Towards Segmentation and Labelling of Motion Data in Manufacturing Scenarios

António Santos[1], João Rodrigues[1,3], Duarte Folgado[1,2], Sara Santos[2], Carlos Fujão[3], and Hugo Gamboa[1,2]

[1] Laboratório de Instrumentação, Engenharia Biomédica e Física da Radiação (LIBPhys-UNL), Departamento de Física, Faculdade de Ciências e Tecnologia, FCT, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

[2] Associação Fraunhofer Portugal Research, Rua Alfredo Allen 455/461, 4200-135 Porto, Portugal

[3] Volkswagen Autoeuropa, Quinta da Marquesa, 2954-024 Q.ta do Anjo, Portugal

{alm.santos@campus.fct.unl.pt, jmd.rodrigues@fct.unl.pt, sara.santos, duarte.folgado, hugo.gamboa}@fraunhofer.pt

Abstract. There is a significant interest to evaluate the occupational exposure that manufacturing operators are subjected throughout the working day. The objective evaluation of occupational exposure with direct measurements and the need for automatic annotation of relevant events arose. The current work proposes the use of a self similarity matrix (SSM) as a tool to flag events that may be of importance to be analyzed by ergonomic teams. This way, data directly retrieved from the work environment will be summarized and segmented into sub-sequences of interest over a multi-timescale approach. The process occurs under 3 timescale levels: Active working periods, working cycles, and in-cycle activities. The novelty function was used to segment non-active and active working periods with an F1-score of 95%. while the similarity function was used to correctly segment 98% of working cycle with a duration error of 6.12%. In addition, this method was extended into examples of multi time scale segmentation with the intent of providing a summary of a time series as well as support in data labeling tasks, by means of a query-by-example process to detect all subsequences.

Keywords: Self-similarity Matrix · Time Series · Industry · Musculoskeletal Disorders, Inertial · Segmentation, · Summarization · Unsupervised · Labeling

1 Introduction

1.1 Work Related Disorders and Risk Evaluation

Musculoskeletal disorders are a broad variety of health conditions that affect the locomotor system. These are usually described by pain and a reduction of people's mobility, with some of the more common conditions including osteoarthritis, back and neck pain, fractures associated with bone fragility, injuries and systemic inflammatory conditions such as rheumatoid arthritis.

The burden of impaired musculoskeletal health can be very deteriorating for individual human lives. Pain and disability lead to the inability of performing daily life activities, precluding an active participation in social activities and possibly limiting the work performance of the population. [6][7] Furthermore these may also promote various physical and mental comorbidities.[8][9][10]

The present work will target work-related musculoskeletal disorders (WMSDs) i.e. MSDs which are induced or aggravated by work and the circumstances of its performance. [11] Workspaces can be pointed out as a significant contributor to MSDs [12], with the added advantage of being also a controlled environment that can be adapted towards health policies concentrated on primary prevention. To project a proper intervention for the work-space it is very important to firs evaluate each separate job-environment relationship. It's not effective to design a "one-fits-all" system for different work settings as it requires specific adaptations designed by trained ergonomist specialists.

In this sense, there have been developed several systems to assess the level of occupational exposure in each workplace. These techniques fall mostly into three main categories, which differ from each other depending on the type of data acquisition method used: (1) Self-report from workers, a protocol that enables workers to individually provide the data about their personal work experience. This can be achieved through the use of questionnaires, diaries, personal interviews or checklist surveys so that they might be further evaluated by experts; (2) Observational methods, which consists on the visual surveillance of the work routine by trained experts, where, conventionally, their metrics are then provided by pre-designed ergonomic risk assessment sheets; (3) Direct Measurements, that propose the use of new technologies to retrieve more precise/accurate information from the work environment. This is made through the usage of sensors usually attached to specific subjects performing work routines.

It has been proven, overall, that an ergonomic assessment of the work environment can indeed reduce the prevalence of MSDs by helping to identify possible risk factors, and in the interest of this work we'll argue that there are also advantages in a transition towards more direct measurement techniques. As a more objective form of describing the interaction between workers and the workplace doesn't suffer from any variability dependent on the worker or specialist observer, as the "Self-reporting" and "Observational" methodologies do.

