<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Metrics for monitoring patients progress in a rehabilitation context: a case study based on wearable inertial sensors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>O'Flynn, Brendan; Dagini, Nektaria; Bottone, Sonia; Urru, Andrea; Tedesco, Salvatore</td>
</tr>
<tr>
<td><strong>Publication date</strong></td>
<td>2018-10</td>
</tr>
<tr>
<td><strong>Type of publication</strong></td>
<td>Conference item</td>
</tr>
<tr>
<td><strong>Link to publisher's version</strong></td>
<td><a href="https://www.iaria.org/conferences2018/HEALTHINFO18.html">https://www.iaria.org/conferences2018/HEALTHINFO18.html</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>© IARIA, 2018.</td>
</tr>
<tr>
<td><strong>Item downloaded from</strong></td>
<td><a href="http://hdl.handle.net/10468/7272">http://hdl.handle.net/10468/7272</a></td>
</tr>
</tbody>
</table>

Downloaded on 2020-09-26T19:47:50Z
Metrics for Monitoring Patients Progress in a Rehabilitation Context
A Case Study based on Wearable Inertial Sensors

Brendan O’Flynn, Nektaria Dagini, Sonia Bottone, Andrea Urru, Salvatore Tedesco
Tyndall National Institute, University College Cork
Lee Maltings, Prospect Row, Cork, T12R5CP, Ireland
email: brendan.oflynn / nektaria.dagini / sonia.bottone / andrea.urru / salvatore.tedesco@tyndall.ie

Abstract— Inertial data can represent a rich source of clinically relevant information which can provide details on motor assessment in subjects undertaking a rehabilitation process. Indeed, in clinical and sport settings, motor assessment is generally conducted through simple subjective measures such as a visual assessment or questionnaire given by caregivers. As part of a mobile health application, wireless sensors such as inertial measurement units and associated data sets can help provide an objective and empirical measure of a patient’s progress through rehabilitation using on body sensors. In this publication, several metrics in different domains have been considered and extrapolated from the 3D accelerometer and angular rate data sets collected on an impaired subject with knee injury, via a wearable sensing system developed at the Tyndall National Institute. These data sets were collected for different activities performed across a number of sessions as the subject progressed through the rehabilitation process. Using these data sets, a novel and effective method has been investigated in order to define a single score indicator which can provide accurate quantitative analysis of the improvement of the subject throughout their rehabilitation. The indicator compares impaired and unimpaired limb motor performance. The present work proves that the defined score indicator can be taken into account by clinicians to study the overall patients’ condition and provide accurate clinical feedback as to their rehabilitative progress.

Keywords- ACL; IMU; Inertial Sensors; Metrics; Motor Assessment; Rehabilitation; Wearables.

I. INTRODUCTION

Motor assessment is the part of biomechanics which studies the process by which the musculoskeletal system can create and control coordinated movements [1]. Voluntary movement requires the transmission of a message from the brain to the appropriate muscle which also controls the smoothness and coordination of the movement. If motor function is intact, muscles can be moved to command allowing symmetrical movements with significant strength levels. With particular reference to the treatment of patients with lower extremity injuries, literature has recently shown a paradigm shift, going from time-dependent concepts to function-based concepts [2], where qualitative and quantitative tests comparing affected and unaffected sides must be met before successfully accessing the following rehabilitation stage.

Qualitative and quantitative motor assessment is typically divided into clinimetrics, balance analysis, and gait analysis. Indexes, rating scales, questionnaires, and observational forms represent the clinical standard for knee joint assessment, including, for instance, Knee Injury and Osteoarthritis Outcome Score (KOOS), Oxford Knee Score (OKS), Tegner Lysholm Knee Scoring Scale, International Knee Documentation Committee (IKDC), Western Ontario & McMaster Universities Osteoarthritis Index (WOMAC) [3]. However, these tools are subjective and, even when utilised by experienced clinicians, may not be adequate or sensitive enough.

