Gait parameter estimation from a miniaturized ear-worn sensor using singular spectrum analysis and longest common subsequence

JARCHI, D., WONG, C., KWASNICKI, R.M., HELLER, Ben <http://orcid.org/0000-0003-0805-8170>, TEW, G.A. and YANG, G.Z.

Gait Parameter Estimation from a Miniaturised Ear-Worn Sensor using Singular Spectrum Analysis and Longest Common Subsequence

Delaram Jarchi, Charence Wong, Richard M Kwasnicki, Ben Heller, Garry Tew and Guang-Zhong Yang, *Fellow, IEEE,*

***Abstract*—This paper presents a new approach to gait analysis and parameter estimation from a single miniaturised ear-worn sensor embedded with a triaxial accelerometer. Singular spectrum analysis (SSA) combined with the longest common subsequence (LCSS) algorithm has been used as a basis for gait parameter estimation. It incorporates the information from all axes of the accelerometer to estimate parameters including swing, stance and stride times. Rather than only detecting local features of the raw signals, the periodicity of the signals is also taken into account. The proposed method for capturing major gait events such as the heel contact and toe off is validated with a high-speed camera, as well as a force-plate instrumented treadmill. The results have been validated with data from ten older adults demonstrating the accuracy of the analysis framework. Considering the average estimates of gait parameters related to the left and right foot, the absolute mean errors are 35.5 milliseconds for the swing time, 36.9 milliseconds for the stance time and 17.9 milliseconds for the stride time. In addition, the method has been applied to a set of healthy subjects and patients with lower limb trauma to evaluate the clinical value of the sensor and its analysis framework in a free living environment.**

***Index Terms*—Singular Spectrum Analysis (SSA), Longest Common Subsequence (LCSS), gait, e-AR (ear-worn activity recognition) sensor.**

I. Introduction

SYSTEMATIC evaluation of bipedal locomotion, or 'gait analysis', can provide useful information regarding human biomechanics, behaviour and pathology. Traditionally, this task is conducted subjectively with a set of predefined observation-based protocols. This has recently been replaced by more objective techniques employing optical tracking, multiple cameras and force plates in dedicated gait laboratories for detailed gait measurements. Recent advances in wearable sensing have further improved the practical use of the technique, allowing small wireless sensors to be integrated into wearable, prosthetic, and assistive devices [1][2][3].

The major advantage of such sensing technologies is in its ability for long-term continuous monitoring of the patient in a free-living environment, rather than specialised laboratory settings. In addition, specific context-aware gait monitoring systems can be developed to help understand the progression of disease, assess the efficacy of treatment and the rehabilitation process, and predict the onset of adverse events such as unstable gait patterns that may lead to high probabilities of falls in elderly patient groups.

D. Jarchi, is with the Hamlyn Center, Imperial College London, London, UK (e-mail: d.jarchi@imperial.ac.uk). This work is supported by the ESPRIT project, UK. C. Wong and R. M. Kwasnicki are with the Hamlyn Center, Imperial College London, London, UK (e-mail: charence.wong05@imperial.ac.uk, richard.kwasnicki07@imperial.ac.uk), and B. Heller and G.A. Tew are with Sheffield Hallam University, Sheffield, UK (email: G.Tew@shu.ac.uk,B.Heller@shu.ac.uk). G-Z. Yang is with the Hamlyn Centre, Imperial College, London, UK (g.z.yang@imperial.ac.uk).

The use of accelerometers for gait analysis has been established since more than a decade ago. For example, Aminian et al. used two accelerometers, one on the heel and the other on the trunk, to detect walking patterns [4]. Recently, a range of wireless devices has been developed due to the popularity of miniaturised MEMS-based multi-axis accelerometers integrated with low-power wireless embedded platforms. This includes a study on gait analysis that multiple sensors including accelerometers, gyroscopes and pressure sensors were used to detect heel contact and toe off with good accuracies [5].

For data analysis, acceleration signals for walking can be classified, for example, using a wavelet-based fractal analysis method [6]. This allows clustering of walking patterns from different patient groups such as those with Parkinson's disease.

One of the major research topics in the use of these sensors is how to balance the complexity (e.g. the number of sensors required and their practical embodiment) against the reliability and underlying information content of the platform. Naturally, the use of multiple sensors provides more information that is directly or indirectly related to the gait patterns. However, this complicates system design in terms of cross-node communication, synchronisation, and modelling. It also affects user compliance. Furthermore, consistent placement of multiple sensors is difficult, thus affecting the reliability and accuracy of the system. Such an approach, therefore, still tends to be limited to laboratory experiments.

Integrating all sensing capabilities into a single wireless sensor node has clear advantages, particularly for patient studies. Existing research has shown that detection of certain spatio-temporal gait parameters is possible with single accelerometers. Lower trunk accelerations can be used to predict the subsequent stride's cycles and left/right steps, allowing estimation of step length and walking speed [7][8]. Changes in gait cycle variability have been explored in musculoskeletal disorders [9]. For detailed gait analysis, other parameters such as swing and stance durations are also required. Thus far, the detection of toe off with a single accelerometer is poorly studied, and most studies are limited to multiple sensor configurations [10][11][12].

With the constraint of using a single sensor, prior research has also been directed to the issue of optimal sensor placement with more interest in patient cohort comparisons. To this end, it is necessary to address the practical requirements of: 1) ease of sensor placement; 2) consistency and repeatability; 3) the underlying information content of signals. Such a problem can be treated as a feature selection problem with a multi-objective function by incorporating the above considerations and other system related constraints.

Previous research has shown that by placing the sensor behind the ear, most of the above constraints can be satisfied [13]. It also takes advantage of the intrinsic capabilities of the skeletal bone in transmitting both high and low-frequency waves to the cranium, which can be picked up by the sensor. This, in essence, reproduces the mechanism of how humans control gait and balance. Based on this concept, we have developed an ear worn activity recognition (e-AR) sensor by using the body sensor network (BSN) platform [14]. It has been shown that gait-related force estimations, including weight acceptance and impulse can be derived. This has been validated with a force-plate instrumented treadmill for both normal participants and

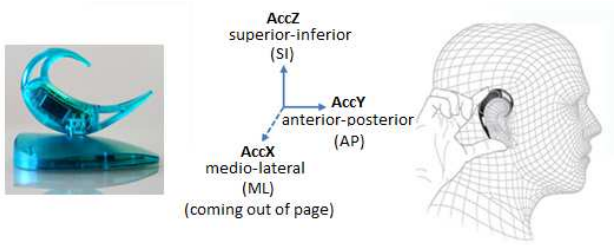

Fig. 1. The light-weight version of e-AR sensor (left), axes of orientations defined for this paper (middle) and how the sensor is worn on the ear (right).

patients after knee surgery [15].

Thus far, the algorithms used for gait pattern analysis are limited to extraction of peaks in the raw acceleration signals mostly to find only heel contacts. Therefore it is necessary to extract both heel contacts and toe off events from acceleration signals with accurate validation using e.g. synchronised video or force-plate related data. Detection of heel contacts and toe off events has been used in many different clinical applications. These applications include improving the gait of stroke patients for Functional Electric Stimulation (FES), evaluating pathological gait impairments, investigating prescription footwear effects on quality of the gait for rehabilitation, investigating influence of increasing age on gait parameters and observing recovery of orthopedic patients after lower limb surgery.

The purpose of this paper is to provide a robust technique for detailed gait analysis by addressing the current drawbacks in single accelerometer-based approaches (e.g. limiting the study to detection of heel contacts only). The proposed method is model-free and based on Singular Spectrum Analysis (SSA) for time-series analysis. It incorporates a time-series matching approach called Longest Common Subsequence (LCSS) to enhance the noise-resilience of the proposed algorithm. Detailed validation of the proposed method with high-speed cameras and a force-plate instrumented treadmill is provided. The results have shown the importance of estimated gait parameters in clinical applications such as monitoring recovery after surgery.

The remainder of the paper is structured as follows. In Section II, details on experimental setup and sensor placement are first described. This is followed by an introduction to the SSA and LCSS algorithms to be used. In Section III, detailed experimental results and validation are provided. Finally, Section IV concludes the paper by summarising the main contribution of the paper, its limitations and potential future improvements.

II. Methods

A. Sensor Hardware and Placement

In this paper, an e-AR sensor is used for recording the acceleration signals arising from the gait patterns. The sensor is based on the BSN platform which contains an 8051 processor that has a 2.4 GHz transceiver (Nordic nRF24E1), a 3D accelerometer (Analog Devices ADXL330), a 2MB EEPROM (Atmel AT45DB161), and a 55mAhr Li-Polymer battery [16]. The light-weight version of the e-AR, having a mass of 7.4 g, that allows the recordings of mobility information and can be used in healthcare and sports applications is shown in Fig. 1. In this figure, the axes orientation of the e-AR sensor and how it is worn by the user are also shown.

B. Singular Spectrum Analysis

SSA is a model-free technique that can be applied to time-series data to decompose it into a number of orthogonal components. These components include slowly varying trend, oscillatory and unstructured noise [17]. SSA has been successfully used in many applications of times series analysis including denoising and prediction. For example, it has been applied to bio-signals such as single channel respiratory signals and the source signals are effectively separated

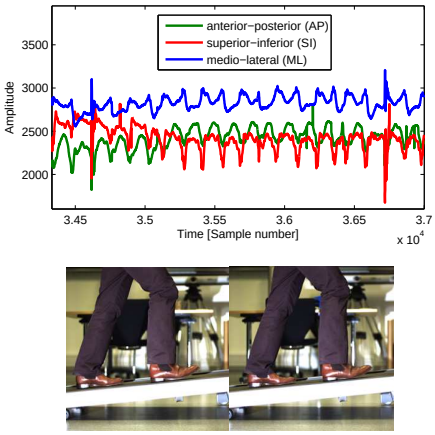

Fig. 2. The raw acceleration signals and the corresponding images for the large heel strikes captured by the high speed camera.

[18][19]. Generally, the SSA algorithm consists of two main stages: decomposition and reconstruction.

Decomposition: This stage consists of two steps called embedding and singular value decomposition (SVD).

