This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2014.2377238, IEEE Internet of Things Journal

Towards Automatic Activity Classification and Movement Assessment During a Sports Training Session

Amin Ahmadi, Edmond Mitchell, Chris Richter, Francois Destelle, Marc Gowing, Noel E. O'Connor, Kieran Moran

***Abstract*—Motion analysis technologies have been widely used to monitor the potential for injury and enhance athlete performance. However, most of these technologies are expensive, can only be used in laboratory environments and examine only a few trials of each movement action. In this paper, we present a novel ambulatory motion analysis framework using wearable inertial sensors to accurately assess all of an athlete's activities in real training environment. We firstly present a system that automatically classifies a large range of training activities using the Discrete Wavelet Transform (DWT) in conjunction with a Random forest classifier. The classifier is capable of successfully classifying various activities with up to 98% accuracy. Secondly, a computationally efficient gradient descent algorithm is used to estimate the relative orientations of the wearable inertial sensors mounted on the shank, thigh and pelvis of a subject, from which the flexion-extension knee and hip angles are calculated. These angles, along with sacrum impact accelerations, are automatically extracted for each stride during jogging. Finally, normative data is generated and used to determine if a subject's movement technique differed to the normative data in order to identify potential injury related factors. For the joint angle data this is achieved using a curve-shift registration technique. It is envisaged that the proposed framework could be utilized for accurate and automatic sports activity classification and reliable movement technique evaluation in various unconstrained environments for both injury management and performance enhancement.**

***Index Terms*—Activity classification; Wearable inertial sensor; Technique assessment; Sensor fusion; Knee joint angle; Curve shift registration; Biomechanics, Smart and connected health**

I. Introduction

SPORT and physical activity have important cardiovascular, musculoskeletal and mental health benefits and are enjoyed by large numbers [1]. However, associated lower body musculoskeletal injuries are very common [2], [3], [4]. Almost all injuries are caused by relative excessive loading on the tissues i.e. high loading relative to tissue strength. One factor that significantly influences this loading is movement technique. Athletes can be biomechanically screened to determine an athlete's predisposition for injury [5] by recording and quantifying both their movement technique (i.e. joint angle and angular velocity) and some measure[1] of loading on their body (e.g. impact accelerations)[2]during a series of actions common to their sport and known to be related to injury (e.g. running [3], jumping and landing [6], agility cuts [9]). Generally, the athlete completes $1-5$ maximum effort trials of each action [6] and their results are compared to normative values, if available [10]. These tests are almost exclusively completed in a laboratory since biomechanics based motion analysis systems tend to be camera based (6+ cameras typically) which must remain spatially fixed during the testing session and tend to be negatively affected by changing lighting conditions. This screening process creates several assessment and comparison challenges, which significantly reduce its ecological validity and usefulness. These include:

1. The athletes are generally highly focused on how they complete the tasks, and therefore may not utilize a movement technique that they would normally use in a training session or match.
2. The controlled laboratory environment does not reflect the conditions of the training environment (e.g. uneven/wet ground, fatigued conditions).
3. The use of only 1 to 5 trials as representative of how an athlete completes a movement technique is highly questionable. The low number of trials is common because of the significant processing time (and cost) associated with optical based systems.
4. There is a lack of normative data for many sports based tasks because of the low number of tested athletes.
5. It is a very expensive process limiting its general application.

A solution to the above assessment challenges would be to use sensors that could be worn throughout a training session or competitive event, detecting an athlete's joint angular motion and impact accelerations. Accelerometers mounted on the body can be used to infer loading based on Newton's second law of motion $(F = ma)$[6], [7], [8] during every foot-ground contact. We estimate that within a 45 minute training session this could involve each foot striking the ground (more than2000) times. By taking advantage of the advancement in

A.Ahmadi, E.Mitchell, C.Richter, F.Destelle, M.Gowing, N.E.O'Connor, K.Moran are with Insight: Center for Data Analytics, Dublin City University, Dublin, Ireland.

A.Ahmadi, C.Richter and K.Moran are with Applied Performance Research, School of Health and Human Performance, Dublin City University, Dublin, Ireland.

[1]Direct loading on individual tissues cannot be measured in a non-invasive fashion but it is possible to determine aggregate loading on a region of tissues or structures.

[2]Technically this should be referred to as deceleration, but the term acceleration is used throughout this manuscript in line with the current biomechanical literature [6], [7], [8]

2327-4662 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2014.2377238, IEEE Internet of Things Journal

microelectronics and other micro technologies, it is possible to build inexpensive, miniaturized, low-mass and non-invasive instruments to monitor the movement and performance of athletes, patients, etc. in sporting or more natural environments and provide near real-time feedback to subjects. These new technologies are sufficiently accurate when compared with optical and video systems [11]. Micro electromechanical systems (MEMS) based inertial sensors including accelerometers and gyroscopes are good examples of using micro technology to monitor, classify and measure various human activities. Wireless/Wearable Inertial Measurement Units (WIMUs) are capable of tracking rotational and translational movements and are gaining in popularity to monitor human movements in a number of sport training [12], rehabilitation [13] and everyday activities [14]. This allows a subject's activities be to continuously monitored and subsequently corrected outside clinical environments. With the recent development of more accurate and relatively cheap WIMUs, combined with improved algorithms to more accurately determine sensor orientation [15], [16] , it has become feasible to deploy wearable body sensor networks in training sessions. Some commercially available systems include X-IMU (http://www.x-io.co.uk), Xsens (www.xsens.com) and Shimmer (www.shimmersensing.com). If WIMUs are to be used in this context, data processing time must be very short and user involvement minimized. This requires a system to automatically and accurately categorize each foot-ground contact based on the type of movement of the user (i.e. walk, jog, sprint, jump, land, agility cut). To the best of the authors' knowledge, such a system is not currently available. Even with low trial numbers, there are a number of challenges associated with comparing data (which are amplified with the larger trial numbers potentially possible with WIMUs):

1. Continuous data (e.g. joint angle) are usually reduced to a single/few discrete measure(s) that purportedly represent a joint's movement technique (e.g. peak flexion), but in reality comprises less than 2% of the available data [17], [18].
2. Continuous data (e.g. angle-time data) contains phase and amplitude variations both between individuals (inter-subject) and within multiple trials by the same individual (intra-subject).

Traditionally normative data is produced by time normalizing a trial to 101 data points and averaging across trials (e.g. mean $\pm 95\%$ confidence intervals) [10]. However, this may result in a distortion of the data as key events are not time aligned across trials [18].