However, despite direct measurements affording a more objective representation of the reality, it still has limitations regarding greater costs associated with its implementations and the narrow success on a practical environment. Being mostly used successfully under controlled simulated laboratories. Another important concern is its interpretability with the analysis of motion sensors being complex and requiring specific insights to retrieve relevant knowledge.

1.2 Motivation

The current information era, in conjunction with technological developments, is promoting a shift in the way industries manage and perform their work. With the inclusion of sensing devices, dashboards and machine learning algorithms, many tasks can be better evaluated with direct quantitative measures. This can be made for machines to prevent breakdowns and optimize their performance, but can also be used for humans, namely their occupational health, to prevent work-related disorders. The prevention of work-related disorders implies a specialized intervention to perform changes in the workplace, the worker's training process and/or other organizational strategies. The process of deciding a need for intervention implies a previous risk evaluation of the worker's routines. The direct measures from inertial sensors provide relevant information for a more objective and personalized risk assessment [1]. However, before the actual ergonomic risk evaluation, a considerable number of steps in preparing the data are needed. Data preparation is an important part of any analysis or task that requires further deployment in supervised methods, namely the time-consumingtime consuming segmentation and labelling process. This is an impediment that has become extensively common in various types of industries and social sectors, stalling sometimes further development. As the rate of information that we acquire from the world tends to increase there is a greater need for processes that can automaticaly retrieve useful knowledge from that data, as people don't have that same capacity.

Tools capable of segmenting time series to support and ease the labelling process of motion and posture data are highly valued, since it is a sensitive and time-consuming process, but highly necessary for the risk analysis and the deployment of semi-supervised and supervised methods [2]. In this case, motion data in real scenarios are even more complex since it is highly rich and diverse in behaviors. For instance, although cyclic and repetitive activities performed in industrial scenarios are mostly consistent from cycle to cycle, perfect conditions cannot be met all the time. Eventually, the working process can inadvertently be stopped or delayed. In addition, ergonomic risk assessment methods have to analyze each working cycle, which means that a previous segmentation of each one of these instances has to be made. Sub-activities that make the sequence of actions comprehended in the working cycle can also be sub-segmented for further recognition and association with risk measures. For instance, the *ergonomist* might have an interest in understanding which actions from the working cycle have a significant association with a high or low risk measure, to adapt the working station with intervention strategies that could help prevent future disorders. Another relevant analysis is the study of changes in the working behavior over time, since by segmenting the working cycle and its sub-activities, we can compare them over time and perform higher level associations (e.g. how the worker adjusted his behavior over the morning and afternoon periods).

This work studies the hypothesis of using the *Self-Similarity Matrix* (SSM), already applied for audio information retrieval [4], for segmentation and assisted labeling of time series from inertial motion and postural data of workers in an assembly line of an automotive industry. Several results are provided showing the ability of using this method for segmentation tasks, and examples with exploratory results are given to extend the segmentation power of the method for multi time-scale segmentation, summarization and assisted labeling. The contributions are towards having an unsupervised tool to perform time series segmentation in multiple and hierarchical time scales that can further be applied for time series summarization and/or personalized labelling-by-example in occupational health datasets.

To provide an illustrative example to the reader of our intent, we show a multi-scale segmentation of an electrocardiogram (ECG) signal with the proposed method in Figure 1.

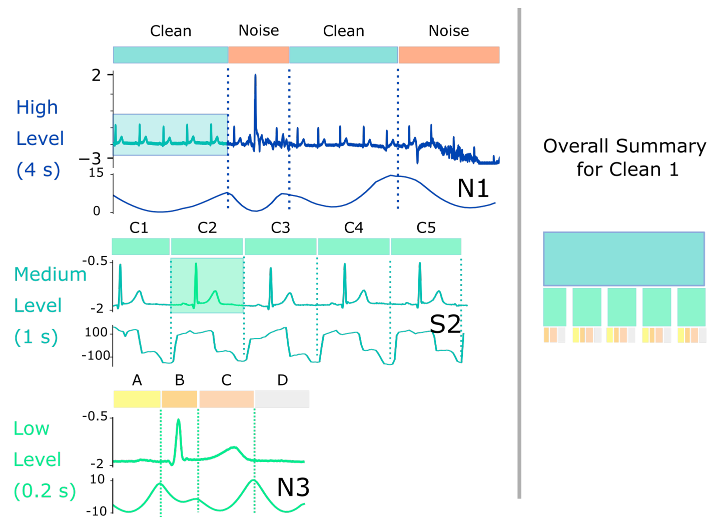

Fig. 1. Example of a multi time scale segmentation of a simple ECG signal with noise. The signal is from the Physionet[5]. Each row of signals indicates a smaller time-scale, zooming-in a specific sub-segment of the previous segmentation. On the right, an example of how we could summarize the first portion of the signal with this type of segmentation.