Embedding: In this step, by using the delayed version of the input time-series, a multidimensional matrix named trajectory matrix is created. Therefore, the time-series $\mathbf{s}$ having the length n is converted into an $l \times k$ matrix as the trajectory matrix:

$$\begin{aligned}\mathbf{X} &= [x_{ij}] = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k] \\ &= \begin{bmatrix} s_0 & s_1 & s_2 & \dots & s_{k-1} \\ s_1 & s_2 & s_3 & \dots & s_k \\ s_2 & s_3 & s_4 & \dots & s_{k+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{l-1} & s_l & s_{l+1} & \dots & s_{n-1} \end{bmatrix}\end{aligned} \quad (1)$$

where $k = n - l + 1$ and l is the window length (embedding dimension) ($1 \leq l \leq n$). Vectors $\mathbf{x}_i = [s_{i-1}, s_i, ..., s_{i+l-2}]^T$ form the columns of the trajectory matrix and are called *l*-lagged vectors. It is evident from Equation (1) that the elements of all the diagonals $i + j = constant$ of the trajectory matrix are the same. Therefore, the trajectory matrix is a Henkel matrix.

Singular value decomposition (SVD): In this step SVD is applied to the trajectory matrix to represent it as a sum of rank-one bi-orthogonal elementary matrices. Let $\mathbf{S} = \mathbf{X}\mathbf{X}^T$ and assume $\lambda_1, \lambda_2, ..., \lambda_l$ are eigenvalues of $\mathbf{S}$ in decreasing order of magnitudes ($\lambda_1 \geq \lambda_2 \geq , ..., \geq \lambda_l \geq 0$) and the corresponding eigenvectors are $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_l$. If $\mathbf{v}_i = \mathbf{X}^T \mathbf{u}_i / \sqrt{\lambda_i}$, then it is possible to write the trajectory matrix as:

$$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2 + ... + \mathbf{X}_d \quad (2)$$

where $d = \underset{i}{argmax}\ \lambda_i > 0$ and $\mathbf{X}_i = \sqrt{\lambda_i}\mathbf{u}_i\mathbf{v}_i^T$. The vector $\mathbf{u}_i$ is the *i*th left eigenvector, the vector $\mathbf{v}_i$ is the *i*th right eigenvector, and the collection $(\sqrt{\lambda_i}, \mathbf{u}_i, \mathbf{v}_i)$ is called the *i*th eigentriple of the SVD in Equation (2). The projection of the time-series onto the direction of each eigenvector gives the corresponding principle component.

Reconstruction: In this step, the eigentriples are grouped into m disjoint subsets. By setting the indices for subsets $(I_1, I_2, ..., I_m)$, the elementary matrices of each group are summed together as:

$$\mathbf{X}_{I_j} = \mathbf{X}_{i_{j1}} + ... + \mathbf{X}_{i_{jp}} \quad (3)$$

where $I_j = i_{j1}, ..., i_{jp}$. The original trajectory matrix can be written as the sum of all the resulted matrices $\mathbf{X}_{I_j}$:

$$\mathbf{X} = \mathbf{X}_{I_1} + ... + \mathbf{X}_{I_m} \quad (4)$$

Grouping of the eigentriples depends on the application and therefore there are no general rules in practice for grouping. The final step

of the SSA algorithm is the diagonal averaging in which the final elementary matrix, which is the sum of all elementary matrices in the group is transformed into a time-series of length n. Each element of the resulted time-series is computed using the average of the matrix elements over the diagonal $i+j = constant$. Considering a general $l \times k$ matrix, the qth element of the time-series is given by averaging over the diagonal $i + j = q + 2$. The reconstruction step of SSA algorithm is potentially useful for time-series denoising. The SSA algorithm for time-series reconstruction based on grouping one set of elementary matrices given the corresponding indices is summarized in Algorithm 1. The input to this algorithm is the original time-series ($\mathbf{s}$), the set of indices (I) to group the elementary matrices and the embedding dimension (l). Then the output is the reconstructed time-series based on grouping the corresponding elementary matrices.

C. Longest Common Subsequence

Let $\mathbf{a}$ and $\mathbf{b}$ be finite discrete time-series. $\mathbf{a}_1^p$ denotes the time-series $\mathbf{a}$ with discrete time index varying between 1 and p. Similarly $\mathbf{b}_1^q$ and denote the time-series $\mathbf{b}$ with discrete time index varying between 1 and q. In addition, let a_i,b_i be the i^{th} sample of time-series $\mathbf{a}$ and $\mathbf{b}$ respectively.

The Longest Common Subsequence (LCSS) algorithm has been initially used for string matching applications [20]. It has been subsequently extended to measure the similarity of two time-series having different length [21], [22]. The algorithm uses dynamic programming and matching regions in time and space. The idea is to avoid matching the regions that are distant or degenerate. The recursive formulation of the LCSS is defined in [22] as follows:

$$\mathrm{LCSS}_{\delta,\epsilon}(\mathbf{a}_1^p, \mathbf{b}_1^q) = \begin{cases} 0 & \text{if} \quad p < 1 \quad or \quad q < 1, \\ 1 + \mathrm{LCSS}_{\delta,\epsilon}(\mathbf{a}_1^{p-1}, \mathbf{b}_1^{q-1}) & \text{if} \begin{cases} d_{LP}(a_p, b_q) < \epsilon \quad and \\ \mid p - q \mid < \delta, \end{cases} \\ Max \begin{cases} \mathrm{LCSS}_{\delta,\epsilon}(\mathbf{a}_1^{p-1}, \mathbf{b}_1^{q}) \\ \mathrm{LCSS}_{\delta,\epsilon}(\mathbf{a}_1^{p}, \mathbf{b}_1^{q-1}) \end{cases} & otherwise \end{cases} \tag{5}$$

where p and q are the length of time-series $\mathbf{a}$ and $\mathbf{b}$ respectively, and $d_{LP}(a_p, b_q)$ is any L_P-norm of the $(a_p - b_q)$. The constant δ provides the flexible control of the matching region in time while the constant ϵ is a threshold for matching in space. The value given as the output of the LCSS depends on the length of its input time-series. Therefore the similarity $S_{\delta,\epsilon}$ of the two series is measured by normalizing the output of the LCSS as:

$$S_{\delta,\epsilon}(\mathbf{a}, \mathbf{b}) = \frac{\mathrm{LCSS}_{\delta,\epsilon}(\mathbf{a}_1^p, \mathbf{b}_1^q)}{min(p, q)} \tag{6}$$

D. Gait Parameter Estimation

For estimation of gait parameters, the gait signals need to be segmented from the acceleration data recorded using the e-AR sensor. In many traditional experiments, the acceleration signals are often labelled manually. However in applications that require continuous monitoring, the gait signals should be automatically segmented. In [23] a method has been proposed to discriminate walking activity from non-walking. This method based on SSA can be further extended for automatic gait segmentation using acceleration signals.

In order to estimate the gait parameters, the essential gait events need to be located on the signals. These events include right heel contact (RHC), left heel contact (LHC), right toe off (RTO) and left toe off (LTO).

Algorithm 1 SSA for time-series reconstruction

$\widetilde{\mathbf{s}}$ = SSA($\mathbf{s}$, $I = (i_1, ..., i_p)$,l)
- Create trajectory matrix $\mathbf{X}$, see Equation (1)
- Apply SVD onto the $\mathbf{X}$ to find $\mathbf{X}_j$ ($\sqrt{\lambda_i}, \mathbf{u}_i, \mathbf{v}_i^T$)
- Group elementary matrices $\mathbf{X}_I = \sum_{j=i_1}^{i_p} \mathbf{X}_j$
- Perform diagonal averaging on $\mathbf{X}_I$ to construct $\widetilde{\mathbf{s}}$

return $\widetilde{\mathbf{s}}$

To locate the gait events, an experiment is first performed in which the e-AR signals are synchronized with a Photron FASTCAM SA3 high-speed camera operating at 250 Hz. Then the gait events extracted from the images are used to locate the events on the signals. To make the synchronisation possible, two big heel strikes are performed in the experiment. These heel strikes produce large peaks in the e-AR signals. Finding the corresponding image of each large peak is helpful to resample the accelerometry data linearly to the high-speed camera rate. Three experiments are performed using a subject walking on the treadmill having the following situations:

- Constant speed, zero incline
- Increased speed, zero incline
- Increased incline, constant speed

As an example, the results for increasing incline and constant speed are shown in Figs. 2 and 3. In Fig. 2, the raw acceleration signals are shown, which illustrates two clear peaks. The corresponding images for the large peaks given by the high-speed camera are shown in Fig. 2. After resampling, two gait cycles are selected and the gait events are located on the acceleration signals. In Fig. 3, the two consecutive gait cycles and the images given by the high-speed camera for the main gait events (RHC, LHC, RTO and LTO) are shown.

In summary, by using the output of all the experiments, as it will be explained in detail, for detection of heel contacts, the signals of the anterior-posterior (AP) and superior-inferior (SI) axes should be used in which there are local minimum peaks in both axes. To determine which heel contact corresponds to the left and which one to the right, the acceleration signal of the medio-lateral (ML) axis is used. The mean value of the acceleration amplitudes from one RHC to the next LHC is larger than the mean value of the acceleration amplitudes from the LHC to the RHC. For detection of toe off events, the best axis to use is the ML axis. The accelerations of ML axis are segmented using RHCs and the toe-off events are then detected. The RTO is the first local minimum after the LHC, and the LTO is a local maximum before the LHC (due to slightly different appearance of the LHCs on the ML axis, throughout the paper, this local maximum is referred to a local maximum before the main valley of the ML axis cycle segmented from RHC to the LHC). Since raw accelerations are noisy, there maybe many local maxima/minima points. Therefore, instead of directly locating the toe-off events on the raw accelerations, a number of gait cycles are grouped and by using different techniques (SVD, LCSS and peak detection), toe-off events are estimated. In fact these events are located on enhanced gait cycles which result in more accurate estimations than locating them on raw accelerations.

Therefore, using all the axes we will be able to locate the essential gait events (RHC, LHC, RTO and LTO) and then to estimate the gait parameters based on the extracted time stamps. In the following a methodology is explained to estimate all the gait parameters.