These last two challenges can potentially be addressed using continuous data analysis techniques (e.g. functional data analysis [19]), although only a small number of biomechanical studies have attempted to do so [20], [17]. The aim of this study is to utilize wearable inertial sensors and develop a method to:

- Automatically and accurately categorize each foot-ground contact based on the type of movement (i.e. walk, jog, sprint, jump, land, agility cut);
- Extract joint angle and impact acceleration data automatically for each foot contact cycle;
- Generate normative data; for joint angle data using a functional data approach, for impact acceleration data using a discrete data point[3];
- Compare an individual to the normative data and identify the phase over which they differ (if any).

This manuscript represents a substantive extension to previous experimental work [1]. In particular, it extends this work by examining hip joint angle and sacrum impact acceleration, as well as comparing and contrasting different methods for movement classification.

II. Proposed Framework

The main components of our framework are illustrated in Fig.1. It consists of three main components: (i) activity classification, (ii) peak impact acceleration identification and calculation of sensor orientation and flexion-extension knee and hip angle, and (iii) technique analysis. We present results only in relation to sacrum impact accelerations as well as knee and hip flexion-extension angles in order to exemplify the process and avoid unnecessary repetition in this paper, but this method can be extended to other variables (e.g. knee valgus-varus, tibia impact accelerations) and other measures (e,g, joint angular velocity).

A. Activity Classification

Automatic activity classification is used to identify different training activities as this would allow training sessions to be more quickly evaluated by sporting and health professionals. It would allow them to quickly segment an athlete's training session by activity and thus allow the desired data to be more easily located. This approach also facilitates the creation of a database containing the evolution of an athlete's movements within and across training sessions.

In this work four different classifiers were investigated in order to create the most accurate classification system. The classifiers employed were Lazy IBk, RBF Network, Naive Bayes and Random Forest. Lazy IBk is a k nearest neighbour (k-NN) classifier. k-NNs has been shown to perform well in human activity problems [21], [22]. An RBF Network is an Artificial Neural Network (ANN) classifier and this family of classifiers has been very successful at discriminating between different human activities [23], [24]. Naive Bayes is a Bayesian classifier which has been used in a wide array of classification problems since the 1990s [25], [26]. Finally, a Random Forest is a relatively new decision tree classifier but has strong theoretical foundations and has been successfully used in recognizing human movements. [27], [27]. All four classifiers are investigated in this work.

Much of the prior research in activity classification has dealt with identifying mundane tasks such as eating, ascending and descending stairs, sitting, brushing teeth as well as motion activities such as being stationary, walking and running [28], [29] and training exercises and sports activities [30], [31].

[3]Only a discrete data point (i.e. peak impact acceleration) was analysed because there is currently no research relating the pattern of the whole deceleration signal to injury.

2327-4662 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2014.2377238, IEEE Internet of Things Journal

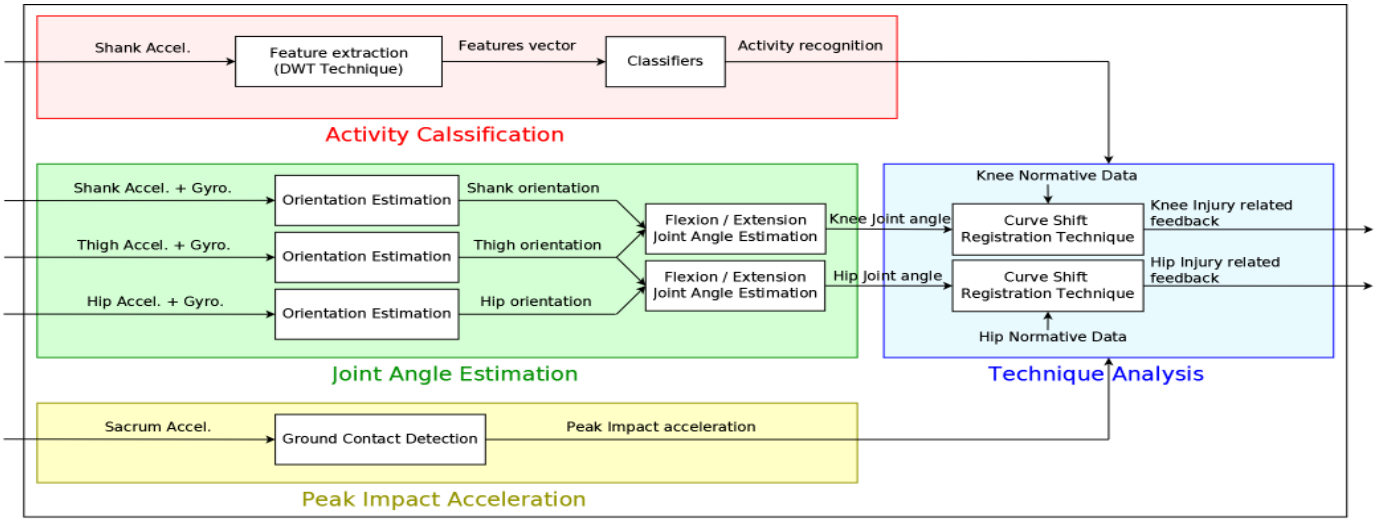

Fig. 1: The main components of the proposed motion analysis framework

Current research has shown that accelerometers can be used to classify human activity for high energy actions such as sprinting, jogging, jumping, etc [32]. In sports, accelerometers have been used to monitor elite athletes in competition and training environments. In swimming applications, accelerometers have allowed the comparison of stroke characteristics for a variety of training strokes and therefore have helped to improve swimming technique [33]. In competitive rowing, they have been used for the recovery of intra- and inter-stroke phases as a means to assess technique [34]. Accelerometers have also been utilized to identify the various phases of kinematic chain during the serve action in tennis [35].

In developing our approach to activity classification, the exercise routine performed by each athlete was first segmented and annotated for all activities and used to create a training set. A window length of three seconds was then chosen as this was sufficient time for each of the selected training activities to be completed. The Discrete Wavelet Transform (DWT) has been used with much success in extracting discriminative features from accelerometer data as the basis for classification. The Wavelet transform works by decomposing a signal into a number of time shifted and scaled versions of a selected mother wavelet. These X,Y,Z vectors have been used to assist in identifying sporting activities in soccer and field hockey [36]. Daubechies 4 wavelet "db4" is a popular mother wavelet choice in signal analysis problems due to its regularity and fast computational time, and was chosen in this work. The outputted coefficients produced by the DWT can be further decomposed to further increase the frequency resolution. Each additional decomposition increases the level i by one. The total energy E_T at level i of the DWT decomposition is given by [28].

$$E_T = A_i A_i^T + \sum_{j=1}^{i} D_j D_j^T \tag{1}$$

where A_i is the approximation coefficient at level i, A_i^T is the transpose of A_i and D_i is the detailed coefficient at level i. One feature proven to be useful in discrimination is the energy