Figure 1 shows three levels of segmentation from higher to lower time scale. The first is segmented by the novelty function (N1) with a time scale of *4 seconds*, which peaks separate clean ECG from noise. The segments of interest can then be segmented into ECG cycles by the valleys of the similarity function (S2), calculated with a time scale of *1 second*. Finally, inside each cycle, the most significant sub-segments are each of the P, T waves and QRS complex, segmented by the novelty function (N3) with a time scale of *0.2 seconds*. This example is illustrative of the goal we want to reach with this work towards a tool that can perform a meaningful multi time-scale segmentation that contributes to transform the data into a summary and that can be used to assist the analyst in performing labelling of the segments identified.

1.3 Towards Occupational Data Segmentation

This work proposes the usage of the *SSM* created by a feature-based representation of the time series from the worker's motion while performing tasks. From this representation, we will present visual evidence of its usage for the identification of events with ergonomic significance and summarize the working activity. The summarization process is performed by segmenting the data into sub-sequences of interest in a multi-timescale approach, as defined below.

1. **Active working periods** *(higher time scale)*: *Active and non-active segments* indicate sub-sequences of the time series where the worker was performing the cyclic working tasks or not doing them, respectively (e.g.: stop on the working line that leads to a pause in the working activity). In this work, transitions between active and non-active work periods are made, with the intent in focusing (*zoom-in*) our attention to each *working cycle*;
2. **Working cycles** *(middle time scale)*: Sub-sequences of highly similar sequences of movements that are being repeated in time. This type of segmentation pre-assumes that the time series under analysis will be mostly (or entirely) defined by the repetition of a motion, which are the examples of tasks analyzed in this work. Having each working-cycle identified, *sub-activities* inside the working cycle can be searched by focusing (*zoom-in*) our attention to a single working-cycle sub-sequence;
3. **In-cycle activities** (lower time scale): A working cycle is structured in a sequence of *sub-activities*. The transition between these might be related with a significant change in the worker's motion behavior or change in posture. In this context, what will be intended is to divide the working cycle into more primitive segments and understand if the segmentation is consistent over each of the analyzed cycles.

In addition to the several levels of segmentation, we will also provide an example in how a labelling process can be performed with the same *SSM* by means of a *query-by-example*. This methodology will be exemplified to identify instant repetitions of a given pre-specified time series' sub-sequence of the user's interest. For example, if the analyst wants to understand the contribution of a specific segment of the working cycle for the occupational risk, this sub-sequence could be used as a query to search all the other equal segments. Moreover, current risk evaluations of occupational exposure are made with the assistance of video. Such query-by-example methods could be integrated into a video system that let's the analyst select an interval of the video that corresponds to a specific part of a working cycle, and the algorithm segments the remaining other similar segments. This would increase the speed of analysis and/or labeling process.

The segmentation process proposed will rely in the detection of significant events in multiple time-scales. These events are the instants in time that segment any of the aforementioned sub-sequences (1, 2 and 3). This work provides additional content to previous experiments [3] by (a) improving the detection performance of *events 1* with the usage of the *novelty function*, (b) extend the event detection process to a multi time-scale approach by segmenting sub-activities of interest inside a working cycle, and (c) discussing how this method can be used as a tool to support labelling of time series.

This document is organized by first (1) discussing related work, then (2) explaining the data we used, (3) describe the proposed methodology, (4) and (5) demonstrate illustrative examples in occupational inertial data, while discussing these findings and (6) provide final thoughts regarding this work.

