
	Midswing	0.10	0	0	0	0.90

TABLE 2. MATRIX OF EMISSION PROBABILITIES GIVEN THE HIDDEN STATE

Emission: (in phase)		**EMIT**				
		Heel strike	Foot flat	Heel off	Toe off	Midswing
IN	Heel strike	0.6923	0.0769	0.0769	0.0769	0.0769
	Foot flat	0.0222	0.8889	0.0222	0.0444	0.0222
	Heel off	0.0526	0.0526	0.7895	0.0526	0.0526
	Toe off	0.0667	0.0667	0.0667	0.7333	0.0667
	Midswing	0.0667	0.0667	0.0667	0.2000	0.6000

When considering sensitivity (True Positive Rate) and specificity (True Negative Rate) metrics, it is important to understand the baselines for comparison. For example, a gait phase detector that classifies all frames of a gait as having the same class may score around 20% sensitivity and 80% specificity. This is because the sensitivity score comes from the class that happened to be the same as the one mistakenly assigned to all frames, so all of that class's positives will have been correctly classified. The specificity score arises from the fact that most of the frames are negative examples of each phase. For example, most frames are not heel strikes, so classifying all frames the same will produce true negatives for most frames.

The sensitivity and specificity values were also computed with a tolerance, +/-n frames [3], so that a gait event was counted as detected even if it was missed by +/-n frames. This does not affect the output of the classifier but quantifies the performance within a given threshold.

D. Results

Figure 5 illustrates the results of applying the FNN/HMM method, for phase detection in the left and right legs, to the inertial sensors dataset for a typical trial, together with the ground truth data obtained from a visual inspection of the animated Vicon data. The figure also highlights the metrics of "sensitivity" and "specificity" for stringent tolerance values of +/- 0 frames.

The HMM ensures that the output classes follow the transitions of states; one does not see toe offs immediately following heel strikes, for example. If the FNN misclassifies a frame but correctly classifies the frames around it, then the Viterbi decoding of the most likely

sequence will correct this. There are some unclassified frames at the start and at the end of the motion data, which are due to the inclusion of seven frames either side of the frame being classified in the sliding window.

By summing the performance counts (true and false positives and negatives) of all the classes, for all iterations of the cross validation and using them in the expressions in Section IV-C, an average performance over the entire dataset is obtained. The average specificity calculated was 88.7% and average sensitivity was 97.2% for zero tolerance. This is equivalent to an accuracy of within 23 milliseconds, for the data capture rate of 42.67 frames per second.

Figure 2 (left) explores the impact of changes in tolerance on the sensitivity and specificity metrics. The performance increases up to a tolerance of +/-2 frames, *i.e.*, a window of 117 milliseconds for an average gait cycle in the dataset of about one second.

TABLE 3. AVERAGE SENSITIVITY AND SPECIFICITY MEASURES

Subject	Sensitivity (%)	Specificity (%)
1	82.33	97.48
2	93.25	96.37
3	80.10	97.18
4	87.47	96.91
5	89.53	97.44
Overall	88.49	97.12

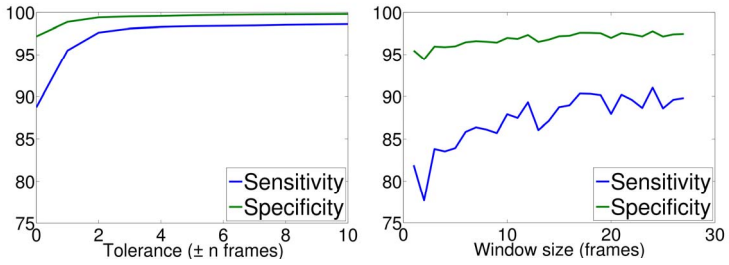

Figure 2. (left) Impact of tolerance on sensitivity & specificity; (right) Impact of FNN window size on sensitivity & specificity.

Having a window of tolerance when calculating performance is useful for determining the range in which the classifier is accurate. What constitutes a useful window of accuracy will vary with the application, and the onus is on practitioners to decide what range is acceptable to them. In this paper, however, results are reported with zero tolerance frames to provide a clear baseline for comparison with other techniques. In addition to varying the tolerance, the effect of changing the size of the sliding window was also explored. Using a smaller window is advantageous in reducing the computational load of the task, which raises the possibility of performing the computation locally on the sensor network. The cross validation was rerun with many different window sizes, using no tolerance frames in the performance calculation, as shown in Figure 2 (right). Performance is slightly higher for larger sliding window sizes, peaking at 24 frames, with good results seen even with individual frames as input.