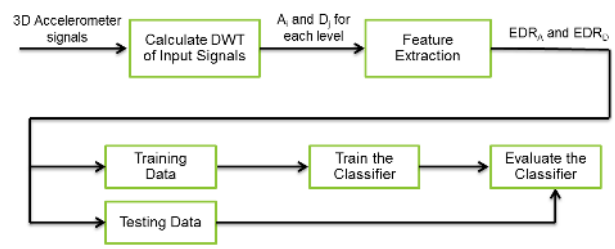

Fig. 2: The overview of the DWT decomposition and classification process.

ratio in each type of coefficient [28]. EDR_A represents the energy ratio of the approximation coefficients while EDR_{Dj} represents the energy ratio of the detail coefficients.

$$EDR_A = \frac{A_i A_i^T}{E_T} \tag{2}$$

$$EDR_{D_j} = \frac{D_i D_j^T}{E_T} \qquad j = 1, \ldots, i \tag{3}$$

In [28], Ayrulu-Erdem and Barshan found that the normalized variances of the DWT decomposition coefficients and the EDRs provided the most informative features for a different albeit similar problem. They contrasted their performance to informational features such as normalized means, minimums and maximums of the EDRs and obtained superior performance. As such we adopt the same approach here. The variances of the coefficients are calculated over each DWT coefficient vector at the ith level. A random forest training algorithm in conjunction with the DWT features was employed to create an appropriate classifier. The overview of the DWT decomposition and classification process is illustrated in Fig.2.

B. Sensor Orientation and Joint Angle Estimation

Measuring accurate orientation plays an important role in sports activity applications as it enables coaches, biomechanists and sports scientists to monitor and investigate athletes'

2327-4662 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2014.2377238, IEEE Internet of Things Journal

movement technique in indoor and outdoor environments. Although there are different technologies to monitor athletes' technique and measure their body orientation, wearable inertial sensors have the advantage of being self-contained in a way that measurement is independent of motion, environment and location. It is feasible to measure accurate orientation in three-dimensional space by utilizing tri-axial accelerometers, gyroscopes and a proper filter.

The Kalman filter has widely been utilized to measure orientation for many applications and commercial inertial orientation sensors, including Xsens and Intersense [37], [38]. However, it has some disadvantages including implementation complexity [39], [40], high sampling rate due to linear regression iteration (fundamental to the Kalman process) and the requirement to deal with large scale vectors to describe rotational kinematics in three-dimensions [38], [16]. There are some other alternatives to address these issues including Fuzzy processing [41] or frequency domain filters [42]. Although these approaches are easy to implement, they are limited to operating conditions. In this paper, we use an algorithm which has been shown to provide effective performance at low computational expense. Using such a technique, it is feasible to have a lightweight, inexpensive system capable of functioning over an extended period of time.

The algorithm employs a quaternion representation of orientation and is not subject to the problematic singularities associated with Euler angles. The estimated orientation rate is defined in the following equations [15]:

$$\begin{cases} {}^S_E q_t = {}^S_E q_{t-1} + {}^S_E \dot{q}_t \Delta t \\ {}^S_E \dot{q}_t = {}^S_E \dot{q}_{\omega,t} - \beta \frac{\nabla f}{||\nabla f||} \end{cases} \tag{4}$$

where

$$\begin{cases} \nabla f({}^S_E q, E_g, S_a) = J^T({}^S_E q, E_g) f({}^S_E q E_g, S_a) \\ S_a = [0, a_x, a_y, a_z] \\ E_g = [0, 0, 0, 1] \end{cases} \tag{5}$$

In this formulation, ${}^S_E q_t$ and ${}^S_E q_{t-1}$ are the orientation of the Earth frame relative to the sensor frame at time t and $t-1$ respectively. ${}^S_E \dot{q}_{\omega,t}$ is the rate of change of orientation measured by the gyroscopes. S_a is the acceleration in the x, y and z axes of the sensor frame, termed a_x, a_y, a_z respectively. The algorithm calculates the orientation ${}^S_E q_t$ by integrating the estimated rate of change of orientation measured by the gyroscope. Then gyroscope measurement error, β, was removed in a direction based on accelerometer measurements. This algorithm uses a gradient descent optimization technique to measure only one solution for the sensor orientation by knowing the direction of the gravity in the Earth frame. f is the objective function and J is its Jacobean (J^T is transpose of J) and they are defined by the following equations:

$$f(q, S_a) = \begin{bmatrix} 2(q_2 q_4 - q_1 q_3) - a_x \\ 2(q_1 q_2 + q_3 q_4) - a_y \\ 2(0.5 - q_2^2 - q_3^2) - a_z \end{bmatrix} \tag{6}$$

$$J(q) = \begin{bmatrix} -2q_3 & 2q_4 & -2q_1 & 2q_2 \\ 2q_2 & 2q_1 & 2q_4 & 2q_3 \\ 0 & -4q_2 & -4q_3 & 0 \end{bmatrix} \tag{7}$$

It is common to quantify orientation sensor performance as the static and dynamic RMS (Root-Mean-Square) errors [16]. The static RMS values of the pitch and roll components of an orientation using the described technique are $0.594°$ and $0.497°$, respectively. The dynamic RMS values of the pitch and roll components of an orientation are $0.623°$ and $0.628°$, respectively [15]. Therefore, the algorithm achieves levels of accuracy matching that of the Kalman based algorithm [15].

Typically a joint rotation is defined as the orientation of a distal segment with respect to the proximal segment. In order to measure flexion-extension joint angle, the orientation of the two wearable inertial sensors attached on the distal segment and the proximal segment were calculated using the described fusion algorithm. The alignment of each sensor unit's frame with the body frame was done using a functional calibration described in [43], [44]. A technique described in [45] was then applied to the shank and thigh segments to measure flexion-extension knee joint angle and to the thigh and pelvis segments to measure flexion-extension hip joint angle. This is described by the following equations:

$$\begin{cases} q_{knee} = {}^S_E q^*_{thigh} \otimes {}^S_E q_{shank} \\ q_{hip} = {}^S_E q^*_{pelvis} \otimes {}^S_E q_{thigh} \end{cases} \tag{8}$$

where ${}^S_E q_{pelvis}$, ${}^S_E q_{thigh}$ and ${}^S_E q_{shank}$ are the quaternion representation of the orientation of the pelvis, thigh and shank respectively. The $\otimes$ denotes the quaternion product and $*$ denotes the quaternion conjugate. The knee and hip joint angles were measured during the entire training session. The results are illustrated and discussed in section III-C.