Gold-standard technology adopted in gait analysis for quantitative movement analysis may include camera-based motion analysis, instrumented treadmills, force platforms [4], but their application is constrained by costs, access to specialist motion labs, as well as practicality of application for larger patient/subject groups.

A viable alternative is represented by the adoption of small-size low-cost wearable sensing units whose consideration for lower-limbs monitoring during rehabilitation, in order to provide objective performance of impaired subjects throughout the process, has been growing lately. Indeed, inertial sensors, typically including accelerometers, gyroscopes, and magnetometers, have been used to derive gait parameters efficiently both in healthy and symptomatic subjects [5].

This paper describes a long-term investigation of post-injury rehabilitation carried out by using a wearable inertial system developed at the Tyndall National Institute, consisting of two sensors per limb, able to provide a complete biomechanics assessment for a series of scripted activities. The work is organised as follows. Hardware platform description and test protocol are described in Section II and III, respectively. The features extracted are illustrated in Section IV. The mathematical model is outlined in Section V. The obtained results are shown in Section VI. Finally, conclusions are drawn in the final section.

II. HARDWARE PLATFORM

The Tyndall biomechanical monitoring system consists of two Tyndall Wireless Inertial Measurement Units (WIMUs) per leg [6]-[9]. The platform measures 44 × 30 × 8 mm and 7.2 g without battery as shown in Figure 1.

The WIMU is equipped with a high-performance low-
power ARM Cortex-M4 32-bit microprocessor operating at a frequency up to 168 MHz part of the STM32F0407 family produced by STMicroelectronics. It also features a floating point unit single precision, high-speed embedded memories (1 Mb of Flash memory, 192 + 4 Kb of SRAM), an extensive range of enhanced I/Os and peripherals, and standard and advanced communication interfaces.

Figure 1. Tyndall Wireless Inertial Measurement Unit (WIMU).

Inertial sensors (3D accelerometer and gyroscope, MPU-9250 from Invensense) are the main sensing components on the platform and are wired to the microcontroller through the I2C communication. Sensor data can be transmitted wirelessly via a communication BLE-complaint module (Broadcom BCM20737S), with integrated ARM CM3 microcontroller unit, radio frequency and embedded Bluetooth Smart Stack, or logged to a removable Micro SD card with sampling rate of 250 Hz.

For measurement of inertial data, the Invensense MPU-9250 was chosen for its low power consumption and the high range (16g for accelerometer and 2000 deg/s for the gyroscope) with limited noise levels.

The platform also features a USB connector, battery charger, fuel gauge, external I/O connectors, three LEDs, and power switch. Even if not considered in the presented investigation, the platform could also provide additional sensing data, such as magnetic field (from the MPU-9250) and environmental data (pressure, humidity, temperature) from the Bosch BME280. All the components fit with mobile applications requirements and, averagely, the overall power consumption in TX/RX mode is 100 mA, dropping to 40 mA (17 mA) for stand-by (sleep mode).

III. PROTOCOL FOR DATA COLLECTION

In conjunction with clinical partners, an experimental protocol for data collection was developed to evaluate patient progress. The rehabilitation exercises considered are walking, half-squat, hamstring curl, and flexion-extension, defined by physiotherapists as indicators of rehabilitation.

These are described as follows:
- In the walking exercise, the subject walks on a calibrated treadmill, which is operated at defined speeds (3-4-6 km/h) for approximately one minute per test.
- In the half-squat exercise, the subject stands with their feet shoulder’s distance apart and arms crossed on the chest. Keeping the chest lifted, the hips are lowered about 10 inches, planting the weight in the heels. The body is then brought back up to standing by pushing through the heels.
- In the hamstring curl exercise, the subject stands and bends their knee, raising the heel toward the ceiling as far as possible without pain, relaxing the leg after each repetition. This is repeated on both legs.
- In the flexion-extension exercise, the subject lies supine on the floor and bends their knee raising it toward the chest as far as possible without pain, relaxing the leg after each repetition. This is repeated on both legs.