TABLE 4. GAIT ANALYSIS FIGURES PER SUBJECT

Subject	Cycle Time (s)	Steps per minute	Left Stance Ratio	Right Stance Ratio	Average Stance Ratio
1	1.06	56.71	0.57	0.58	0.58
2	0.97	61.65	0.62	0.64	0.63
3	1.01	59.31	0.64	0.64	0.64
4	0.95	63.21	0.60	0.60	0.60
5	0.90	66.78	0.64	0.63	0.63

This is due to the high dimensionality within each frame: there are seven Orient devices with six sensors each. A single frame therefore contains enough information about the posture and momentum of the legs for the FNN to identify the most likely gait phase. This is not the case when using a smaller number of sensors, and a correspondingly larger window size is required for accurate results. In order to illustrate this, the cross-validation was rerun using just uniaxial accelerometer data from the feet (two dimensions per frame). This produced a sensitivity figure of 85% and a specificity figure of 96% for a window size of 20 frames, which dropped to 74% sensitivity and 93% specificity for a single frame. Using a smaller number of frames, where possible, is useful in reducing the computational load of training and testing the classifier, but with a detrimental effect on the accuracy of the classifier. The detection of gait phases in both legs can be used to compare the gaits of different subjects. This provides extrinsic validation of the method as shown in Table 4 for each subject averaged over both legs. The ratio of the swing to stance periods is close to the expected value of around 60% as shown in Figure 1.

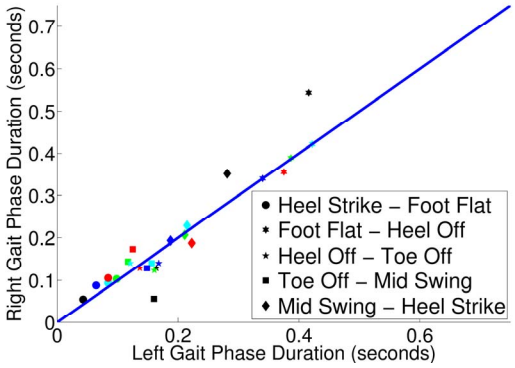

Figure 3. Gait analysis for phase duration asymmetry.

Figures 3 and 4 illustrate the analysis of the durations of the gait phases for each leg for the five subjects. The straight line indicates where data for a perfectly symmetrical gait would lie. The results show a fairly even distribution around the symmetry line on average, although there is some variability between subjects as would be expected.

While there is not sufficient data to draw strong conclusions, this type of analysis can be used to perform personalised evaluation of subjects in comparison to the population average. For example, asymmetry in the swing and stance phases is revealed for the subject with the black dots, who is left-handed. They appear to have a left-biased swing phase and a right-biased stance phase, where the stronger side drives the gait and hence a longer swing phase, and the weaker side is used to balance, indicating a longer stance period.

The performance of the method for gait phase detection was further characterised as shown in Figure 6 by varying the numbers and placements of the Orient devices. We hypothesised that reducing the number of Orients where possible would reduce the computational complexity at a cost to detection performance. We also expected that the Orients would be more effective in some positions than others, and that the effectiveness might vary for each type of gait phase. The results presented so far were obtained using seven Orients placed on the feet, shins, thighs and pelvis; Figure 6 shows the sensitivity and specificity of the method for each gait phase when trained using the Orients on the feet, shins, thighs or pelvis alone. The results are computed from the sum of performance counts over all folds of the cross validation, as described previously.

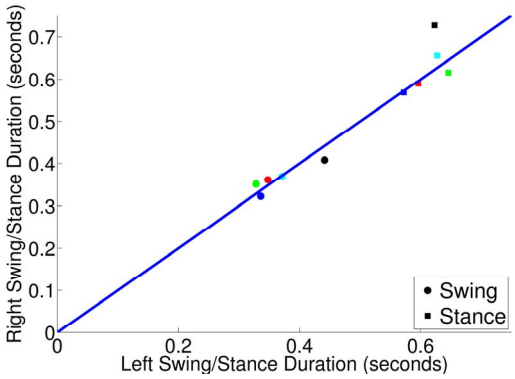

Figure 4. Swing/Stance duration asymmetries (colour coded by subject).

It should be noted that the performance of the gait phase detection method remains high even when using a single Orient placed at the base of the spine. The performance is highest when using all of the sensors in combination, but any of the sensors in isolation also produces good results. In general, however, the differences were marginal and the phase detection was possible using any lower body placement. The classifier was tested with the accelerometer and gyroscope data in isolation. Reducing the number of sensors is advantageous for ease of use, cost, and computational complexity, if performance can be maintained. Figure 6 shows the sensitivity and specificity by placement, gait phase, and type of sensor. Using the classifier with data from only the accelerometer or gyroscope in isolation produces more variable results, by sensor placement, than using the sensors together. The sensitivity is more variable in part because the quantity of data used is reduced to a single pair of triaxial accelerometers, and split further by gait phase. We, therefore, cannot draw firm conclusions about the relative merits of each sensor placement, except to note that similar levels of performance can be achieved with any of them.