- **Trend Removal**
  The e-AR acceleration signals usually contain artefact that is created by head motion. This artefact is added to the accelerometer data as a trend that can be removed by the SSA algorithm. To remove the trend of the signal, SSA is applied into the e-AR signals. The first eigentriple that relates to the trend of the data is selected, its elementary matrix is formed and then it is used to reconstruct a time-series. This time-series is subtracted from the original signal. In the following the signal of the AP axis ($\mathbf{s}_2$) is used and its trend is removed by the SSA:

$$\begin{aligned} \mathbf{s}_2 &\longleftarrow \mathrm{AccY(AP)} \\ \widetilde{\mathbf{s}} &= \mathrm{SSA}(\mathbf{s}_2, I = (1), l), \ see\ Algorithm\ 1 \\ \widetilde{\mathbf{s}}_2 &= \mathbf{s}_2 - \widetilde{\mathbf{s}} \end{aligned} \tag{7}$$

  where $\widetilde{\mathbf{s}}$ is the trend of the signal of AP the axis, $\widetilde{\mathbf{s}}_2$ is the acceleration signal of AP axis after trend removal. Since the SSA algorithm is based on linear combination of elementary matrices related to eigentriples, the trend removal can be also performed in another way by grouping all eigentriples except

the first one as:

$$\widetilde{\mathbf{s}}_{\mathbf{2}} = \mathrm{SSA}(\mathbf{s}_2, I = (2, ..., l), l) \tag{8}$$

The above equation is used to remove the trend from all axes. When SSA is applied to the new signal (after removing trend from AP axis), the first two eigenvectors correspond to the most dominant oscillation of the signal. If there is no artefact by head motion then the reconstructed signal after removing the trend is the same signal shifted around zero to have a mean value of zero or very close to zero. In Fig. 4., the gait signal from the AP axis is shown before and after trend removal. In this figure, the generated eigenvalues are shown for each signal separately after applying the SSA algorithm. It is evident that the first eigenvalue before trend removal has a large value that after normalization makes other eigenvalues very close to zero. However after trend removal, the first two eigenvalues have close values related to the most dominant oscillation of the signal from the AP axis. In the next step these two first eigenvectors from the AP axis are reconstructed using SSA which helps to find the RHC and LHC.

- **Detection of RHC and LHC**
To detect the RHC and LHC, the signal of the AP axis is used to find an interval in time domain for detection of heel contacts. After removing the trend from the AP axis, a peak detection method is applied to the dominant oscillation of AP axis (obtained by reconstructing the signal using the first two largest eigenvalues) to find the local maximum and local minimum points. Most existing methods have applied the peak detection method to the raw signals to find the heel contacts [24][11][15][7]. There are also some methods based on autocorrelation analysis to estimate the gait cycle considering variations in the speed of walking [8].

  Here the objective of applying the peak detection method to the first dominant oscillation of the AP signal is to exploit the periodicity of the signal in detecting RHC/LHC, not only relying on the explicit peaks in the data. This is important in that not all of the heel contacts produce distinctive peaks in the AP/VT signals. Therefore we do not rely only on the amplitudes of the signal and apply a peak detection method to extract the RHC and LHC. Since walking generates repetitive patterns, the analysis of periodic features of the signal will reveal important information regarding periodicity of the underlying system. The periodicity of the accelerations is taken into account but also the amplitudes of accelerations in two axes (AP and SI) are used to detect heel contacts towards more accurate estimations.

  Once the local maxima and minima of the dominant oscillation of the AP signal have been found, if there is some deviation on dominant oscillation which forms two very close local maximum points, a local minimum in between may appear. As an example it can be seen from Fig. 5 that the amplitude of such local minimum can be close to the amplitude of each local maxima. Therefore it is an invalid point as the local minimum point. The same situation may happen in which invalid local maximum and several close local minima points are formed. A simple process is performed to remove invalid local maxima/minima. After finding all these extrema of the first dominant oscillation of the AP axis, a further validation process is performed. In order to estimate the LHC and RHC more accurately, a short interval is constructed using the corresponding local minima points of the dominant oscillation. Since after removing the trend from AP or VT axis the signal is centered around zero, it is beneficial to add the mean of the accelerations before trend removal to reconstruct the accelerations centered around their mean. Then the signals of the AP and VT axes, after trend removal and mean correction, are multiplied in each specified interval (based on the detected local minima of the dominant oscillation) and the point that gives the minimum value is considered as the index of heel contact.

  After finding the heel contacts, the signal of the ML axis is used to determine the left/right heel contacts (RHC/LHC). First three local minimum indices (given by using the dominant oscillation of AP axis) are selected. If the mean amplitude value of the signal of the ML axis from the first local minimum's index to the second one is bigger than the mean value from the second to third one, the first local minimum is RHC, the second one is LHC and the third one is RHC. In a similar manner if the mean amplitude value of the signal of the ML axis from the first local minimum's index to the second one is smaller than the mean value from the second to third one, the first local minimum is LHC, the second one is RHC and the third one is LHC. In Fig. 6, the signal of the ML axis, the dominant oscillation of the signal from AP axis and its local minima/maxima points for determination of RHC and LHC are shown. Therefore, the dominant oscillation of AP axis is used for determination of intervals which contain the time index of the heel contacts. Then based on these intervals, heel contacts are extracted using accelerations of both AP and SI axes. Finally, ML axis is used for determination of RHC and LHC.

- **Extraction of gait cycles**
In this step, the information given by the previous step is used to segment the signal from the ML axis. Segmentation is important for finding the toe-off points, as the synchronisation experiment showed that the ML axis is the best axis to estimate toe-off points. Therefore the signal of the ML axis is segmented between each two consecutive RHCs.

- **Applying SVD**
Since the acceleration signals can be noisy for practical applications, filtering of the raw signals is required.
The proposed technique in this paper is to apply SVD on a number of gait cycles (segmented from one RHC to the next RHC, with possible length extension, using the ML axis) and extract the main gait cycle:

$$\mathbf{X}_c = \begin{bmatrix} \widetilde{\mathbf{s}}_1(RHC(i) : RHC(i) + q - 1) \\ \vdots \\ \widetilde{\mathbf{s}}_1(RHC(i + N - 1) : RHC(i + N - 1) + q - 1) \end{bmatrix} = \begin{bmatrix} \mathbf{c}_0 \\ \vdots \\ \mathbf{c}_{N-1} \end{bmatrix} \tag{9}$$

$$[\mathbf{U}, \mathbf{\Sigma}, \mathbf{V}] = \mathrm{SVD}(\mathbf{X}_c^T) \tag{10}$$

$$\mathbf{g}_c = \mathbf{v}_1 / \mathrm{norm}(\mathbf{v}_1) \tag{11}$$

where $q = max\{\ RHC(i + j + 1) - RHC(i + j)\} + 1$, $\widetilde{\mathbf{s}}_1$ is the acceleration signal of ML axis (after trend removal and mean correction), $c_j (j = 0, ..., N - 1)$ denotes each individual gait cycle, $\mathbf{X}_c$ is an $(N \times q)$ matrix of N gait cycles, SVD performs singular value decomposition, $RHC(i)$ denotes the time-point index of the i^{th} RHC, $\mathbf{v}_1$ is the first vector (having the largest eigenvalue) of the $\mathbf{V}$ matrix which contains the right singular vectors and $\mathbf{g}_c$ is considered to be the main gait cycle. The main gait cycle is used as a template in the next step to estimate the toe-off events and also re-align the gait cycles to obtain enhanced (filtered) gait cycles.

- **Applying LCSS**
For every N groups of gait cycles, the main gait cycle $\mathbf{g}_c$ which is obtained by applying SVD on the N gait cycles, is compared to each individual cycle and their similarity value is estimated by the LCSS algorithm. If the similarity given by the LCSS is less than a specified threshold, then the corresponding gait cycle is removed. All the other gait cycles are then resampled to produce the enhanced gait cycles. The re-sampling technique used in this study uses the point correspondence of the main gait cycle and the raw cycle. It keeps only the samples corresponding to the main gait cycle.

- **Detection of RTO and LTO**
As the synchronisation results suggest, the RTO event is the first local minimum peak after the LHC and the LTO is a local maximum before the LHC. However detection of RTO and LTO from raw gait cycles are difficult since there maybe several local maxima or minima points. One strategy that practically found to be very effective, is to estimate the RTO and LTO from the main gait cycle $\mathbf{g}_c$ (obtained for every N groups of gait cycles) by considering the expected shape of the gait cycle and applying LCSS algorithm as detailed in below.
Here for detection of LTO on the main gait cycle $\mathbf{g}_c$, it is estimated based on a peak before the maximum of LHCs (in time domain based on the time stamps) for N grouped cycles. RTO is detected on g_c as the first peak after the minimum of LHCs in time domain for the N grouped cycles.
Detection of LTO has shown to be more difficult than RTO due to different appearance of LHC on the ML axis which creates extra local maxima/minima before LHC that are not related to LTO (see Figs. (3 and 7)). Therefore, for a more robust detection of LTO, a template based on expected shape of cycles from ML axis is constructed and the point on the main gait cycle g_c which is related to the first peak of the template, using the point correspondence output of LCSS, is considered as the LTO. Ten extracted main gait cycles (red color) and the constructed template (blue color) with the results from LCSS algorithm are shown in Fig. 7. The last point of the each main gait cycle is based on the maximum of LHCs for all grouped cycles. From Figs. (3 and 7), the importance of using a template for a more robust detection of LTO can be seen.
Finally after detecting representative LTO and RTO on the main gait cycle, by applying the LCSS method, the corresponding point of the peaks for LTO and RTO are located on each raw cycle. Therefore, RTO and LTO are not directly located on raw gait cycles. They are first located on the main gait cycle g_c, then on the raw cycles based on the outputs of LCSS. LTO is detected on g_c using a template and applying LCSS ($LCSS(r, g_c(1:g_i)), see\ Algorithm\ 2$) and RTO is detected on g_c using a peak detection technique. Then LTO and RTO are located on the enhanced cycle which is obtained by resampling the raw gait cycle based on the main gait cycle and the point correspondence output of LCSS ($LCSS(g_c, c_j), see\ Algorithm\ 2$).

- **Estimating gait parameters**
Once all the gait events (RHC, LHC, RTO and LTO) have been detected from the accelerometer signals, it is possible to estimate the gait parameters using the corresponding time stamps. These parameters include:

  – Left swing time: $t_s(LHC(i)) - t_s(LTO(i))$
  – Right swing time:$t_s(RHC(i+1)) - t_s(RTO(i))$
  – Left stance time:$t_s(LTO(i+1)) - t_s(LHC(i))$
  – Right stance time:$t_s(RTO(i)) - t_s(RHC(i))$
  – Left stride time: $t_s(LHC(i+1)) - t_s(LHC(i))$
  – Right stride time:$t_s(RHC(i+1)) - t_s(RHC(i))$

  where $t_s(.)$ denotes the corresponding time stamp given the time-point index of its input time-series, i denotes the index of the event, then $i+1$ is the index of the next event. Other important gait parameters can be obtained using the estimated heel contacts and toe-off events. In this paper we define the step asymmetry as:

  – step asymmetry:

$$\frac{t_s(LHC(i)) - t_s(RHC(i))}{t_s(RHC(i+1)) - t_s(LHC(i))}$$

The Pseudo-code for the estimation of gait parameters is shown in Algorithm 2.