The system has been tested with an impaired subject. The impaired subject is a female athlete, age: 44, height: 161 cm, and weight: 52 kg, with good general health status, with a history of knee injuries and surgery (reconstructed anterior cruciate ligament in the left leg following a sporting injury). The tests were carried out during the course of the rehabilitation program, e.g., starting 1 month before surgery and finishing 7 months after surgery. Overall, the subject has been evaluated in 8 sessions through three periods: once in pre-surgery conditions (e.g., 1 month before surgery), then 6 times in a range of 20 weeks starting one month after surgery (namely short-term post-surgery), and finally once 3 months after the last data capture (e.g., during long-term post-surgery period).

A number of repetitions has been collected for each exercise, so as to provide an accurate picture of the overall conditions, and each exercise was evaluated during the majority of the data captures. Hamstring curl, as well as walking at 3 and 4 km/h, were performed at every session. Similarly, flexion-extension was always recorded except in the pre-surgery session due to subject’s impairment of movement. For the same reason, half-squat and walking at 6 km/h were not recorded in the first 2 sessions after surgery.

IV. FEATURES EXTRACTION

The metrics considered for the patient’s assessment are divided into seven categories described below. More details on the computation of the features are reported in [6]-[8].

A. Gait Metrics

Well-known gait measures are calculated from the data recorded by the inertial sensors attached on the shanks, including: Gait Cycle Time (GCT), Stance Phase (StP), Swing Phase (SwP), Stride Length (SL), Stride Speed (SS), Stride Clearance (SC). This information is obtained for both legs only for walking. This category includes 6 features.

B. Range of Motion (RoM) Metric

Knee Range of Motion (RoM), defined as the peak-to-peak amplitude of the knee joint angle over the x-, y-, and z-axis during a single exercise repetition, is obtained for both limbs and for all the exercises taken into account. This category includes 3 features.
C. Kinematic Metrics

Kinematic metrics, which have been occasionally adopted for gait analysis, can provide useful information on the movement analysis. Those metrics include: Range of Angular Velocity (RAV), Vertical Acceleration (VA), Vertical Velocity (VV), Fluency (along the three axes), Kinetic Value (KV). All those features are calculated for each of the 4 sensors used for data collection and for all the exercises. This category includes 7 features.

D. Stability Metric

Stability is defined as the dynamic time warping of the x-, y-, z-axis of the acceleration and angular rate signals measured at two consecutive repetitions/strides, then averaged based on all the repetitions present in a test session. Those features are calculated for each of the 4 sensors used for data collection and for all the exercises. This category includes 3 features.

E. Jerk-based Metrics

Jerk is the rate of change of the acceleration in a repetition. Several jerk-based metrics have been investigated in literature, including: Integrated Squared Jerk (ISJ), Mean Squared Jerk (MSJ), Cumulative Square Jerk (CSJ), Root Mean Square Jerk (RMSJ), Mean Square Jerk normalized by peak speed (N_MSJ), Integrated Absolute Jerk (IAJ), Mean Absolute Jerk normalized by peak speed (N_MAJ), Dimensionless Square Jerk (DSJ). Those features are calculated on the three axis, for each of the 4 sensors used for data collection, and for all the exercises. This category includes 24 features.

F. Statistical Metrics

This category takes into account various well-known statistical features extrapolated from the time-domain. Those variables are applied on every segmented walking stride/exercise repetition for both legs performed during the sessions. The selected features are described below:
- Mean, standard deviation, skewness, kurtosis, root mean square, calculated over the acceleration and angular velocity magnitudes,
- Mean, standard deviation, skewness, kurtosis, root mean square, minimum, maximum, Coefficient of Variation (CV), and Peak-to-Peak (p-p) amplitude over the x-, y-, and z-axis of the acceleration and angular rate signals.

All those features are calculated for each of the 4 sensors used for data collection. This category includes 64 features.