2 Related Work

Processes of computer analysis under acquisitions retrieved from a real work environment have already been proven to be very useful in the design of evaluation tools to facilitate ergonomic studies. Under the work of [13], by firstly calculating the orientation of anatomical joints, it was possible to design an adjusted ergonomic risk score. This had the final objective of automatically assessing the operator risk exposure during the work on an assembly line.

In the context of summarizing a work period acquisition by segmenting it into smaller motion segments, the work of [14,15] has analysed multidimensional angular joint motion time series with each motion segment being represented by dynamic model approximations. The segmentation methodology was based on univariate active forgetting segmentation methods, with a two-step recursive least square algorithm predicting change points in the dynamic behavior of the system. The motion data segments are then represented by parameters derivative from a dynamic model fit, which as the advantage of the features being insensitive to small-time variations, very common in this type of data, and allows for a comparison based on a "kinetic energy-like" measure.

Meanwhile [16], in the process of summarizing human motion, temporally segmented the repetitive human motions, much like we also propose to do. This work used a time series representation of the joint angle of the subjects. This motion data was then converted into a generic full-body kinematic model, by using an unscented Kalman filter, and then retrieved kinematic features by performing a primary frequency analysis to the transformed data. The data was then segmented by based on the zero crossing of these retrieved features, followed by an adaptive k-means clustering to identify which segments are repetitions of each other.

[17] proposed the use of a mask-based Neural Network (NN) capable of extracting desired patterns of interest from a large time series database, without the requirement of a predefined template. Validated under electrocardiogram and human motion signals, this work was an algorithm proposal that could automatically detect specific patterns in biosignals.

The use of an SSM applied on the analysis of human motion, more specifically human motion at work, is still a subject with a short exploration in the literature. However, this work will make the case that it might be a useful tool adapted to this concern.

Notwithstanding, there has been some works on this subject, with [18] discussing in much detail a method that relies on a neighborhood graph to partition the dataset into 1) distinct activities and 2) motion primitives according to a self-similar structures. Alternatively, we can also see the matrix profile algorithm[19], an well established algorithm for an optimized processing of time series, which by focusing on the similarity join problem can additionally compute the answer to the time series motif and discord problem. A matrix profile is in a short summery a "time serie", whose values are the euclidean distance between a subsequence and its nearest neighbor. Under this work, its also proven the validity of this methodology to analyse human motion data.

The use of SSM is a well explored tool for the a rapid analysis of music data. The reason why this type of database is especially fit for this analysis is that it is (1) recurrent in nature, (2) complex enough in nature than other types of processing of the raw data would be demanding enough that converting to a similarity matrix isn't as demanding in comparison. Fortunately, these characteristics also apply to our case. We are using inertial data in time series format, retrieved from various Inertial sensors attached to the subjects' body, which is a complex type of data. Besides, it is recurrent in nature, because the tasks being analysed are cyclic. Therefore, the proposed method is highly inspired by the same method used in audio signals analysis for audio thumbnailing or summarization [20].

The developm

3 Dataset Description

3.1 Population

The inertial data available to develop this methodology was acquired in the context of validating an inertial measuring system that would guarantee access to direct measures in occupational industrial environments. This system was previously used to deliver an ergonomic risk assessment based on the angular information retrieved by the raw data of these sensors [1].

The population in which the system was tested is described in this work [1]. The in-field data used was acquired in an industrial environment from an automotive assembly plant. More specifically, the subjects were working in assembly tasks and the data was acquired while the subjects were performing the tasks of a specific workstation. The dataset includes six participants, each monitored while working at two different workstations. In this scenario, each workstation has a specific set of tasks that have to be performed by the worker. These tasks are repeated throughout the working period, being divided into working cycles.

3.2 Instruments and Setup

An inertial measurement unit (IMU) is an electronic device composed internally by three 3-axial sensors: accelerometer (Acc), gyroscope (Gyro) and magnetometer (Mag). When attached to strategic points of the human body, it allows to register that object's specific force, angular momentum, and position, simultaneously. Combined, these provide much information about the movement and posture of the subject.

In this study data from the dominant upper limb of the subjects was measured. The system comprehends a set of four 9-DoF IMUs. These were attached on the upper dominant limb of the subjects, namely:

- **IMU 1** posterior side of the hand
- **IMU 2** posterior side of the forearm (wrist)
- **IMU 3** posterior side of the arm (elbow)
- **IMU** 4 thorax area

All devices were attached so that the Y-axis was aligned upwards. Figure 2 shows how the sensors were placed on the subjects.