C. Technique Analysis

The exercise reported in detail in this section is the jogging task. This was selected because it incorporates three activities that can make up most actions: an impact (with the ground), a loading phase and a swing (unloaded) phase. The jogging task was extracted based on the information given by the classification approach reported above. Foot contact cycles (heel strike to heel strike) were subsequently identified using knee joint angles and tibia acceleration. Heel strike was defined as the sudden change in acceleration after every cyclic local maximum in knee joint angle data (i.e. the swing phase). Subsequently, knee and hip joint angles were extracted. The separated knee and hip joint angle curves demonstrated similar patterns across trials and athletes, which as expected differed in their temporal characteristics. To maintain all the information of the curve shapes (magnitude and timing of local maxima and minima) the normative (representative) curve was created using two approaches: (a) averaging across the foot contact cycle without registration (unregistered curve), which is the most common approach in biomechanics [46], [47] and (b) performing a phase shift registration approach before averaging across the foot contact cycles as described by the following equations [19]:

2327-4662 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2014.2377238, IEEE Internet of Things Journal

$$x_i^*(t) = x_i(t + \delta_i) \tag{9}$$

$$\begin{aligned} SSE &= \sum_{i=1}^{N} \int_{\tau} ([x_i(t+\delta_i) - \hat{\mu}(t)]^2 ds) \\ &= \sum_{i=1}^{N} \int_{\tau} ([x_i^*(t+\delta_i) - \hat{\mu}(t)]^2 ds) \ . \end{aligned} \tag{10}$$

The phase shift registration alters the time domain by δ_j for each waveform x within a foot contact cycle i for multiple δ_j to find the δ_j where a registration criteria is at its minimum [19].

The used criterion (squared standard error; SSE) was calculated for each waveform relative to the overall mean $\hat{\mu}(t)$ over its specific time interval t. This process was applied for every foot contact cycle to identify the optimal δ_j for each foot contact cycle i. Subsequently, these curves were registered using the optimal δ_j. After all waveforms were registered, the overall mean was updated and the whole process was iterated n times until no significant change ($SSE_{n-1} \ll SSE_n \approx SSE_{n+1}$) in the registration criteria occurred. This procedure of estimating a transformation by transforming to an iteratively updated average is often referred to as the Procrustes method [19][4]. To examine if differences exist between the mean curve and the registered mean curve, we examined the curves using Analysis of Characterizing Phases [17]. This approach offers a more comprehensive comparison than discrete point analysis or functional principal component analysis, as it identifies phases of variation of the data that are subsequently used to generate subject scores.

To explore the ability of the proposed process to identify individuals with abnormal movement biomechanics, an individual with low back pain was also assessed. Clinical differences were explored both visually and statistically. Statistical differences were identified by examining the boundaries of the confidence intervals of the single athlete with the 95% confidence intervals of the normal group data. Waveforms were considered statistically different when the confidence intervals did not overlap [48]).

III. Experiments and Evaluation

A. Data Collection

To evaluate the proposed framework, actions of nine healthy subjects and one injured subject with low back pain were captured using six wearable inertial sensors. Subjects with different levels of skill proficiency were chosen for this study in order to provide a wide range of variations (i.e. speeds, movement techniques, etc.) to examine the framework. WIMUs were placed on the left/right shank, left/right thigh and pelvis and sacrum of a subject as shown in Fig.3. The location of the sensor on each body segment was chosen to avoid large muscles; as soft tissue deformations due to muscle contractions and foot-ground impacts may negatively affect impact accelerations and the accuracy of joint orientation estimates. The sensors were affixed to the subject with double sided tape and velcro straps with some elasticity in the fabric, so as not to restrict the subject's movement and performance. Next, the subject was asked to perform a series of actions as they normally do during outdoor training sessions. Each subject performed a predefined exercise routine on a large outdoor grass soccer pitch. The exercise routine consisted of the following motions: agility cuts, walking, sprinting, jogging, box jumps and football free kicks. Each motion lasted approximately 60 seconds for a total of approximately $9-10$ minutes for the entire session.

[4]The reader should note that for some biomechanical data (or waveforms) phase shift registration might not lead to a representative curve shape. For such cases a dynamical time warping approach can be applied, which uses specific landmarks (global maxima and minima) to define a warping function h to which the waveforms are evaluated (Equation $x_i^*(t) = x_i[h_i(t)]$)

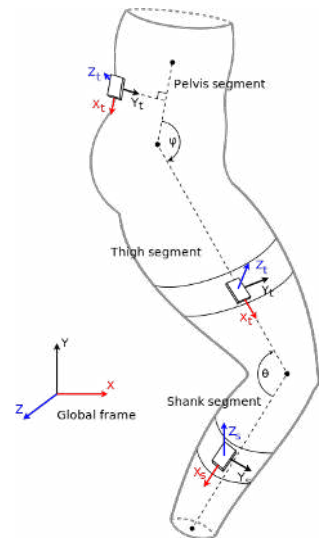

Fig. 3: Placement of three inertial sensor units on the pelvis, thigh and shank as well as their local coordinate system in a global coordinate system is illustrated.

The data from each sensor was recorded to an internal SD card on board the device. As each sensor recorded data independently, a physical event was required to synchronize all devices together. This was achieved by instructing each subject to perform five vertical jumps, ensuring large acceleration spikes would occur simultaneously on each device, that would be clearly visible in the accelerometer stream. In a post processing step, peak alignment was automatically performed and all data streams were cropped to two seconds before the first vertical jump landing. Video footage of each data capture session was also recorded and annotated, to be used as ground truth for the automatic segmentation and recognition of movement categories (i.e. jogging, agility cuts, sprinting etc.).

B. Classification Evaluation

Table I shows the time taken to train and test different classification models. Each classifier was very quick at testing instances of data however the Multilayer Perceptron took almost 10 mins to train a single model. This creates a significant bottleneck when wishing to compare different classification parameters. The Random Forest allowed data to be classified almost instantaneously which is extremely desirable in real world scenarios. Each classifier's ability to distinguish between the different activities in the Dataset is shown in Fig.4.

2327-4662 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2014.2377238, IEEE Internet of Things Journal

TABLE I: Time taken to train and test different classifiers is illustrated. In this table RF, NB and MP stand for Random Forest, Naiive Bayes and Multilayer Perceptron respectively

Classifiers	RF	NB	Lazy iBK	MP
Testing Time(s)	0.00	0.15	0.38	0.08
Training Time(s)	0.38	0.10	0.00	538.9

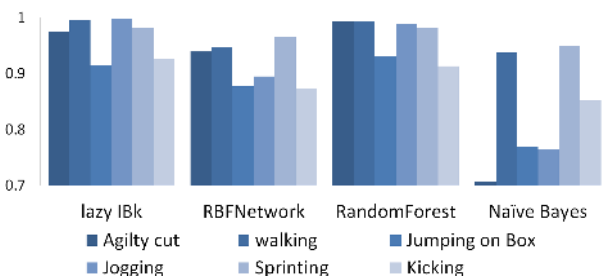

Fig. 4: Performance of different classifiers to classify various activities is illustrated.