Algorithm 2 Gait parameter estimation

-Segment the gait signals from the acceleration data
-$\mathbf{s}_1 \longleftarrow AccX(\mathrm{ML})$, $\mathbf{s}_2 \longleftarrow AccY(\mathrm{AP})$, $\mathbf{s}_3 \longleftarrow AccZ(\mathrm{SI})$
-Detect RHC and LHC
$\mathbf{y}$ = SSA($\mathbf{s}_2, I = (2,3), l$) dominant oscillation
$[i_{min}\ i_{max}]$ = Detect-local-min-max($\mathbf{y}$) returns indices
Apply constraint $y(i_{min}) < 0$ and $y(i_{max}) > 0$,
remove invalid local peaks
$\widetilde{\mathbf{s}}_1$ = SSA($\mathbf{s}_1, I = (2,...,l), l$) $+ mean(\mathbf{s}_1)$ trend removal
$\widetilde{\mathbf{s}}_2$ = SSA($\mathbf{s}_2, I = (2,...,l), l$) $+ mean(\mathbf{s}_2)$ trend removal
$\widetilde{\mathbf{s}}_3$ = SSA($\mathbf{s}_3, I = (2,...,l), l$) $+ mean(\mathbf{s}_3)$ trend removal
for $n = 1 : num\ (length(i_{min}))$ number of local minima
$q = \widetilde{\mathbf{s}}_2(i_{min}(n) - \tau_1 : i_{min}(n) + \tau_1) \times \widetilde{\mathbf{s}}_3(i_{min}(n) - \tau_1 : i_{min}(n) + \tau_1)$
$h(n) = min(q)$
end
Determine left/right heel contacts:RHC, LHC,$j = 1, k = 1, i = 1 : num - 2$
if $mean(\widetilde{\mathbf{s}}_1(h(i) : h(i+1))) > mean(\widetilde{\mathbf{s}}_1(h(i+1) : h(i+2)))$
$RHC(j) = h(i), j = j + 1$
else $LHC(k) = h(i), k = k + 1$
$i = 1, ..., length(RHC) - 1,\ j = 0 : N - 1$
-Segment the ML axis data ($\widetilde{\mathbf{s}}_1$) using $RHC(i+j) : RHC(i+j) + q - 1$
-Group N gait cycles into a matrix, $\mathbf{X_c}$
-Apply SVD on the matrix to get the main gait cycle $\mathbf{g}_c$, see eq. (9-11)
similarity = LCSS($\mathbf{g}_c, \mathbf{c}_j$)
if (*similarity*$< \tau_2$) remove the raw cycle, find the next one
otherwise Re-align the raw cycle by resampling
-Detect $LTO(i+j)$
- Set template $r = a_1 exp^{\frac{(t-b_1)^2}{2z_1^2}} + a_2 exp^{\frac{(t-b_2)^2}{2z_2^2}}$
find the local maximum of the $\mathbf{g}_c$ before $g_i = \max\limits_{j=0,..,N-1} LHC(i+j)$
-which is the corresponding point of the first local maximum of the template on $\mathbf{g}_c$ using LCSS($r, \mathbf{g}_c(1 : g_i)$)
find the corresponding point on the raw cycle $\mathbf{c}_j$ using LCSS($\mathbf{g}_c, \mathbf{c}_j$)
-Detect $RTO(i+j)$
find first local minimum of the $\mathbf{g}_c$ after $\min\limits_{j=0,..,N-1} LHC(i+j)$
find the corresponding point on the raw cycle $\mathbf{c}_j$ using LCSS($\mathbf{g}_c, \mathbf{c}_j$)
-Extract the time stamps using the $RHC(i)$, $LHC(i)$, $LTO(i)$, $RTO(i)$
estimate the gait parameters, e.g. swing, stance and stride times

E. Real-time detection of gait events

For some applications such as generating locomotion for paraplegic patients using FES, it is necessary to detect gait events in real time. In this section a general framework is proposed for real time detection of heel contacts. The proposed method can be extended in future studies for detection of both heel contact and toe-off events with proper validation to be used in appropriate applications. An experiment is performed in which the acceleration signals are recorded using the e-AR sensor (e-AR lite version which has a real-time clock) with sampling frequency of 100 Hz from a healthy subject walking along the corridor and turning at the end of corridor to continue walking. The raw accelerations of three axes are shown in Fig. 8(a). As evident from this figure and confirmed experimentally, the turning period creates smoothed accelerations on SI axis.

In Fig. 8(b) a window with the size of 300 samples (from the 1^{st} sample to the 300^{th} sample) from AP axis is used to create an initial trajectory matrix representing a subspace based on $\mathbf{U}$, $\mathbf{\Sigma}$, and $\mathbf{V}$ matrices ($I_j = 1, ..., 30$) (see Appendix). For another window of 300 samples having zero overlap with the first window as shown in the Fig. 8(b), the $\mathbf{U}''$, $\mathbf{\Sigma}''$, and $\mathbf{V}''$ matrices are updated to estimate the new trajectory matrix using the 30 first eigenvectors, then its columns from $(k+1)^{th}$ column to the $(2k)^{th}$ column are used to reconstruct the signal of AP axis in the new segment without applying the SSA algorithm (see Appendix). This is equivalent to applying SSA to the trajectory matrix obtained from the second window. It can be seen from Fig. 8(b) that subspace learning is effective for reconstruction of the acceleration signals by projecting the trajectory matrix to the learned subspace.

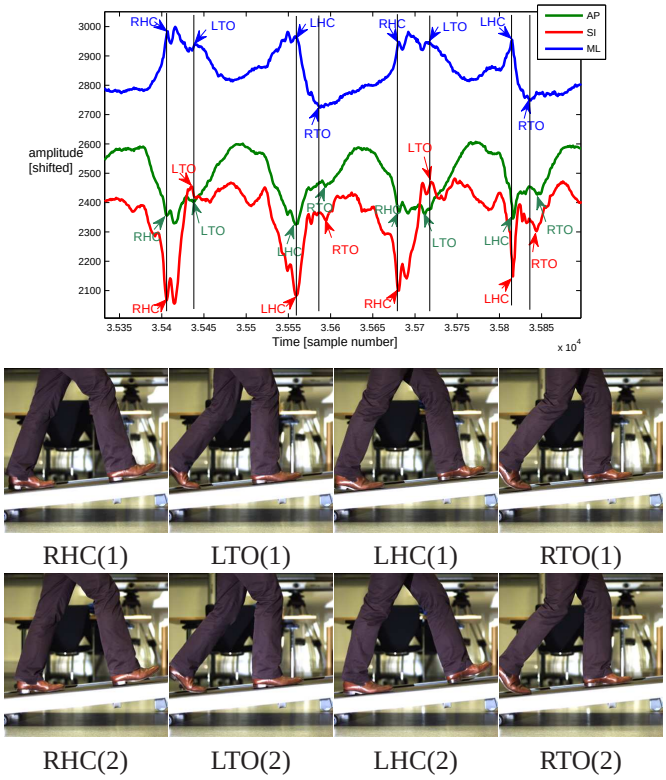

Fig. 3. *Top row*: Example 3-axis e-AR signals corresponding to two consecutive gait cycles. *Second and third rows*: The images given by the high-speed camera for main gait events which are shown using arrows on the signal of ML axis.

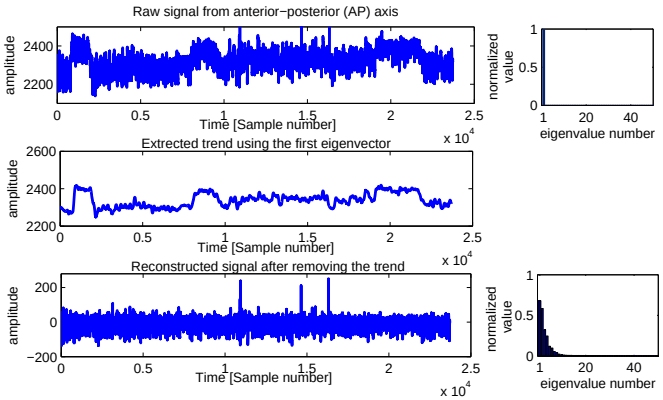

Fig. 4. The signal of the AP axis, before and after trend removal and the corresponding eigenvalues.

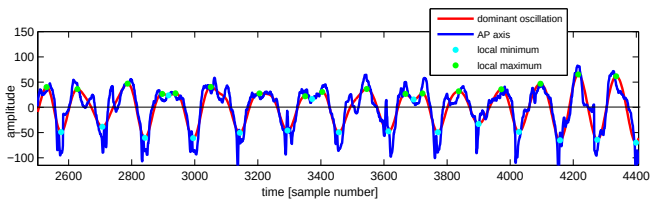

Fig. 5. The dominant oscillation of AP axis and the local minima and maxima points. The invalid local minima/maxima points for detection of heel contacts should be removed.

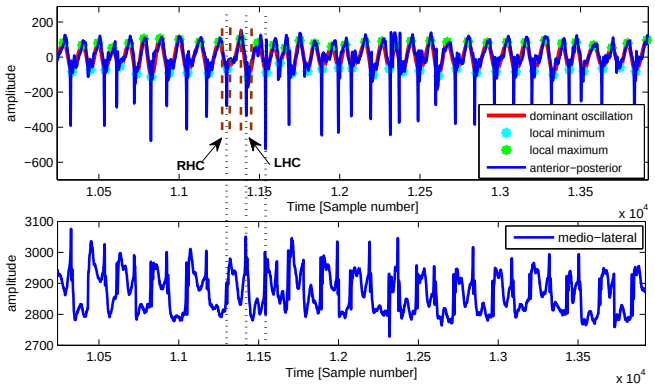

Fig. 6. *Top row*: The signal of the AP axis and its dominant oscillation, and the local minimum and maximum points are shown. It can be seen from this plot that there are no distinctive peaks for all the heel contacts. *Bottom row*: The signal from the ML axis is used to determine the left and right heel contacts.

To find the heel contacts the signal of AP axis is used. The dominant oscillation plus the trend of the AP axis is formed from the first window (of size 300 samples) to create the initial subspace $(r = 1, ..., 3)$. Since the sampling frequency is low we move each window by one point. Then for each new time point the trajectory matrix is re-generated based on:

$$\mathbf{X}^{t-1} = [x_{ij}] = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k]$$

$$\mathbf{X}^{t} = \begin{bmatrix} s_1 & s_2 & s_3 & \dots & s_{k-1} & s_k \\ s_2 & s_3 & s_4 & \dots & s_k & s_{k+1} \\ s_3 & s_4 & s_5 & \dots & s_{k+1} & s_{k+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ s_l & s_{l+1} & s_{l+2} & \dots & s_{n-1} & s_n \end{bmatrix} \quad (12)$$

$$\mathbf{X}^{t} = [\mathbf{X}^{t-1}(\mathbf{x}_2 : \mathbf{x}_k) \ [\mathbf{x}_k(2) \ \mathbf{x}_k(3) \dots \mathbf{x}_k(l) \ s_n]^T] \quad (13)$$

Where $\mathbf{X}^{t-1}$, $\mathbf{X}^{t}$ are the trajectory matrix of the previous and current iteration (see Equation (1)) and s_n is the new time point. The iteration index t increases as new new point s_n is arrived. Therefore, based on the Appendix the $\mathbf{U}''$, $\mathbf{\Sigma}''$, and $\mathbf{V}''$ matrices will be updated for each new trajectory matrix constructed using the new time point.