V. Conclusion

A method combining feed-forward neural network (FNN) and hidden Markov model has been proposed for detecting the phases in the stance and swing stages in normal human gait. The method can handle high dimensionality (*e.g.,* the dimensionality with 25 frames in the window is greater than 1000), but still works with small numbers of sensors. Even though this technique has been shown to work with low dimensional data, future research will exploit the high dimensionality to classify gait phases and other activities engaged in by the subjects which other methods capable of operating only on low dimensional data cannot handle. Training the FNN can be performed on a GPU for speed; each fold of the cross validations in the experiment took around 2 seconds on an Nvidia GeForce GT 650M graphics card. The test phase requires minimal computation for each frame and could be performed on the sensor network in the future. The current method performs the gait phase detection off line and future work will explore real time gait phase detection on the device, using less data (fewer sensors/frame rate/sliding window step size). Success with a small neural network (just 30 hidden units) suggests that much of the information from the 42 sensors on the legs is redundant or well correlated, and the dimensionality can be reduced and tailored to specific applications. The method does not require careful alignment of sensors in the sagittal plane: although sensor orientation should be consistent throughout test and training data. Future work will extend this method to the analysis of people suffering from gait abnormalities such as antalgic gait, lateral trunk bending, functional leg-length discrepancy, and excessive knee extension, where normal walking is the standard against which the pathology is measured.

VI. Acknowledgement

We wish to thank Smita Sasindran Pochappan for collecting the data and the staff at the SMART Centre, Astley Ainslie Hospital for their support. This research was supported by the Centre for Speckled Computing (www.specknet.org).

References

[1] Bourlard, H., and Morgan, N.. "Merging Multilayer Perceptrons and Hidden Markov Models: Some Experiments in Continuous Speech Recognition" International Computer Science Institute TR-89-033, 1989.

[2] Bae, Joonbum and Tomizuka, Masayoshi J. "Gait Phase Analysis based on Hidden Markov Model", in Proc. of 5th IFAC Symp. on Mechatronic System, Cambridge, MA, 2010.

[3] Mannini A, and Sabatini A "Gait phase detection and discrimination between walking-jogging activities using hidden Markov models applied to foot motion data from a gyroscope. Gait & Posture 36 (2012) 657 – 661.

[4] Guenterberg et al, "A method for extracting temporal parameters based on hidden Markov models in body sensor networks with inertial sensors", in IEEE Transactions on Information Technology in Biomedicine, Vol. 13, No. 6, Nov. 2009.

[5] Pappas *et al*, "A reliable gyroscope-based gait-phase detection sensor embedded in a shoe insole", IEEE Sensors Journal, Vol. 4, No. 2, April 2004.

[6] Uustal H, Baerga E. Gait Analysis. In: Cuccurullo S, editor. Physical Medicine and Rehabilitation Board Review. New York: Demos Medical Publishing; 2004.

[7] Sasindran S Pochappan, D K Arvind, Jennifer Walsh, Alison M. Richardson, Jan Herman, "Mobile clinical gait analysis using Orient specks" in Proc BSN 2012, London, UK, May 2012.

[8] Abaid N *et al*, "Gait Detection in Children with and without Hemiplegia Using Single-Axis Wearable Gyroscopes", PLoS ONE 8(9): e73152. doi:10.1371/journal.pone.0073152.

[9] A.D. Young, M. J. Ling, D. K. Arvind, "Orient-2 - A Wireless Real-time Posture Tracking System using Local Orientation Estimation", 4th Int. Workshop on Embedded Networked Sensors, pp. 25 - 26 June, Cork, Ireland, July 2007, ACM.