Overall the random forest classifier was the most accurate. Walking is a very low energy activity and therefore each classifier was well able to distinguish it from the other activities. Similarly, the agility cut while a high energy activity like sprinting and kicking, was recognized by most classifiers as it involved very distinct frequency and movement patterns. Every activity had distinctive features that allow each classifier to differentiate between them. Due to the efficiency of the DWT a relatively small amount of features were inputted into each classifier making the classification process very quick. Although other classifiers were investigated, the Random Forest achieved the highest classification accuracy within acceptable computational limits. Using the approach described in section II-A, we achieved a classification accuracy of 98.3% utilizing the Random Forest classifier in conjunction with the DWT technique. This value was computed using a ten-fold cross validation leave one out method. The F-measure score, as a harmonic mean of precision and recall that reaches its best value at 1 and worst score at 0, was calculated. Precision is calculated as the number of correct results divided by the number of total results while recall is the number of correct results divided by the number of results that should have been returned positive. These metrics are often described in terms of the metrics true positive (T_p), false positive (F_p) and false negative (F_n). Since the classifier was trained with classes which had different instance populations the F-measure scores are given in table III. The F-measure score gives a better indication of a models ability to correctly identify an activity than standard classification accuracy alone.

Table II shows the confusion matrix from the classification procedure. There is only one area of confusion using this model which is kicking the football. This difficultly lies with the variation in kicking styles from person to person. As can be seen in Table III, F-measures vary between 0.913 to 0.992. Walking and agility cut have the highest F-measures followed by jogging, sprinting, jumping on the box and football kicks.

TABLE II: Confusion matrix for the Random Forest classifier

Activity	a	b	c	d	e	f
a = Agility cut	180	0	0	0	0	0
b = Walking	0	399	0	0	0	0
c = Jumping on box	0	0	27	2	0	0
d = Jogging	0	0	0	205	0	0
e = Sprinting	0	0	0	0	28	0
f = Kicking	3	5	2	3	1	73

TABLE III: The precision, recall and F-measure from the Random Forest classifier applied to the activities during training sessions.

Activity	Precision	Recall	F-measure
Agility cut	0.984	1	0.992
Walking	0.988	1	0.994
Jumping on Box	0.931	0.931	0.931
Jogging	0.976	1	0.988
Sprinting	0.966	1	0.982
Kicking	1	0.839	0.913

C. Technique Evaluation

In the simulated training intervention the subjects were asked to jog for one minute where about 30 foot contact cycles could be identified for each subject. It can be seen in Fig.5 and Fig.7 that the generated knee angle curves show the classic bimodal shape, with a small ($0 - 35\%$ cycle) and large ($35 - 90\%$ cycle) sequencing of flexion-extension. The statistical analysis of the knee angle curves indicated significant differences between the unregistered and the registered mean curves. The unregistered mean curve demonstrated significantly higher ($p = 0.002$) and lower magnitudes ($p < 0.001$), for ($11 - 17\&87 - 96\%$) and ($61 - 75\%$) of the foot contact cycle, respectively. For the hip angle, Figures .6 and .8 show the classic extension-flexion sequencing during the stance and early swing phases ($0{-}75\%$) followed by a smaller extension-flexion sequencing during the later swing phases ($75 - 100\%$).

As shown in Fig.6, the statistical analysis for the hip angle curves also indicated significant differences between the unregistered and the registered mean curves. The unregistered mean curve was significantly lower ($p < 0.010$) magnitudes, for ($14 - 22\%$) and ($74 - 86\%$) of the foot contact cycle. Differences are similarly evident at an intra-subject level, for both the knee and hip joints.

For the knee angle, it can be seen that in the first phase ($1{-}40\%$) of the examined foot contact cycle the registered and unregistered curves are very similar (except for the magnitudes between $10 - 20\%$). However, for phases beyond 40%, both mean curves start to show differences in magnitude, timing characteristics and standard deviation. For the hip joint the registered and unregistered curves show, as for the knee joint, similarities for most phases but differ clearly for phases before

2327-4662 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2014.2377238, IEEE Internet of Things Journal

and after local maximas. The magnitude and position of the peak hip flexion (approx. 80%) demonstrates clearly the effect of intra-subject variability of the movement cycle, which affects the mean curve. By solely averaging the foot contact cycles (unregistered approach), the generated mean curve is altered by the intra- or/and inter-subject variability and can lose very valuable information about the subject. This can be extremely important in injury studies, where small differences from normal healthy subjects or small intra-subject differences over time may indicate a predisposition to injury or the early stages of injury, requiring the implementation of an appropriate training intervention. The more complicated or oscillating the collected biomechanical data, the more important it is likely to register the data, especially if derivatives are examined (i.e. joint angular velocities).

It can be seen in Fig.7 and Fig.8 that the runner with low back pain exhibited clear differences from the normative data, for both the knee and hip joint angles. In normal subjects, the knee generally flexes during initial loading $(0-10\%)$ and early mid stance $(10-15\%)$ while in the injured subject it clearly extends. The initial loading response involves the bi-articular hamstring muscle acting concentrically to extend the hip to keep the trunk upright, Fig.8, and as a consequence

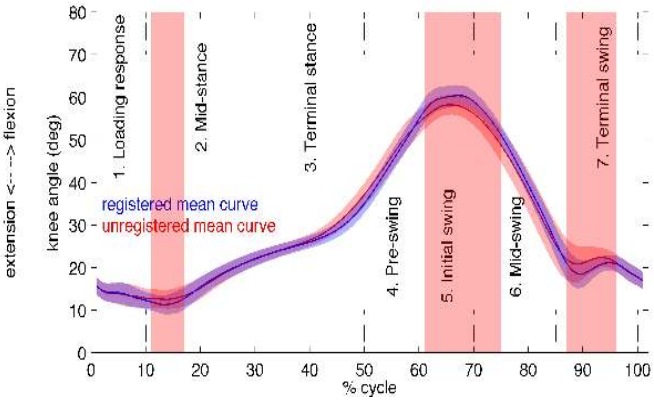

Fig. 5: Registered and unregistered mean knee curves of an injured subject.