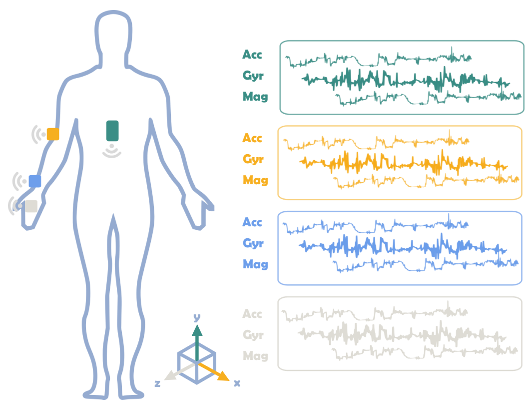

Fig. 2. Schematic of the placement of Inertial sensors, used for the dataset acquisition protocol. Based on [1] and [3]

.

The signals available for analysis are the 3-axis accelerometer, gyroscope and magnetometer of all IMUs used, collected with a sampling rate of 100 Hz. The raw data has all the events described in the **??**, namely (1) active working periods intercalated with non-active working periods; (2) working cycles and (3) sub-activity segments. For the ground-truth position of these events, all signals were annotated by means of video-records of the acquisition sessions.

4 Methods

The following section describes the proposed methodologies to detect the various events of interest for the segmentation process. It will, as such, be structured as follows:

1. SSM construction

2. Similarity function
3. Novelty function
4. Sub-sequence search-by-example

, with the first point describing how the *SSM* is structured such that all the remaining strategies are capable of finding relevant event types presented in Section 1.

4.1 SSM construction

Pre-Processing Before any analysis, the data has to be prepared. This includes tasks such as synchronizing and filtering. The synchronization tends to be an essential process, as the framework usually involves several sensor devices that can have divergent internal clocks. The filtering process is also a fundamental step. In this case, a second-order low pass Butterworth filter of 40 Hz was used on the inertial data. The choice of this cut frequency was because experimental studies [21] have already proved that human motion and posture could be well represented by a frequency up to 20Hz.

Feature Retrieval After all the sensor dimensions have been pre-processed, a new representation of the dataset is made by applying a moving window function, which retrieves a set of predefined features of the temporal, statistical and spectral domain. Extracting relevant features is of great importance to have a rich characterization of the morphology of each signal [22].

This process has the result of turning a multivariate time series with n dimensions and m data points, into a multivariate time series with $n \times f$ feature dimensions and $< m$ number of data points. This is a process of feature retrieval which doesn't necessarily reduces the volume of data, but instead tries to reduce its complexity into simpler components that describe the shape of the data.

This entire process will, of course, be parameterized by the windows length ($Wind_{len}$) and an overlapping fraction ($Overlap_{frac}$). Both $Wind_{len}, Overlap_{frac}$ have a large influence on the results, as they define the time scale at which features are extracted and consequently will also define the time resolution of the *SSM*. This means that an adjustment of these parameters changes the time dimension of the events which are gonna be highlighted. In other words, a larger $Wind_{len}$ will result in highlighting similarities between longer sub-sequences of the set of time series, while a shorter $Wind_{len}$ will have the inverse effect. Most studies in the field of human activity recognition consider the use of 1 to 10 second time length windows.[23] However, despite this being a good starting value in an ergonomic analysis, it is required to consider events along the multiple time scales, and without any previous information, there cannot be made any assumptions.

SSM calculation An *SSM* is a graphical representation of the similarity between each data sample (or window) and all the remaining samples (or windows) in the rest of a time series. This allows to highlight similar and (dis)similar structures, that can be patterns or dynamic behaviors of the signal.

Obtained the previously described feature matrix

$$F = (x_1, \ldots, x_n)_{(n \in \mathbb{N})}$$

, where each coordinate x is also described by a vector of dimensions ($m \in \mathbb{N}$). Then the *SSM*, used for this work, will have a size $n \times n$, with each of its position being defined by

$$SSM(i, j) = s(x_i, x_j)$$

Where $s(x_i, x_j)$ is a distance measure, which takes into account the vector points $x_i, x_j \in X$, both of length m and returns a real value, which represents a score of how closely similar are these two point coordinates.