A matrix considering the second and third largest eigenvalues is generated to find the new point of the dominant oscillation based on the learning algorithm. In this step there is no need to apply the diagonal averaging as in the SSA algorithm since only one time point is added to the system at each iteration to extract a new time point of the dominant oscillation (y_n) from the last element of the constructed matrix (matrix $\mathbf{Y}$ in Algorithm 3). The matrix $\mathbf{Y}$ does not need to be constructed for all its elements since only one (the last) element is enough to estimate the new point of dominant oscillation however for simplicity of understanding, its formulation is shown in Algorithm 3. The differentiation of the two consecutive points of dominant oscillation has a value very close to zero at local maxima and minima points. A constraint can be set to detect a point on the AP axis having a negative value for the dominant oscillation (local minimum) and a value very close to zero for the differentiation of dominant oscillation (see Fig. 8(c)). Then a window with the size of τ centered at the detected local minimum of the dominant oscillation has been formed to detect the local minimum point using both AP and SI axes representing one heel contact. An example is shown in Fig. 8(d) where the right and left heel contacts are detected in an online manner in which τ is set to 30 samples. Also we implemented the algorithm using Matlab software (version R2012a(7.14.0.739) with 64-bit) on a CPU Intel Core (version i7-2600 with 8-core running at 3.4GHz) and for processing 600 samples (as shown in Fig. 8),

Algorithm 3 Real time detection of heel contacts

- create the $\mathbf{U}$, $\mathbf{\Sigma}$ and $\mathbf{V}$ based on the trajectory matrix $\mathbf{X}^0_{I_j} = [\mathbf{x}_1, ..., \mathbf{x}_k]$, $I_j = 1, ..., 3$, set $t = 0$
- for each new time point s_n do the following
- $t = t + 1$
- update the trajectory matrix $\mathbf{X}^t_{I_j} = [\mathbf{X}^{t-1}_{I_j}(\mathbf{x}_2 : \mathbf{x}_k)\ [\mathbf{x}_k(2)\ \mathbf{x}_k(3) \dots \mathbf{x}_k(l)\ s_n]^T]$
- update $\mathbf{U}''$,$\mathbf{\Sigma}''$ and $\mathbf{V}''$ based on $\mathbf{X}^t_{I_j}$ (as matrix $\mathbf{C}$, see Appendix) $r = 1, ..., 3$
- $\mathbf{Y} = \mathbf{U}''(:, 2:3)\mathbf{\Sigma}''(2:3, :)\mathbf{V}''(k+1:2k, :)^T$ as dominant oscillation, y_n = $\mathbf{y}_k(l)$
  calculate $y_1 = y_n - y_{n-1}$
  if ($y_1 \simeq 0$ and $y_n < 0$)
  detect local minima from multiplication of AP and SI axis
  using a window with the size of τ
- update the learned subspace

the elapsed time was obtained as 4.988205 seconds. Since the time difference for arrival of each new sample is 10 milliseconds, the whole processing time obtained (4.988205 seconds) is less than the total timing required for arrival of all measurements (600 $\times$ 10 milliseconds = 6 seconds). That means a real-time implementation of algorithm.

The error of subspace reconstruction based on the norm of $\mathbf{H}$ matrix (see Appendix) is calculated for each new trajectory matrix. It can be seen from Fig. 8(e) that the error of reconstruction is higher for the AP axis at the points that the subject is turning where there are small changes in the AP axis signal with more changes for the SI axis. Although the use of LCSS helps to get more accurate estimations for toe-off events, for real-time applications they can also be located on the ML axis as the first peak after each detected heel contacts.

The objective in this section was to propose a method that can be used for real-time detection of gait events using SSA as a basis for future studies. The advantage is that it is not necessary to apply SSA sequentially to the data segments instead to use the learned subspace a nd projecting the trajectory matrix.

Also using subspace learning and the reconstruction error it is possible to detect changes in the walking behavior. The proposed framework in this section requires proper validation using pressure sensing platform in future studies for online detection of gait events. In the next section, the method explained in Section II-D for estimation of gait parameters is validated using three different datasets.

III. Experimental Results

In order to validate the proposed method for gait parameter estimation, three sets of experimental data have been analysed. The first dataset has been used to validate the accuracy of the estimated gait parameters versus force-data while the second and third datasets have been used to investigate the feasibility of gait parameters estimation for future clinical applications. In the following sections, the results of applying the proposed method to the datasets are provided.

A. Validation of estimated gait parameters

For validation of gait parameters estimated by analysing the e-AR signals, a laboratory-based experiment has been performed. In the experiment 10 healthy adults walked on a force-plate instrumented treadmill (Gaitway Treadmill, Kistler Instrument Corp., Amherst USA) at a speed of 3.2 km/h for 20 minutes. After each 2 minute interval, the treadmill incline was increased by 2%. Detection of gait events using an inclined treadmill is more difficult than horizontal line treadmill since walking on an inclined treadmill results into smoothed accelerations that make it more difficult to detect the gait events. In this experiment the gait parameters are mainly validated on inclined treadmill since only 10% of the data are related to straight walking on the treadmill.

The raw acceleration signals given by the high-speed e-AR sensor having a sampling frequency of 130 Hz were segmented into 1 minute windows. For each minute, which also related to the same percentage

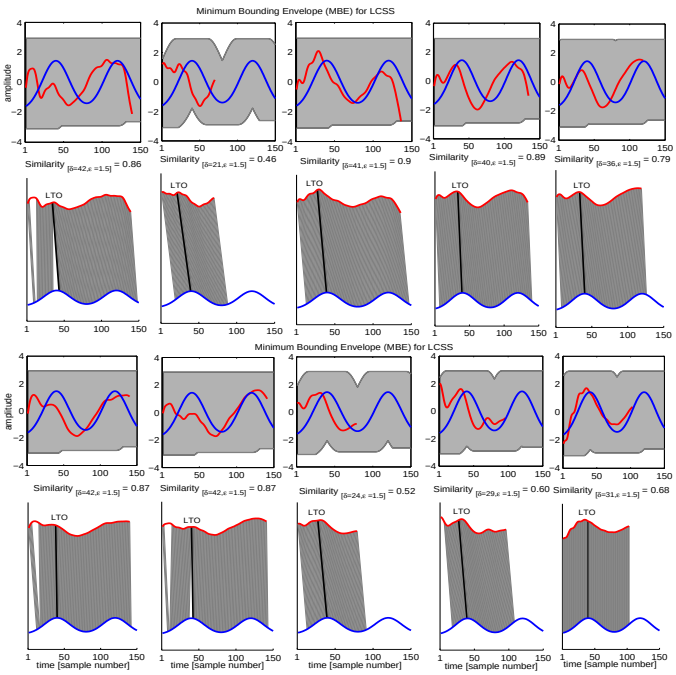

Fig. 7. Ten extracted main gait cycles (red color) segmented to the maximum of LHCs in time domain for their corresponding N grouped cycles and the constructed template (blue color) with the point correspondence output of LCSS to detect LTO on the main gait cycle. The Minimum Bounding Envelop (MBE) is shown in grey color which is a region that covers all matching areas within δ in time and ϵ in space.

of incline, the gait parameter estimation method as explained in Section II-D is used to estimate the essential gait parameters. Since for each 1 minute of the data, the percentage of the incline is constant, 15 gait cycles (N) have been selected and the parameters are then estimated. Therefore, the trend of the gait signals are first removed then RHC and LHC are estimated. The SVD is applied to 15 gait cycles. For grouping gait cycles, the parameter q in Equation (9) is set to 310. The LCSS is then evaluated by having the main gait cycle derived by SVD and that from the raw data as the input.

The value of τ_2 parameter in Algorithm 2 controls the number of discarded cycles. The value of τ_2 is set to be 0.65 in this study, which is empirically defined. Therefore, after segmenting the gait cycles using the estimated RHCs from the ML axis, the cycles that have less than 0.65 similarity based on the point correspondence output of the LCSS will be removed. To apply the SSA algorithm, the embedding dimension is set to be 100 in this study.

In the LCSS algorithm, the value of the ϵ parameter which represents matching in the space can be set to half of the standard deviation of the input time-series which has been shown to provide good results [22]. The warping length δ in the LCSS algorithm is a percentage of the time-series length. Based on the given experiments in [22], the performance of the warping from 5% to 20% of the time-series length is evaluated. By increasing the value of δ, the space search will be larger and less accurate results will be obtained. In the proposed algorithm, the LCSS algorithm is called for two different purposes:

- Finding the similarity of the main gait cycle and each raw cycle (to remove less useful cycles based on similarity value) and also to locate the RTO/LTO detected using main gait cycle on each raw cycle.
- To find the LTO on the main gait cycle, based on a predefined template, to be located later on each raw cycle.

For the first purpose, the value of the ϵ is set to 0.3 multiplied by the minimum of the standard deviation of the two input time-series.

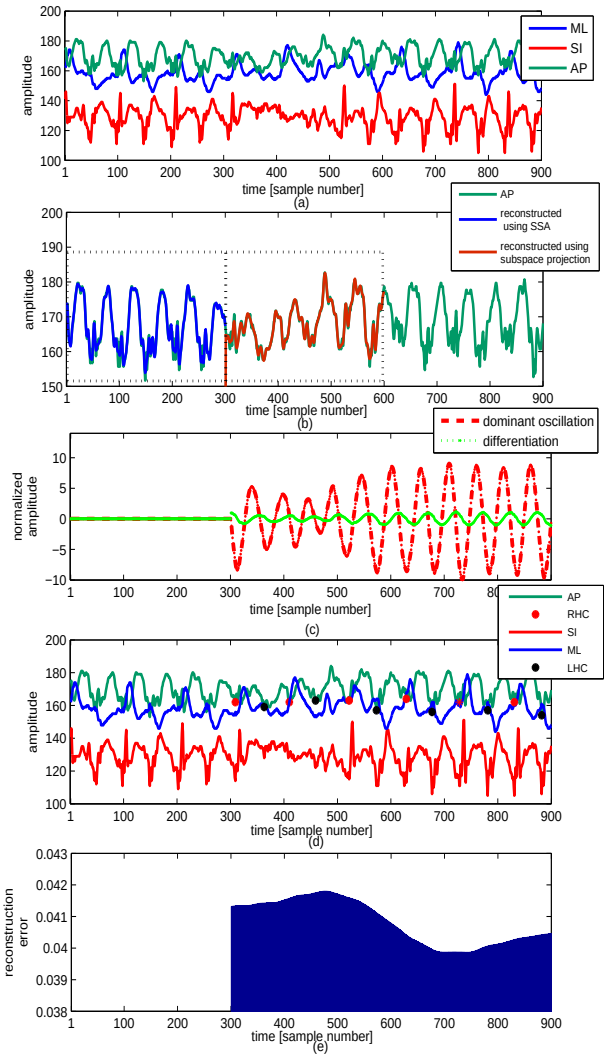

Fig. 8. (a) raw accelerations related to waking on a corridor. (b) reconstructing a segment of the AP axis based on projecting the trajectory matrix to the learned subspace using the previous segment. (c) estimation of dominant oscillation to detect the local minimum points using differentiation of dominant oscillation. (d) located RHC and LHC on the AP axis based on real time detection of heel contacts. (e) calculated error (normalised) of reconstruction in the learned subspace using the constructed trajectory matrix for each time point.