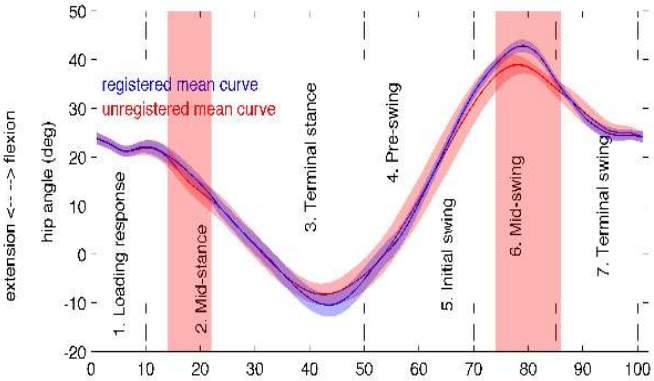

Fig. 6: Registered and unregistered mean hip curves of an injured subject.

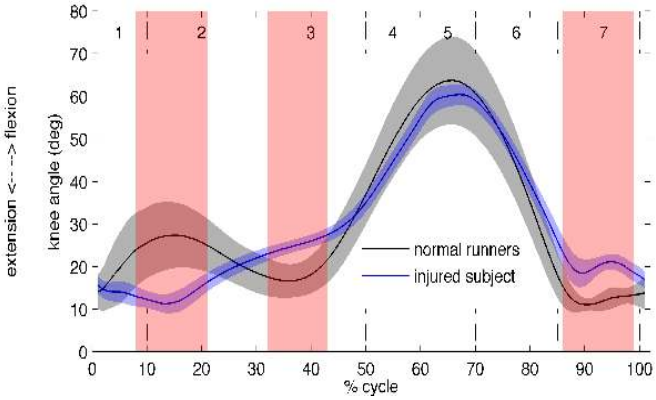

Fig. 7: Registered mean knee curves of an injured subject in comparison to the overall registered mean of normal subjects. Phases of statistically significant differences are indicated by the vertical bands.

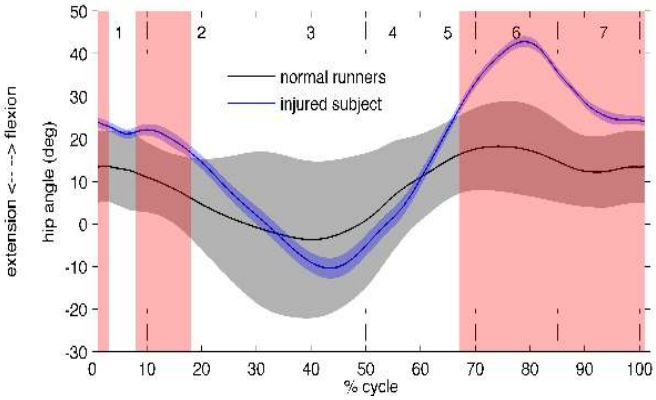

Fig. 8: Registered mean hip curves of an injured subject in comparison to the overall registered mean of normal subjects. Phases of statistically significant differences are indicated by the vertical bands.

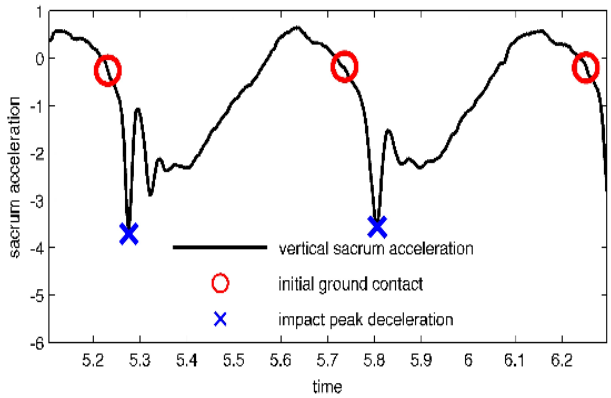

Fig. 9: Illustrates a typical sacrum acceleration curve and key events: o initial ground contact and x peak impact acceleration

of the hamstring also being a knee flexor muscle, this results in knee flexion. Therefore, the abnormal knee extension in the injured subject appears to indicate either a compensatory

2327-4662 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2014.2377238, IEEE Internet of Things Journal

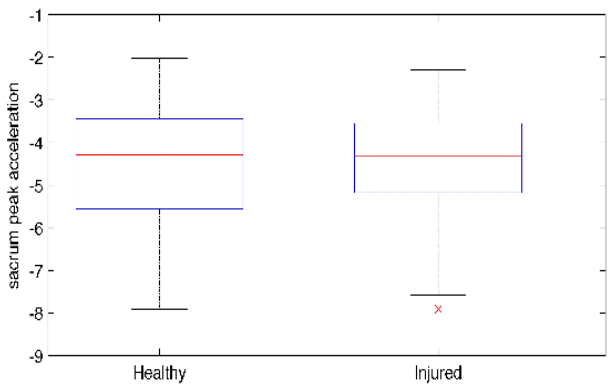

Fig. 10: Boxplot (mean +/- 95% CI) of the peak impact acceleration for the healthy (n = 9) and injured subject(s) (n = 1)

or injury causing movement strategy indicative of the trunk inappropriately flexing during the initial loading response. This is supported by greater hip flexion angles at and post heel strike Fig.8. From a compensatory perspective, this may be a strategy to reduce lower back impact loading with the trunk extensors acting eccentrically to cushion the action. Possibly in response to the abnormal early knee extension, knee flexion is initiated much earlier in the injured subject (at 15% of the cycle) compared to normal $(at\ 35\%)$. The greater knee flexion in the injured subject during the terminal swing phase $(85-100\%)$ combination with greater hip flexion $(68-100\%)$ may be indicative of a crouched ("Groucho") running style aimed at reducing impact loads and hence reducing back pain and further injury [49], [50].However, in contrast to previously reported crouched running strategies [49], [50], our injured athlete did not increase their knee flexion during early/mid stance, they actually extended their knee more than the uninjured athletes. Peak impact accelerations of the sacrum were automatically extracted by identifying the smallest magnitude within a 70 frames $(0.28sec)$ after initial ground contact. The sacrum impact accelerations (mean [upper to lower 95% CI]; -4.6 [-4.5 to -4.7], Fig.9 and Fig.10) were comparable with previous reported values [51]. The injured athletes values (mean [upper to lower 95% CI]; -4.5 [-4.1 to -4.9], Fig.10) did not differ from those of the unijured group. Given that these impact accelerations represent spinal axial loading (level with the sacram) would imply that the athletes injury is vertebrae control based (i.e. muscle, tendon, ligament, neural) rather than vertebrae (i.e. bone) or vertebrae support (i.e cartilage, itervertabrae disc) based.