As such, to build an *SSM*, it is necessary to define the distance function s, with the literature on the subject pointing usually to three options: cosine of the angles, euclidean distance and Kullback Leibler distance. This work chose to use the first distance function. To calculate it, the feature representation signal F positions were normalized. Then, the distance function could be described by the simple inner product between two vectors

$$s(x_i, x_j) = \langle x_i, x_j \rangle$$

, which means that the *SSM* can be calculated by the dot product of F and its transposed [25].

$$SSM = X^T X$$

4.2 Similarity function

A similarity function S_f is a univariate time serie where each value represents how different is a specific time instant when compared with the remaining time series.

Given the *SSM*, built as described in the previous section, the similarity functionS_f can be calculated by the sum throughout one of the axes of the *SSM*.

$$S_f(t) = \sum_{i=0}^{n} SSM_{ti}$$

, with n being the length of the SSM. Due to the property of symmetry of the *SSM*, this operation can either be applied over the lines or the columns of the matrix, with the final result being the same S_f. This will then be smoothed to facilitate the following processing stages.

In a previous work, it was used to search for highly dissimilar sub-sequences of the signal [3]. In this context, this method was mostly successful in identifying sub-sequences corresponding to periods of non active work. Therefore, by using a simple minimal threshold, it was possible to identify these sub-sequences and remove them from the signal. In addition, the similarity function is useful for the segmentation of periodic events, as is exemplified in Figure 1. As the dataset being observed is only composed by working cycles, periodically repeated along time, the S_f will approximate to a sinusoidal function. Then, the valleys of S_f will indicate the time instances where there is a transition between the working cycles, which also can be identified by a simple valley detection operation.

4.3 Novelty function

The Novelty function N_f is also a univariate time serie calculated from the processing of the *SSM*. However, N_f intends to instead provide information on how much the data within a neighborhood distance to the left of that point is different from the data within a neighborhood distance to its right. Being especially relevant in the detection of change points.

When the signal displays a significant change between two states of motion, the corresponding *SSM* will display two distinct blocks along the main diagonal. As such, the novelty function N_f can be calculated by making a *checkerboard* kernel convolution centered along the main diagonal. This kernel in its simplest form can be described as the sum between a kernel that measures coherence and anti-coherence on either side of the center point. In other words, the first component presented in the following equation highlights when the two regions are homogeneous or coherent within each other. Meanwhile, the second component will be highlighted whenever these two regions are also similar within each other. As this last component is negative, the opposite will be expected [25].

$$K_{Box} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$$

The resulting novelty function N_f is expected to have higher values of intensity whenever it is in the middle of two different blocks [25]. The *checkboard* kernels' dimension will be defined by ($M \times M$, where $M = 2L+1$, for $L \in \mathbb{N}$. The central column and central row coordinates of the kernel K_{Box} will have the values of 0, followed by four planes which will either be 1 or -1 according to the same pattern as previously seen. For example, if L=2, then [25]

$$K_{Box} = \begin{bmatrix} -1 & -1 & 0 & 1 & 1 \\ -1 & -1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & -1 & -1 \\ 1 & 1 & 0 & -1 & -1 \end{bmatrix}$$

The final kernel is then the result of further smoothing by a radially symmetric Gaussian function ($\phi(s,t)$) and normalization (dividing by the sum of all absolute values of the kernel) [25].

$$K_{checkerboard}(k,l) = \frac{\phi(s,t) \cdot K_{Gauss}(k,l)}{\sum_{k,l\in[-L,L]} \mid K_{Gauss}(k,l) \mid}$$

The detection of change points is one of the most commonly used techniques in event detection, as by definition they search time instances that display a significant change in the properties of the signal. This work intends to take advantage of this algorithm to also structure the time series into different non overlapping states. In occupational context, we'll use this to separate between *active* and *non-active periods*, and also to divide the work cycles into a sequence of *sub-activities*.