The value of the δ is set to 0.15 multiplied by the minimum size of its inputs (15% of the time-series length). For the template r in Algorithm 2, a_1,b_1 and $2z_1^2$ are set to 10, 40 and 600, a_2,b_2 and $2z_2^2$ are set to 10, 120 and 600 where the time index t starts from 1 to 150. Since the template length is about half the cycles (310), then for the LCSS algorithm (second purpose) the δ is set to 0.30 multiplied by the minimum size of its inputs (30% of the time-series length) and ϵ is set to 1.5 multiplied by the minimum of the standard deviation of the two input time-series.

The RTO and LTO are estimated using the output of LCSS and the peaks of the main gait cycle. In Fig. 9, three example raw cycles and the corresponding main gait cycles are shown. The last raw cycle is removed due to a very low similarity value to the main gait cycle while the first raw cycle has a big similarity to the main gait cycle. In the middle plot of the Fig. 9, it can be seen why we choose to detect the peaks from the main gait cycle rather than the raw cycle to ensure the robustness of the algorithms. In addition LTO can be better detected using the predefined template and applying the LCSS algorithm as shown in Fig. 7.

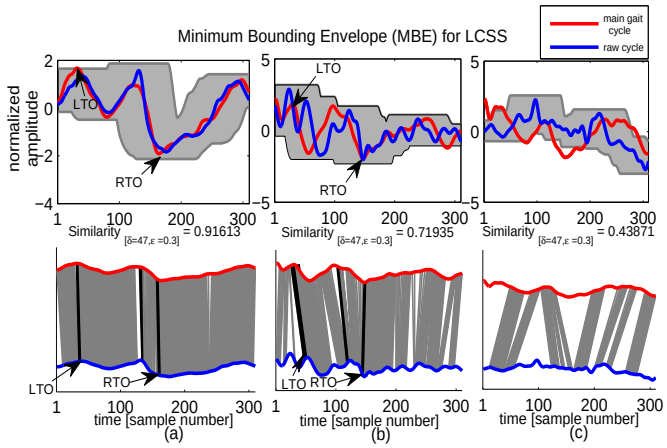

Fig. 9. The main gait cycle given by the SVD algorithm is compared to each cycle using the LCSS method. In the top plot the main gait cycle is very similar to the raw cycle, the point correspondence is shown in the second plot. For the raw cycle shown in the third plot, it has about 71% similarity to the main gait cycle. The raw cycle shown in the fifth plot has about 43% similarity to the main cycle therefore it will be removed not to have contribution in estimation of gait parameters.

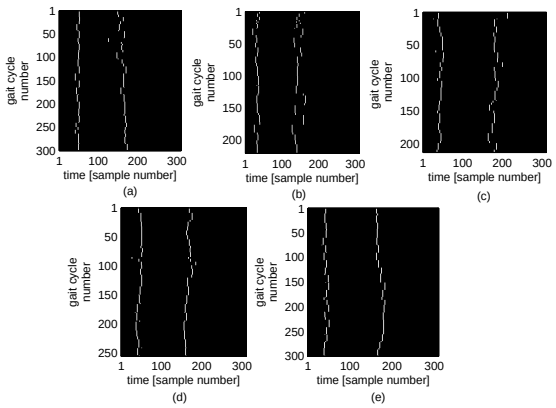

Fig. 11. The peaks detected as LTO and RTO for the 5 healthy adults are shown in white colour which create a trace in the vertical axis.

By using the resampling technique, based on the main gait cycle and output of LCSS, the gait cycles are enhanced. The results are shown in Fig. 10. In this figure, cycles from each 1 minute data are selected, then all the cycles towards the end of the experiment have been added to a matrix and their corresponding 2D pattern is created. It can be seen from this figure that the enhancement of the gait cycle preserves the overall trend of gait peaks while making it easier to estimate the essential gait events. For the 5 healthy adults shown in Fig. 10, the estimated LTO and RTO indices are shown in white in Fig. 11. Having the index of all gait events, the gait parameters (e.g. swing, stance and stride) for the left and right foot are estimated.

The same gait parameters are also estimated using the force data. Force plate detection of heel contact and toe-off was performed on the treadmill sampled at 500 samples/s. The algorithm was implemented in Matlab (The Mathworks Inc.) to partition force traces into right and left steps. A threshold of 10% of bodyweight was used to determine heel contact, and 5% was used to determine toe-off. These relatively large thresholds were needed to account for the increased noise on the force traces caused by the moving belt. Trials in which there was no clear separation of consecutive strides (i.e. when the front foot contacted either the anterior or posterior force-plate before the rear-foot left the same plate) were discarded.

For each subject having the same percentage of incline (and per

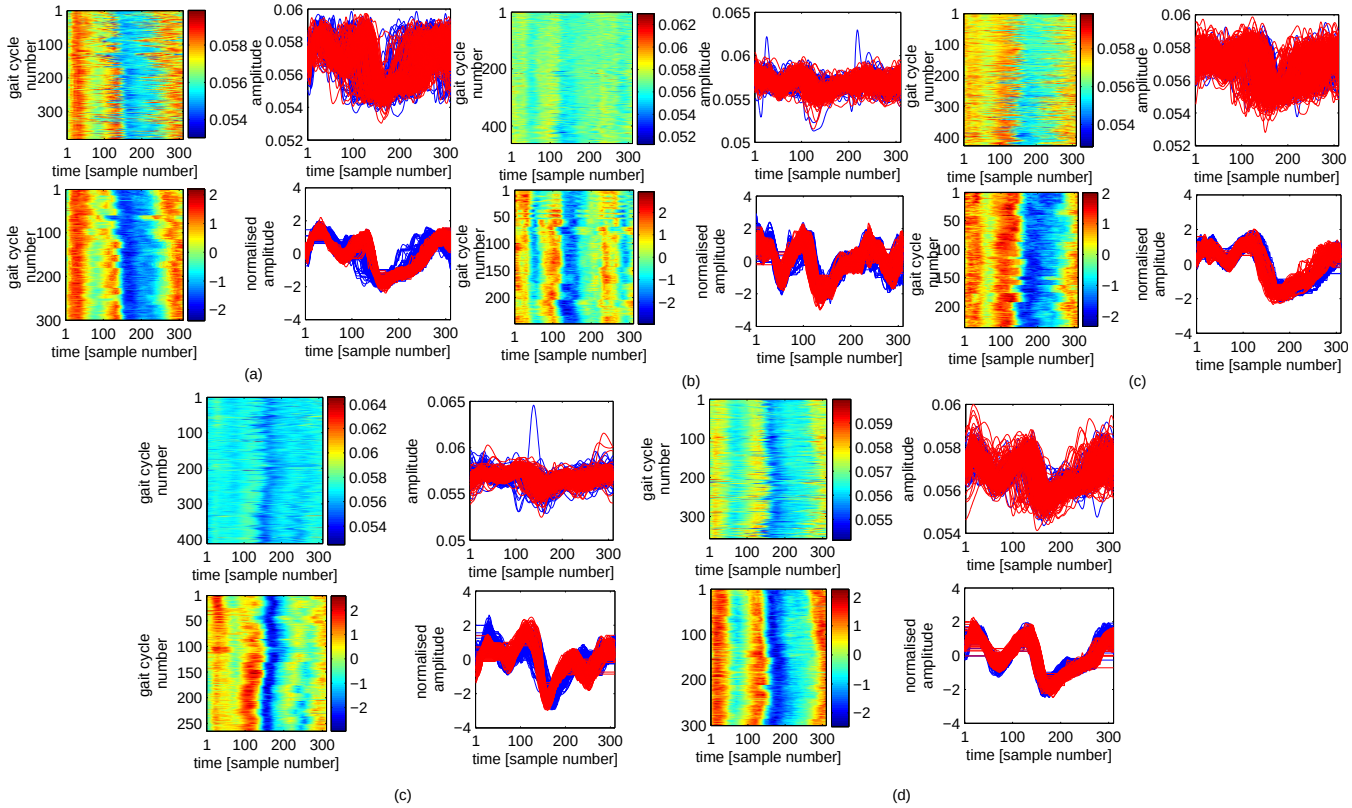

Fig. 10. The raw gait cycles and the enhanced signals by LCSS and SSA. The results are shown for 5 healthy adults walking on the treadmill. In each plot, the raw cycles (top) and the enhanced ones (bottom) from the ML axis are grouped together having an increased incline angle as the gait cycle number increases. The created images of the consecutive cycles are also shown. The first half of cycles are shown in blue and the second half in red.

minute), there is one estimation from the e-AR and one estimation using the force data. The corresponding Bland Altman plot is depicted in Fig. 12. for right/left swing, right/left stance and right/left stride times. In Table I, the mean and standard deviation of the relative and absolute difference between estimations are shown. It can be seen from this table and Bland Altman plot (Fig. 12.) that the best estimation is for the stride times. The absolute error of estimations for each subject in different incline degree is shown in Fig. 13. There are some outliers for some subjects, in which the visual inspection of accelerations for 10^{th} subject shows ML accelerations very much far away normal ones that can be due to inconsistent/incorrect walking of the person on the treadmill. Also for this subject due to discard of many gait cycles less estimated values are obtained.