IV. Conclusion

In this paper, we described a novel body worn inertial sensor framework capable of automatically segmenting and classifying various actions in outdoor unconstrained environments, extracting sacrum peak impact accelerations, and calculating extension-flexion knee and hip joint angles that uses continuous data analysis to both generate accurate normative data and compare individuals to this normative data. The proposed novel framework employed a Random Forest training algorithm in conjunction with a DWT feature extraction technique to successfully classify training session activities with up to 98% overall accuracy. Using the body-worn inertial sensors on the sacrum, thigh and shank of and applying the gradient descent based filter, the local orientation of each sensor and associated body segment orientation were estimated and hence the extension-flexion knee and hip angles were obtained. The calculated knee and hip joint angles, along with the sacrum impact accelerations were input to a data analysis tool at the end of the pipeline to provide accurate movement technique assessment. In examining the continuous joint angle data it is necessary to register the trials before averaging them to ensure the true magnitude and shape of the data is preserved for both group and individual based data. If this is ensured, the presented framework has significant potential for monitoring athletes throughout training and competition to (a) identify injury and performance related determining factors, (b) identify individuals early in an injury pathway prior to extensive tissue damage, and (c) identify individuals predisposed to injury because of their movement technique.

Acknowledgment

The research leading to these results has received funding from the European Community's Seventh Framework Programmes (FP7/2013-2016) under grant agreement no. ICT-2011-8.2.601170 (REPLAY project) and (FP7/2007-2013) under grant agreement no. ICT-2011-7-287723 (REVERIE project).This work was also carried out with the financial support of Science Foundation Ireland (SFI) under grant number SFI/12/RC/2289.

References

[1] A. Ahmadi, E. Mitchell, F. Destelle, M. Gowing, N. O'Connor, C. Richter, and K. Moran, "Automatic activity classification and movement assessment during a sports training session using wearable inertial sensors," in *Body Sensor Networks (BSN), 2014 IEEE International Conference on*. IEEE, 2014, pp. 98–103.

[2] J. L. Dragoo, H. J. Braun, J. L. Durham, M. R. Chen, and A. H. Harris, "Incidence and risk factors for injuries to the anterior cruciate ligament in national collegiate athletic association football data from the 2004-2005 through 2008-2009 national collegiate athletic association injury surveillance system," *The American Journal of Sports Medicine*, vol. 40, no. 5, pp. 990–995, 2012.

[3] A. Hreljac, "Impact and overuse injuries in runners," *Medicine and Science in Sports and Exercise*, vol. 36, no. 5, pp. 845–849, 2004.

[4] R. Van Gent, D. Siem, M. van Middelkoop, A. Van Os, S. Bierma-Zeinstra, and B. Koes, "Incidence and determinants of lower extremity running injuries in long distance runners: a systematic review," *British journal of sports medicine*, vol. 41, no. 8, pp. 469–480, 2007.

[5] B. T. Zazulak, T. E. Hewett, N. P. Reeves, B. Goldberg, and J. Cholewicki, "Deficits in neuromuscular control of the trunk predict knee injury risk a prospective biomechanical-epidemiologic study," *The American journal of sports medicine*, vol. 35, no. 7, pp. 1123–1130, 2007.

[6] K. A. Moran and B. M. Marshall, "Effect of fatigue on tibial impact accelerations and knee kinematics in drop jumps," *Medicine and science in sports and exercise*, vol. 38, no. 10, pp. 1836–1842, 2006.

[7] K. A. Moran, M. Clarke, F. Reilly, E. S. Wallace, D. Brabazon, and B. Marshall, "Does endurance fatigue increase the risk of injury when performing drop jumps?" *The Journal of Strength & Conditioning Research*, vol. 23, no. 5, pp. 1448–1455, 2009.

[8] T. R. Derrick, "The effects of knee contact angle on impact forces and accelerations." *Medicine and science in sports and exercise*, vol. 36, no. 5, pp. 832–837, 2004.

2327-4662 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2014.2377238, IEEE Internet of Things Journal

[9] J. Iguchi, H. Tateuchi, M. Taniguchi, and N. Ichihashi, "The effect of sex and fatigue on lower limb kinematics, kinetics, and muscle activity during unanticipated side-step cutting," *Knee Surgery, Sports Traumatology, Arthroscopy*, pp. 1–8, 2013.

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

[11] A. Ahmadi, D. D. Rowlands, and D. A. James, "Development of inertial and novel marker-based techniques and analysis for upper arm rotational velocity measurements in tennis," *Sports Engineering*, vol. 12, no. 4, pp. 179–188, 2010.

[12] H. Ghasemzadeh, V. Loseu, E. Guenterberg, and R. Jafari, "Sport training using body sensor networks: A statistical approach to measure wrist rotation for golf swing," in *Proceedings of the Fourth International Conference on Body Area Networks*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2009, p. 2.

[13] K. Culhane, M. OConnor, D. Lyons, and G. Lyons, "Accelerometers in rehabilitation medicine for older adults," *Age and ageing*, vol. 34, no. 6, pp. 556–560, 2005.

[14] H. Junker, O. Amft, P. Lukowicz, and G. Tröster, "Gesture spotting with body-worn inertial sensors to detect user activities," *Pattern Recognition*, vol. 41, no. 6, pp. 2010–2024, 2008.

[15] S. O. Madgwick, A. J. Harrison, and R. Vaidyanathan, "Estimation of imu and marg orientation using a gradient descent algorithm," in *Rehabilitation Robotics (ICORR), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1–7.

[16] A. M. Sabatini, "Quaternion-based extended kalman filter for determining orientation by inertial and magnetic sensing," *Biomedical Engineering, IEEE Transactions on*, vol. 53, no. 7, pp. 1346–1356, 2006.

[17] C. Richter, N. O'Connor, B. Marshall, and K. Moran, "Analysis of characterizing phases on waveforms-an application to vertical jumps." *Journal of applied biomechanics(in press)*, 2013.

[18] E. A. Crane, R. B. Cassidy, E. D. Rothman, and G. E. Gerstner, "Effect of registration on cyclical kinematic data," *Journal of biomechanics*, vol. 43, no. 12, pp. 2444–2447, 2010.

[19] J. O. Ramsay, *Functional data analysis*. Wiley Online Library, 2006.

[20] S. Ullah, C. F. Finch *et al.*, "Applications of functional data analysis: A systematic review," *BMC medical research methodology*, vol. 13, no. 1, p. 43, 2013.

[21] Z. Zhou, X. Chen, Y.-C. Chung, Z. He, T. X. Han, and J. M. Keller, "Activity analysis, summarization, and visualization for indoor human activity monitoring," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 18, no. 11, pp. 1489–1498, 2008.