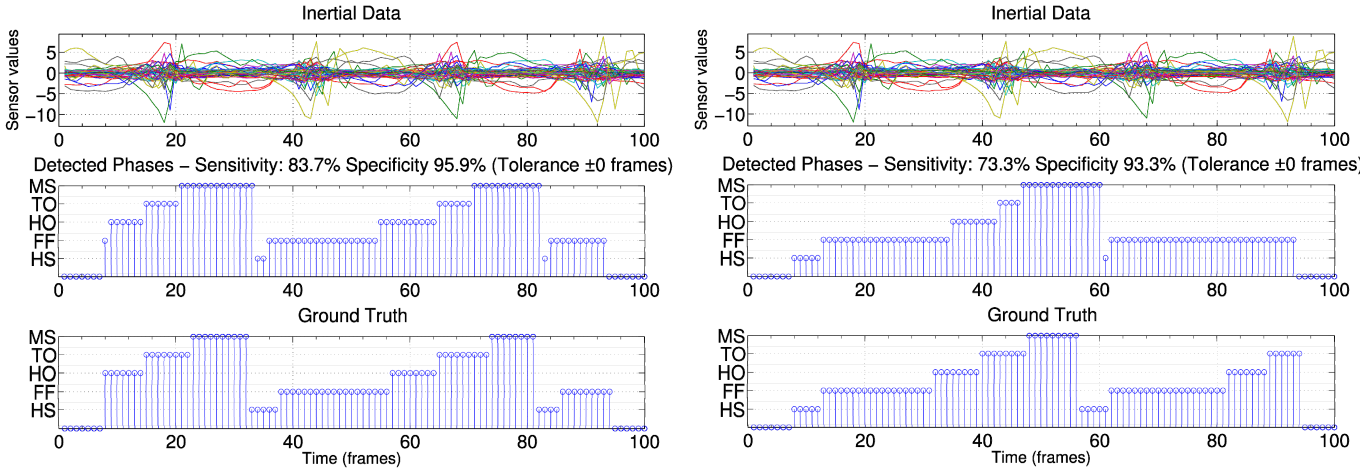

Figure 5. Phase detections for a typical trial, for the left and right legs (left and right figures.)
Top – raw inertial sensor data for both legs; Middle – phase detections of FNN/HMM method; Bottom – ground truth derived from the Vicon data.

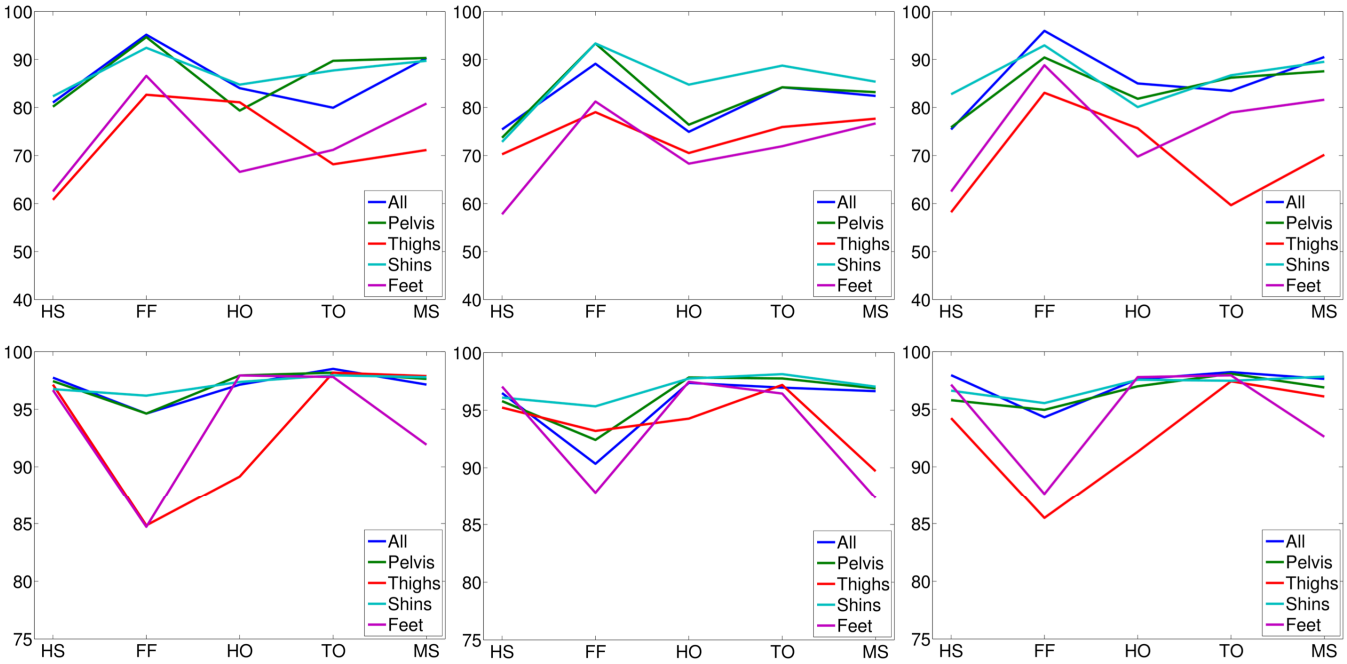

Figure 6. Impact of sensor placement on the detection of each gait phase. (Tolerance +/- 0)
Left column – Accelerometer and gyroscope; Middle column – Accelerometer only; Right column – Gyroscope only.
Top row – Sensitivity; Bottom row – Specificity.