TABLE I

THE MEAN AND STANDARD DEVIATION OF THE DIFFERENCE BETWEEN THE ESTIMATED GAIT PARAMETERS (IN SECONDS) USING E-AR AND FORCE DATA

	relative error		absolute error	
Parameter	mean	SD	mean	SD
left swing	0.0231	0.0348	0.0332	0.0253
right swing	0.0082	0.0605	0.0379	0.0478
left stance	−0.0271	0.0380	0.0358	0.0299
right stance	−0.0095	0.0563	0.0380	0.0426
left stride	−0.0034	0.0281	0.0175	0.0222
right stride	−0.0013	0.0285	0.0183	0.0218

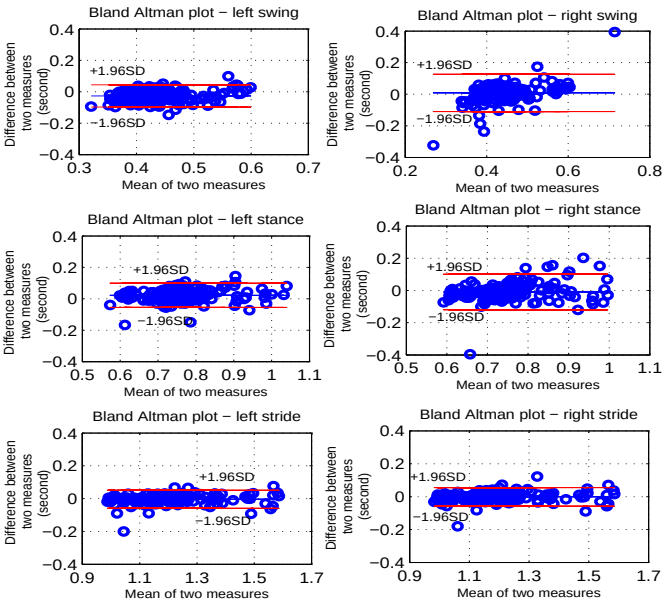

Fig. 12. The Bland Altman plot for estimation of left/right swing, left/right stance and left/right stride using the e-AR and force-plate treadmill. The horizontal axis for each plot is the average of estimation using the two methods and the vertical axis is the difference between the two methods.

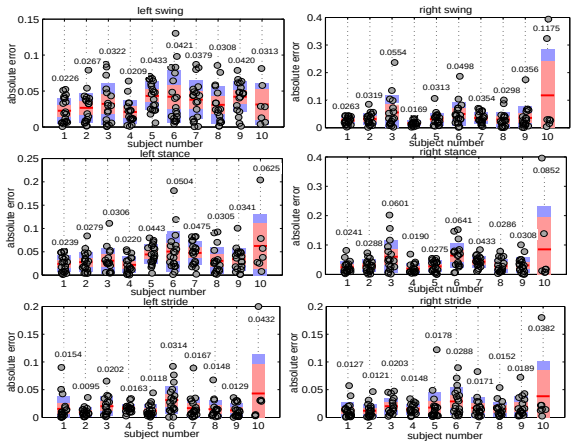

Fig. 13. The absolute error of gait parameters for each subject using e-AR having the force data as the reference. The mean absolute value for each subject is shown.

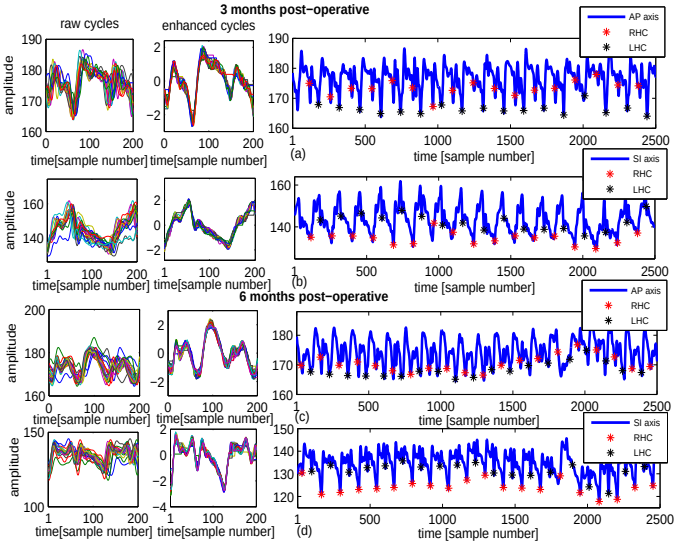

Fig. 14. Demonstration of changes in gait patterns and detected RHCs and LHCs 3 and 6 months post operative. (a) The raw and enhanced cycles of AP axis 3 months post-operative. (b) The raw and enhanced cycles of SI axis 3 months post-operative. (c) The raw and enhanced cycles of AP axis 6 months post-operative. (d) The raw and enhanced cycles of SI axis 6 months post-operative.

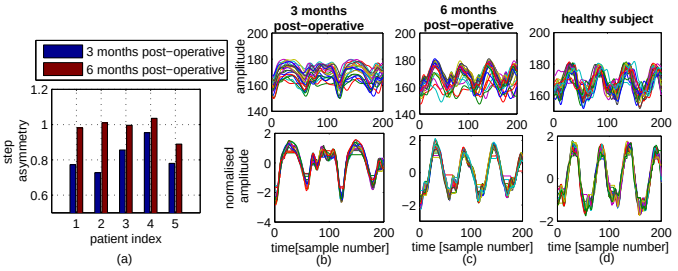

Fig. 15. (a) step time asymmetry estimated for 5 patients 3 and 6 months after operation. (b,c) raw and enhanced cycles of AP axis 3 and 6 months post-operative. (d) raw and enhanced cycles of AP axis from a healthy subject.

B. Gait Analysis for Rehabilitation

Estimation of gait parameters has been used in many studies to quantify the differences between healthy subjects and patients with gait impairments. One important application in gait analysis is to monitor recovery of orthopedic patients after surgery. Having the information from wearable or ambient sensors, the estimated gait parameters as useful features can be used to monitor rehabilitation of orthopedic patients. By analysis of signal characteristics of accelerations, more features can be obtained. Adding more useful features is helpful to create reliable monitoring systems for rehabilitation.

For the next experiment the e-AR lite sensor a the sampling frequency of 100 Hz was used to record the acceleration signals in three axes. The data was recorded from patients who were recovering from reconstructive surgery following severe lower limb trauma (open tibial fracture). This cohort were assessed in a clinical setting using the 6 minutes walk on a 15m corridor. Data was recorded 3 months and 6 months post-operative.

The accelerations from AP and SI axes for one patient 3 and 6 months post-operative are shown in Fig. 14. The detected left and right heel contacts are shown as asterisks in this figure. It can be seen from this figure that despite of a big level of asymmetry in walking, the algorithm can detect heel contacts. This is due to using all three axes in determination of heel contacts. In fact the ML axis for determination of the side of heel contact (left/right) plays an important role which prevents invalid detection of heel contacts. The gait cycles of AP and SI axis are segmented from the RHCs with a window of 200 sample size. The enhanced cycles by using the LCSS algorithm are also shown. It can be seen from this figure that the gait cycles are changing in shapes in different time points after operation which can be used for extraction of additional features. This can be done in future studies by using appropriate template matching approaches having the template for normal gait cycles. As another example the raw cycles and enhanced cycles of AP axis for one patients 3 months and 6 months post-operative are shown in Fig. 15(b,c). The AP axis cycles for healthy subjects are shown in Fig. 15(d). These figures show the importance of direct analysis of accelerations for monitoring rehabilitation.

The proposed method has been applied to five healthy subjects and five patients to estimate the step time asymmetry. For five healthy subjects, the mean and standard deviation of the step asymmetry were obtained as 1.0082 and 0.0309 respectively. In Fig. 15(a), for each patient the estimated step time asymmetry are shown for 3 and 6 months post-operative. As it can be seen from this figure, the step time asymmetry is approaching to 1 in 6 months post-operative.

In forthcoming studies, we further validate the sensor and the analysis framework for free living environments having pressure sensing platform as the reference. This enables creating a reliable system for home based monitoring of rehabilitation.

IV. Discussion and Conclusion

In this paper, a new method is proposed for estimating gait parameters using e-AR device. In this study, the gait parameter estimation is based on the combined use of SSA and LCSS. Other gait features such as spatial features (signal amplitudes for heel contacts events) are also derived. Experimental validation has shown the practical value of the method for healthcare applications.

The accuracy of the algorithm for gait parameter estimation is validated using a force-plate instrumented treadmill. Depending on the clinical application and the purpose of estimating gait parameters, the reliability of gait variables can be more assessed. This includes test-retest reliability assessments and obtaining inter- and intra-class correlation coefficients (ICC) for further analysis. In addition, the results for patients with lower limb trauma are provided to demonstrate the practical use of the sensor in none-lab environment.

It is worth noting that for gait analysis using wearable technologies, it is important to reduce the number of sensors to ensure simplicity while maintaining reliability and accuracy of the system. This is particularly relevant in clinical applications for long term continuous monitoring of patients, as well as the elderly population. In this paper, information from reference technologies such as a high-speed camera and a force-plate instrumented treadmill has been

used to validate the estimated gait parameters derived from the e-AR sensor. In addition a real time implementation is proposed to be further improved in future studies which require online detection of gait events. From the results derived, it is evident that the proposed method in this paper for accelerometry based gait analysis has the following advantages.

- Use of a single light-weight ear worn sensor
- All processing can be performed on-node and therefore significantly reduces wireless-transmission overhead
- Ergonomic design of the sensor to ensure long-term patient compliance

Another important feature of the proposed algorithm is that for the filtering process, there are no strong assumptions on the statistics of the signals. In summary, we have developed a practical gait analysis platform that is suitable for both laboratory and free-living environments. The results derived demonstrate the practical clinical value of the method.

The proposed method in this paper can provide a framework for accelerometry based gait analysis using a single sensor to create clinically reliable future monitoring systems.