[22] F. Foerster, M. Smeja, and J. Fahrenberg, "Detection of posture and motion by accelerometry: a validation study in ambulatory monitoring," *Computers in Human Behavior*, vol. 15, no. 5, pp. 571–583, 1999.

[23] M. Ermes, J. Parkka, J. Mantyjarvi, and I. Korhonen, "Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 12, no. 1, pp. 20–26, 2008.

[24] D. R. Bassett Jr, B. E. Ainsworth, A. M. Swartz, S. J. Strath, W. L. O'Brien, and G. A. King, "Validity of four motion sensors in measuring moderate intensity physical activity." *Medicine and science in sports and exercise*, vol. 32, no. 9 Suppl, pp. S471–80, 2000.

[25] I. Kononenko, "Inductive and bayesian learning in medical diagnosis," *Applied Artificial Intelligence an International Journal*, vol. 7, no. 4, pp. 317–337, 1993.

[26] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," in *AAAI*, vol. 5, 2005, pp. 1541–1546.

[27] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[28] B. Ayrulu-Erdem and B. Barshan, "Leg motion classification with artificial neural networks using wavelet-based features of gyroscope signals," *Sensors*, vol. 11, no. 2, pp. 1721–1743, 2011.

[29] N. Alshurafa, W. Xu, J. J. Liu, M.-C. Huang, B. Mortazavi, M. Sarrafzadeh, and C. Roberts, "Robust human intensity-varying activity recognition using stochastic approximation in wearable sensors," in *Body Sensor Networks (BSN), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1–6.

[30] M. Gowing, A. Ahmadi, F. Destelle, D. S. Monaghan, N. E. OConnor, and K. Moran, "Kinect vs. low-cost inertial sensing for gesture recognition," in *MultiMedia Modeling*. Springer, 2014, pp. 484–495.

[31] A. Ahmadi, D. Rowlands, and D. A. James, "Towards a wearable device for skill assessment and skill acquisition of a tennis player during the first serve," *Sports Technology*, vol. 2, no. 3-4, pp. 129–136, 2009.

[32] X. Long, B. Yin, and R. M. Aarts, "Single-accelerometer-based daily physical activity classification," in *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*. IEEE, 2009, pp. 6107–6110.

[33] Y. Ohgi, H. Ichikawa, and C. Miyaji, "Microcomputer-based acceleration sensor device for swimming stroke monitoring." *JSME International Journal Series C*, vol. 45, no. 4, pp. 960–966, 2002.

[34] D. A. James, N. Davey, and T. Rice, "An accelerometer based sensor platform for insitu elite athlete performance analysis," in *Sensors, 2004. Proceedings of IEEE*. IEEE, 2004, pp. 1373–1376.

[35] A. Ahmadi, D. D. Rowlands, and D. A. James, "Investigating the translational and rotational motion of the swing using accelerometers for athlete skill assessment," in *Sensors, 2006. 5th IEEE Conference on*. IEEE, 2006, pp. 980–983.

[36] E. Mitchell, D. Monaghan, and N. E. O'Connor, "Classification of sporting activities using smartphone accelerometers," *Sensors*, vol. 13, no. 4, pp. 5317–5337, 2013.

[37] H. Luinge, P. Veltink, and C. Baten, "Estimating orientation with gyroscopes and accelerometers," *Technology and health care*, vol. 7, no. 6, pp. 455–459, 1999.

[38] J. L. Marins, X. Yun, E. R. Bachmann, R. B. McGhee, and M. J. Zyda, "An extended kalman filter for quaternion-based orientation estimation using marg sensors," in *Intelligent Robots and Systems, 2001. Proceedings. 2001 IEEE/RSJ International Conference on*, vol. 4. IEEE, 2001, pp. 2003–2011.

[39] H. Luinge and P. H. Veltink, "Measuring orientation of human body segments using miniature gyroscopes and accelerometers," *Medical and Biological Engineering and computing*, vol. 43, no. 2, pp. 273–282, 2005.

[40] M. Haid and J. Breitenbach, "Low cost inertial orientation tracking with kalman filter," *Applied Mathematics and Computation*, vol. 153, no. 2, pp. 567–575, 2004.

[41] S. K. Hong, "Fuzzy logic based closed-loop strapdown attitude system for unmanned aerial vehicle (uav)," *Sensors and Actuators A: Physical*, vol. 107, no. 2, pp. 109–118, 2003.

[42] R. A. Hyde, L. P. Ketteringham, S. A. Neild, and R. J. Jones, "Estimation of upper-limb orientation based on accelerometer and gyroscope measurements," *Biomedical Engineering, IEEE Transactions on*, vol. 55, no. 2, pp. 746–754, 2008.

[43] J. Favre, R. Aissaoui, B. Jolles, O. Siegrist, J. de Guise, and K. Aminian, "3d joint rotation measurement using mems inertial sensors: Application to the knee joint," *Valenciennes, France: ISB-3D*, 2006.

[44] A. G. Cutti, A. Ferrari, P. Garofalo, M. Raggi, A. Cappello, and A. Ferrari, "outwalk: a protocol for clinical gait analysis based on inertial and magnetic sensors," *Medical & biological engineering & computing*, vol. 48, no. 1, pp. 17–25, 2010.

[45] D. Roetenberg, H. Luinge, and P. Slycke, "Xsens mvn: full 6dof human motion tracking using miniature inertial sensors," *Xsens Motion Technologies BV, Tech. Rep*, 2009.

[46] I. Epifanio, C. Ávila, A. Page, and C. Atienza, "Analysis of multiple waveforms by means of functional principal component analysis: normal versus pathological patterns in sit-to-stand movement," *Medical & biological engineering & computing*, vol. 46, no. 6, pp. 551–561, 2008.

[47] O. A. Donoghue, A. J. Harrison, N. Coffey, and K. Hayes, "Functional data analysis of running kinematics in chronic achilles tendon injury." *Medicine and science in sports and exercise*, vol. 40, no. 7, pp. 1323–1335, 2008.

[48] J. Devore, *Probability and Statistics for Engineering and the Sciences*. Cengage Learning, 2011.

[49] T. A. McMahon, G. Valiant, and E. C. Frederick, "Groucho running," *Journal of Applied Physiology*, vol. 62, no. 6, pp. 2326–2337, 1987.

[50] H. P. Crowell, C. E. Milner, J. Hamill, and I. S. Davis, "Reducing impact loading during running with the use of real-time visual feedback," *journal of orthopaedic & sports physical therapy*, vol. 40, no. 4, pp. 206–213, 2010.

[51] J. Mizrahi, O. Verbitsky, and E. Isakov, "Shock accelerations and attenuation in downhill and level running," *Clinical Biomechanics*, vol. 15, no. 1, pp. 15–20, 2000.

2327-4662 (c) 2013 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.