G. Spectral/Entropy/Information-Theoretic Metrics

This category takes into account various well-known spectral, entropy, and information-theoretic feature. Spectral metrics are obtained using the Fast Fourier Transform (FFT). Those variables are applied on the raw 3-axis of the accelerometer/angular rates data collected for both legs on each session. All these features are calculated for each of the 4 sensors used for data collection. This category includes 74 features. The features are: Dominant Frequency (DF) and its Width (FWHM), Spectral Centroid (SpC), Power in 1.5-3 Hz (LFP), Power in 5-8 Hz (MFP), 25-50-75% Quartile Frequency (QF), Spectral Edge Frequency (SEF) at 95% (calculated on the magnitude signal), Harmonic Ratio (HR), Ratio High-Low bands (RHL), Frequency-Domain Entropy (FER), Lempel-Ziv Complexity (LZC).

While LZC is calculated on the single repetitions/strides, the other features are not extrapolated for each segmented walking strides/repitions but are obtained for a sliding window covering the 50% of the whole signal, with 10% overlapping. The data analysis is implemented off-line over the data collected using a commercial software package (MATLAB R2015a, The MathWorks Inc., Natick, MA, 2015). Each repetition/stride was visually segmented.

V. Score Modeling

Preliminary analysis described in [6]-[8] have highlighted that several parameters are seen to be potentially relevant to provide indications on patient’s performance during rehabilitation. However, to support clinicians during their clinical practice, it is essential to obtain a single indicator scoring regarding patient’s performance, so as to avoid analyzing all the parameters separately.

The Mahalanobis distance is typically adopted to describe how much a patient’s performance deviates from the control group. However, this distance does not allow the comparison between the data distribution related to the affected and unaffected side, which is essential for rehabilitation. Moreover, this distance assumes that control and patient group have comparable standard deviations, which cannot be assumed in patients following orthopedic injuries. A more reliable extension of this metric is the Bhattacharyya Distance $D_B$, which measures the similarity of two discrete distributions. $D_B$ obtained as follows:

- Given a specific session and a specific exercise, every feature listed in Section IV is extrapolated from the raw inertial data of all 4 sensors.
- Potential outliers in the feature distribution are then detected and replaced via Winsorization.
- Once the outliers are replaced, the feature vectors for the left and right leg are considered as input of the $D_B$ calculation with their associated averages and standard deviations. This process is repeated for every exercise, session and feature. As a result, after $M$ sessions, for every feature the distance vector $C$ is obtained.

An accurate assessment of a patient’s performance requires the selection of the informative features from every category, excluding those uninformative or redundant. Some features can be informative for some exercises and being redundant for others; thus, it is important to define an automatic method for selecting those features.

A common technique for feature selection is the Least
Absolute Shrinkage and Selection Operator (LASSO) [10]. This regression tool requires to define an output in order to adjust the weights of a linear model which defines the features to be selected. As shown in [10], this output was defined as linearly increasing from the first to the last test session, with this period ranging from 4 to 12 days in the experiments carried out by the authors. However, even though this assumption can be accepted for the short period of time immediately following surgery, it may be unrealistic when analyzing rehabilitation outcomes for a longer period post-surgery and also pre-surgery.

An alternative feature selection approach recently studied is the Clustering Coefficients of Variation (CCV) [11], which is a light-weight and efficient method based on feature variability. The features are clustered according to their CV, and then the optimal cluster of features for the model is chosen. Features showing the most variation between limbs and between different sessions over the course of the rehabilitation represent informative features to be chosen. Therefore, for each M-dimensional distance vector $C$ calculated, the associated CV is obtained. If the CV is lower than one, than the associated distance vector is discarded. Following this initial selection, the remaining distance vectors are normalized through the standard score approach and then considered as points in an M-dimensional space where they are clustered via a weighted K-means clustering ($K = 2$ [11]). The normalization step before the clustering is important in order to guarantee that different scaling between the features could not impact the clustering.