V. APPENDIX

The SVD of the trajectory matrix can be written as:

$$\mathbf{X}_{\mathbf{I}_j} = \mathbf{U}\Sigma\mathbf{V}^T \tag{14}$$

where $\mathbf{U}$ is an $l \times l$ matrix which contains the left singular vectors, Σ is an $l \times k$ matrix of eigenvalues and $\mathbf{V}^T$ is a $k \times k$ matrix which contains the right singular vectors (l is the embedding dimension, and k is defined in Section II-B) and I_j denotes the indices to select the eigenvectors.

$$\begin{aligned} \mathbf{L}_{(r\times k)} &= \mathbf{U}^T_{(r\times l)}\mathbf{C}_{(l\times k)} \\ \mathbf{H}_{(l\times k)} &= \mathbf{C}_{(l\times k)} - \mathbf{U}_{(l\times r)}\mathbf{L}_{(r\times k)} \end{aligned} \tag{15}$$

$$\mathbf{J}_{(l\times l)}\mathbf{K}_{(l\times k)} \xleftarrow{QR} \mathbf{H}_{(l\times k)} \tag{16}$$

where QR denotes the QR decomposition that is performed on the $\mathbf{H}$ matrix as the error of reconstruction after projecting the $\mathbf{C}$ matrix (trajectory matrix of a new observation) onto the learned subspace ($\mathbf{U}$). Matrix $\mathbf{B}$ is defined in [25]:

$$\mathbf{B} = [\mathbf{U}_{(l\times r)} \quad \mathbf{J}_{(l\times l)}] \begin{bmatrix} \mathbf{\Sigma}_{(r\times r)} & \mathbf{L}_{(r\times k)} \\ \mathbf{0}_{(l\times r)} & \mathbf{K}_{(l\times k)} \end{bmatrix} \begin{bmatrix} \mathbf{V}_{(k\times r)} & \mathbf{0}_{(k\times k)} \\ \mathbf{0}_{(k\times r)} & \mathbf{I}_{(k\times k)} \end{bmatrix}^T \tag{17}$$

Considering Equations (14-17), matrix $\mathbf{B}$ is simplified as:

$$\begin{aligned} \mathbf{B} &= [\mathbf{U}_{(l\times r)}\mathbf{\Sigma}_{(r\times r)} \;\; \mathbf{U}_{(l\times r)}\mathbf{L}_{(r\times k)} + \mathbf{J}_{(l\times l)}\mathbf{K}_{(l\times k)}] \begin{bmatrix} \mathbf{V}_{(k\times r)} & \mathbf{0}_{(k\times k)} \\ \mathbf{0}_{(k\times r)} & \mathbf{I}_{(k\times k)} \end{bmatrix}^T \\ &= [\mathbf{U}_{(l\times r)}\mathbf{\Sigma}_{(r\times r)}\mathbf{V}^T_{(r\times k)} \quad \mathbf{U}_{(l\times r)}\mathbf{L}_{(r\times k)} + \mathbf{H}_{(l\times k)}] \\ &= [\mathbf{U}_{(l\times r)}\mathbf{\Sigma}_{(r\times r)}\mathbf{V}^T_{(r\times k)} \quad \mathbf{U}_{(l\times r)}\mathbf{L}_{(r\times k)} + \mathbf{C}_{(l\times k)} - \mathbf{U}_{(l\times r)}\mathbf{L}_{(r\times k)}] \\ &= [\mathbf{U}_{(l\times r)}\mathbf{\Sigma}_{(r\times r)}\mathbf{V}^T_{(r\times k)} \quad \mathbf{C}_{(l\times k)}] \\ &= [\mathbf{X}_{\mathbf{I}_j(l\times k)} \quad \mathbf{C}_{(l\times k)}] \end{aligned} \tag{18}$$

Matrix $\mathbf{B}$ contains the trajectory matrix of the initial subspace and the new observation that can be written as (considering the SVD of the middle matrix in Equation (17)):

$$\begin{aligned} \mathbf{B} &= [\mathbf{X}_{\mathbf{I}_j(l\times k)} \;\; \mathbf{C}_{(l\times k)}] \\ &= [\mathbf{U}_{(l\times r)} \quad \mathbf{J}_{(l\times l)}]\mathbf{U}'\mathbf{\Sigma}'\mathbf{V}'^T \begin{bmatrix} \mathbf{V}_{(k\times r)} & \mathbf{0}_{(k\times k)} \\ \mathbf{0}_{(k\times r)} & \mathbf{I}_{(k\times k)} \end{bmatrix}^T \end{aligned} \tag{19}$$

where

$$\mathbf{U}'\mathbf{\Sigma}'\mathbf{V}'^T = \begin{bmatrix} \mathbf{\Sigma}_{(r\times r)} & \mathbf{L}_{(r\times k)} \\ \mathbf{0}_{(l\times r)} & \mathbf{K}_{(l\times k)} \end{bmatrix} \tag{20}$$

Finally matrix $\mathbf{B}$ is obtained as:

$$\begin{aligned} \mathbf{B} &= \mathbf{U}''\mathbf{\Sigma}''\mathbf{V}''^T \\ \mathbf{U}'' &= [\mathbf{U}_{(l\times r)} \quad \mathbf{J}_{(l\times l)}]\mathbf{U}' \\ \mathbf{\Sigma}'' &= \mathbf{\Sigma}' \\ \mathbf{V}'' &= \begin{bmatrix} \mathbf{V}_{(k\times r)} & \mathbf{0}_{(k\times k)} \\ \mathbf{0}_{(k\times r)} & \mathbf{I}_{(k\times k)} \end{bmatrix} \mathbf{V}' \end{aligned} \tag{21}$$

REFERENCES

[1] J. Parkka, M. Ermes, P. Korpipaa, J. Mantyjarvi, J. Peltola, and I. Korhonen, "Activity classification using realistic data from wearable sensors," *IEEE Transactions on information technology in biomedicine*, vol. 10, no. 1, pp. 119–128, 2006.

[2] M. Mojarradi, D. Binkley, B. Blalock, R. Andersen, N. Ulshoefer, T. Johnson, and L. D. Castillo, "A miniaturized neuroprosthesis suitable for implantation into the brain," *IEEE Transactions on neural systems and rehabilitation engineering*, vol. 11, no. 1, pp. 38–42, 2003.

[3] E. Jovanov, A. Milenkovic, C. Otto, and P. de. Groen, "A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation," *J Neuroengineering Rehanil.*, vol. 1, pp. 2–6, 2005.

[4] K. Aminian, P. Robert, E. Jequier, and Y. Schutz, "Incline, speed, and distance assessment during unconstrained walking," *Med. Sci. Sports Exerc.*, vol. 27, no. 2, pp. 226–234, 1995.

[5] S. J. Bamberg, A. Y. Benbasat, D. M. Scarborough, D. E. Krebs, and J. A. Paradiso, "Gait analysis using a shoe-integrated wireless sensor system," *IEEE Transactions on information technology in biomedicine*, vol. 12, no. 4, pp. 413–423, 2008.

[6] M. Sekine, T. Tamura, M. Akay, T. Fujimoto, T. Togawa, and Y. Fukuo, "Discrimination of walking patterns using wavelet-based fractal analysis," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 10, no. 3, pp. 188–196, 2002.

[7] W. Zijlstra and A. L. Hof, "Assessment of spatio-temporal gait parameters from trunk accelerations during human walking,," *Gait Posture*, vol. 18, pp. 1–10, 2003.

[8] R. M. Nilssen and J. L. Helbostad, "Estimation of gait cycle characteristics by trunk accelerometry,," *Journal of Biomechanics*, vol. 37, pp. 121–126, 2004.

[9] Y. Tochigi, N. A. Segal, T. Vaseenon, and T. D. Brown, "Entropy analysis of tri-axial leg acceleration signal waveforms for measurement of decrease of physiological variability in human gait,," *J. Orthop. Res.*, vol. 30, no. 6, pp. 897–904, 2012.

[10] A. T. Willemsen, F. Bloemhof, and H. B. K. Boom, "Automatic stance-swing phase detection from accelerometer data for peroneal nerve stimulation," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 12, pp. 1201–1208, 1990.

[11] J. M. Jasiewicz, J. H. H. Allum, J. W. Middleton, A. Barriskill, P. Condie, B. Purcell, and R. C. .Li, "Gait event detection using linear accelerometers or angular velocity transducers in able-bodied and spinal-cord injured individuals,," *Gait Posture*, vol. 24, no. 4, pp. 502–509, 2006.

[12] R. W. Selles, M. A. G. Formanoy, J. B. J. Bussmann, P. J. Janssens, and H. J. Stam, "Automated estimation of initial and terminal contact timing using accelerometers; development and validation in transtibial amputees and controls,," *IEEE Trans. Neural. Syst. Rehabil. Eng.*, vol. 13, no. 1, pp. 81–88, 2005.

[13] L. Atallah, B. Lo, R. King, and G. Z. Yang, "Sensor positioning for activity recognition using wearable accelerometers," *IEEE transactions on biomedical circuits and systems*, vol. 4, no. 4, pp. 320–329, 2011.

[14] G. Z. Yang, *Body Sensor Networks.*, Berlin, Germany: Springer-Verlag, 2006.

[15] L. Atallah, A. Wiik, G. G. Jones, B. Lo, J. P. Amis, and G. Z. Yang, "Validation of an ear-worn sensor for gait monitoring using a force-plate instrumented treadmill," *Gait Posture*, vol. 35, no. 4, pp. 674–676, 2012.

[16] B. Lo, J. Pansiot, and G. Z. Yang, "Bayesian analysis of sub-plantar ground reaction force with BSN," *6th international workshop on wearable and implantable body sensor networks, BSN 2009*, pp. 135–139, 2009.

[17] N. Golyandina, V. Nekrutkin, and A. Zhigljavsky, "Analysis of time-series structure: SSA and related techniques," *Chapman and Hall/CRC*, 2001.

[18] S. Sanei, T. K. M. Lee, and V. Abolghasemi, "A new adaptive line enhancer based on singular spectrum analysis," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 2, pp. 428–434, 2012.

[19] F. Ghaderi, H. R. Mohseni, and S. Sanei, "Localizing heart sounds in respiratory signals using singular spectrum analysis," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 12, pp. 3360–3367, 2011.

[20] D. S. Hirschberg, "A linear space algorithm for computing maximal common subsequences," *Comm. ACM*, vol. 18, no. 6, pp. 341–343, 1975.

[21] G. Das, D. Gunopulos, and H. Mannila, "Finding similar time series," *Proc. Conf. Principles of Knowledge Discovery and Data Mining*, pp. 454–456, 1997.

[22] M. Vlachos, M. Hadjieleftheriou, D. Gunopulos, and E. J. Keogh, "Indexing multi-dimensional time-series with support for multiple distance measures," *Proc. ACM Special Interest Group Knowledge Discovery and Data Mining (SIGKDD 03)*, pp. 216–225, 2003.

[23] D. Jarchi, L. Atallah, and G. Z. Yang, "Transition detection and activity classification from wearable sensors using singular spectrum analysis," *Ninth international workshop on wearable and implantable body sensor networks, BSN 2012*, pp. 136–141, 2012.

[24] R. R. Torrealba, J. Cappelletto, L. Fermin-Leon, J. C. Grieco, and G. Fernndez-Lpez, "Statistics-based technique for automated detection of gait events from accelerometer signals," *Electronics Letters*, vol. 46, no. 22, pp. 1483–1485, 2010.

[25] G. H. Golub and C. F. Van. Loan, *Matrix Computations*, The Johns Hopkins University Press, 1996.