As a result of the weighted K-means, the features are divided in two clusters. For both clusters, the normalized distance vector $C$ are averaged among all the features, resulting in two M-dimensional scoring vectors. These scoring vectors are then rescaled so as to be within the range [0-1]. A high score indicates a large distance in performance between limbs, and vice versa a low score represents a small difference. One of the two clusters (and associated scoring vectors) provides the optimal feature subset, and this selection is realized by using Hyper-Pipes [11].

To the best of the authors’ knowledge, it is the first time that a combination of Bhattacharyya distance and CV-based weighted K-means clustering is investigated for monitoring patients’ progress in a rehabilitation context. A summary of the scoring algorithm used is illustrated in Figure 2.

VI. RESULTS

In each session, each exercise was divided in two separate tests (both logged for 60 sec), and in each of the two tests a series of repetitions have been carried out by the subject. The overall number of repetitions recorded for all the sessions was: 184 hamstring curls (92 left / 92 right), 134 flexion/extension (67 left / 67 right), 66 half squats, 478 strides for both legs when walking at 3 km/h, and similarly 544 strides when walking at 4 km/h, and 512 strides when walking at 6 km/h.

WIMUs have been attached to the anterior tibia, 10 cm below the tibial tuberosity, and to the lateral thigh, 15 cm above the tibial tuberosity using surgical adhesive tape. For each test, the features, separated for every category as described in Section IV, were extrapolated and compared among the different sessions after applying the scoring method defined in Section V.

Finally, in order to have the same reference system for both WIMUs worn on the same leg, the method proposed by Seel et al. [12] has been adopted to virtually rotate around an axis the raw inertial data recorded on the shank. As a result, for all the WIMUs involved, the x-axis represents the mediolateral axis, the y-axis is the anteroposterior one, while the z-axis is the vertical axis. Thus, the plane y-z represents the sagittal plane.

Results for the metrics associated with gait, RoM and kinematics are described in the following subsections. Results for additional metrics are still under analysis and will be described in future works.

A. Gait Metrics

Considering the gait results at 3 km/h, the metrics which are clustered and show patients’ progress during rehabilitation are SwP and StP. The resulting score shows a clear increase in association to the second session (due to the early stage of the recovery process post-surgery), with a consecutive decreasing trend converging toward the score value reported in the pre-surgery session. A similar behavior is also shown in the gait exercise walking at 4 km/h, also including GCT in the selected features.
The score highlights a peak in the post-surgery early stage with an evident convergence towards zero in the following sessions. Finally, SC is the only feature selected for the gait at 6 km/h. The score calculated for the pre/post-surgery is comparable (although this exercise was not recorded in the first two sessions after surgery because of patient’s impairment) even though an unexpectedly large value is obtained in the long-term session as shown in Figure 3.

B. RoM Metrics

In the hamstring curl exercise, the mediolateral RoM (e.g., over the x-axis) shows a clear trend with a score steadily decreasing following surgery. The RoM over the z-axis is, instead, selected for the flexion/extension exercise. Even though there is a general tendency of the score to decrease starting from the second sessions, two large scores are obtained for the 6th and 8th session, indicating a non-monotonic improvement. Similar considerations can be also drawn for the squat exercise (z-axis RoM), that is a general reduction of the calculated score with an exception reported in the 7th session.

Walking tests at 3 and 4 km/h have both selected the x-axis RoM. In the former case, the score has a large difference from the first two sessions, while generally decreasing to low score values in the following sessions, with results lower than the pre-surgery period. In the latter, instead, the RAV score is also showing a decreasing trend with an exception in the 7th session. Finally, anteroposterior axis RoM is the feature selected for the gait at 6 km/h. The score calculated for the pre-surgery session has a much larger value in comparison with the remaining session following surgery, even though the trend in this period is clearly not monotonic. The discussed results are shown in Figure 4.

C. Kinematics Metrics

In the hamstring curl exercise, the score obtained from the features selected considering the sensor attached on the shank (e.g., VV and Fluency over the x- and z-axis) shows a clear trend decreasing following surgery, even though the tendency in the long-term is not monotonic. Identical considerations can be drawn for the thigh sensor (with chosen features being RAV, VV, and y-axis Fluency) despite an even less flat trend in the late session; indeed, it is evident a large score value on the 6th session.

The flexion-extension exercise is described by similar conclusions, but the metrics taken into account are VV and z-axis Fluency from the shank, and only z-axis Fluency from the thigh.

Regarding the squat exercise, again the score obtained by considering VA and z-axis Fluency from the shank generally shows a decreasing trend throughout the sessions. However, when observing the score extrapolated from the thigh sensor data, no particular correlation is evident due to several large values, indicating that the thigh sensor placement is not beneficial when analyzing squats.

For walking at 3 km/h RAV and y-axis Fluency, and KV are selected for the shank and thigh, respectively. The score trend is comparable for both sensor locations, with a score presenting a large difference between pre-surgery and immediate post-surgery, and with a decreasing score reaching its minimum at the 6th session. However, large scores are shown again for the last two data collections.

Walking at 4 km/h is not significantly different, even though in this case the selected features are RAV, x-axis Fluency, and KV for the shank. The only dissimilarities compared to the 3 km/h exercise are evident in the larger score on the shank in the pre-surgery session and the flatness shown in the thigh-related score.
Finally, RAV and z-axis Fluency are the feature selected from the shank (thigh) for the walking exercise at 6 km/h. The score has a general decreasing trend when considering the shank data, even though two large values are shown on the 6th and 7th session. The data from the thigh, instead, show a score comparable between the pre-surgery and the long-term period, without any particular trend. Thus, the thigh sensor placement may be not beneficial when analyzing faster speeds. The discussed results are shown in Figures 5-6.

VI. RESULTS SUMMARY AND CONCLUSIONS

The focus of this work-in-progress has been on the analysis of certain metrics associated with gait analysis in an effort to develop a score-based system to aid clinicians in the diagnosis and evaluation of gait in a rehabilitative context. To summarize, this work analyzed the body-worn inertial data collected from a patient over the course of rehabilitation defining a score metric from a number of features for better understanding and monitoring patient’s progress and limbs comparison in several tests. It has been also shown that several metrics, gait, joint angle-related, and kinematics variables, obtained from acceleration and angular rate of the shank and thigh have proved their sensitivity for a number of exercises.

This work presented a wearable inertial system for an objective assessment of lower-limbs in patients over the course of. The hardware platform adopted for the system realization and the data analytics involving inertial data collected from thighs and shanks have been described. The present study proved that a novel scoring method involving Bhattacharyya distance metrics and Clustering Coefficient of Variation for feature selection, based on a number of well-known metrics extrapolated from inertial data collected on the lower-limbs, can be used for defining quantitatively patients’ progress when involved in a rehabilitation program. Accurate results have been shown in a number of exercises. The proposed method is able to indicate which features are more informative regarding patients’ performance and group them in a single indicator which can be easily taken into account by clinicians during their analysis. This score indicator represents an important step towards the development of an objective model for patients’ assessment during rehabilitation.

As only a single subject has been analyzed for the present study, an enhanced number of athletes, with homogeneous characteristics, will also be tested to have a more robust base and further validate the drawn conclusions. Results associated with additional metrics, such as jerk, statistical, spectral features, are currently under investigation and will help assist in the development of such a score-based system as is envisaged by the authors. Additional clinical trials are currently being planned to further validate the developed model in statistical terms.

ACKNOWLEDGMENT

This publication has emanated from research supported by a research grant from Enterprise Ireland (EI) under grant number CF-2017-0685-P. Aspects of this work have emanated in part from research supported in part by Science Foundation Ireland (SFI) and is co-funded under the European Regional Development Fund, Grant Number 13/RC/2077-CONNECT and SFI-16/RC/3918-CONFIRM.

REFERENCES