4.4 Sub-sequences searched-by-example

Another aspect of occupational risk evaluation in industrial scenarios is to compare the occupational risk of sub-segments of the working cycle during the working period. This strategy supports professionals in identifying specific sequences of sub-activities that occur in-cycle, and search them over the entire set of working cycles for comparative purposes. In addition, if in need of labeling the data, this method can be quite fruitful as it searches for the exact sub-sequence match in the matrix. In addition, this process does not match the shape but rather the sub-diagonal of the matrix, that is, being the sub-diagonal one cycle, we are matching the exact portion of the cycle being used as an example.

The search procedure is made with an example that is a sub-sequence of interest in the signal. The search procedure works by sliding the selected example along the *SSM*. The distance, D, between the example and the segments it slides over is calculated as the sum of absolute differences:

$$D(x) = \sum_{x=0}^{x=M} \sqrt{(SSM(x) - SSM_t)^2} \tag{1}$$

where $SSM(x)$ is the segment of the *SSM* over which the example, SSM_t, slides at moment x, starting from 0 to the size of the *SSM*, M. The resulting function has valleys at the instants where the example matches.

4.5 Illustrative Example

Figure 3 shows an *SSM* representation of the set of inertial signals acquired while a worker was performing 2 different workstations (A and C). The *SSM* also shows an interruption in the working line (B). With this illustration, we can highlight which structures are mainly present in the *SSM*, namely blocks and parallel sub-diagonals. The first indicates transitions between homogeneous blocks and can be detected by the novelty function. The second indicates presence of periodicity or cyclic behavior, which can be detected with the similarity function. In the next section, these functions will be used in examples of signals from the manufacturing scenario to provide evidence of applying this methodology to perform the segmentation of time series and possible applicability in summarization and labeling.

The methods explained in this section are deeply inspired by the work from Meinard Müller in the context of information retrieval from audio records [4][24].

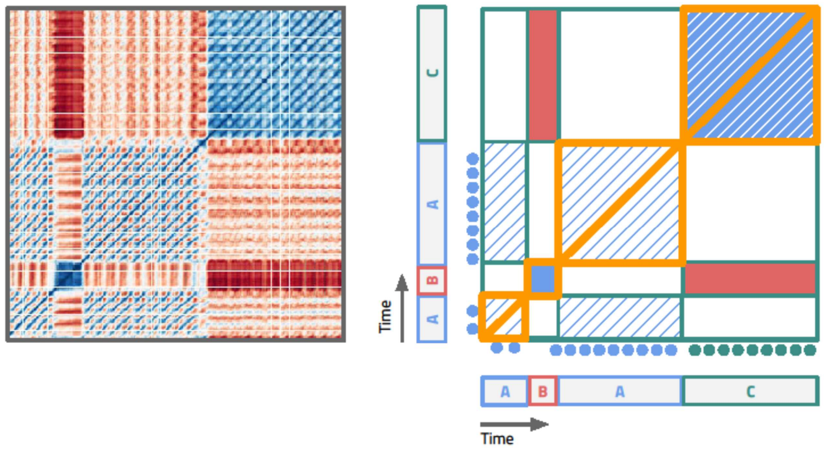

Fig. 3. At the left is the SSM designed from the signals acquired while an operator was performing 2 different workstations. At the right is a simplification of the original SSM, with highlights on the main structures present (Blocks and Diagonals). A being workstation 1 and C workstation2. B is the interruption in the working line.[3]

We performed the *SSM* calculation and novelty searched by means of the available *libfmp* python library [20] [25] [26].

5 Results and Discussion

The following section delivers the results obtained by applying the aforementioned methodologies to demonstrate the validity of the proposed objectives. As presented as an introductory simple example in Figure 1, we can use the novelty function to segment homogeneous sub-sequences of the time series, and the similarity function to identify periodic segments. With this in mind, we will now demonstrate the application to more complex scenarios, such as motion data in industrial settings, for time series segmentation, summarization and labeling.

The proposed method is applied in real motion data from an industrial setting to perform the (1) detection of active working periods and (2) segmentation of working cycles. In addition, it provides examples to use these methods to perform a multi timescale analysis and summarization of the data, as well as how this can support a labeling process.

5.1 Active Working Periods Segmentation

In order to detect the transition between *active* and *non-active work* we used the novelty function N_f. This function will display peaks in sections where there is a transition between coherent blocks. To identify these transitions we applied a smoothing and peak detection technique to the N_f.

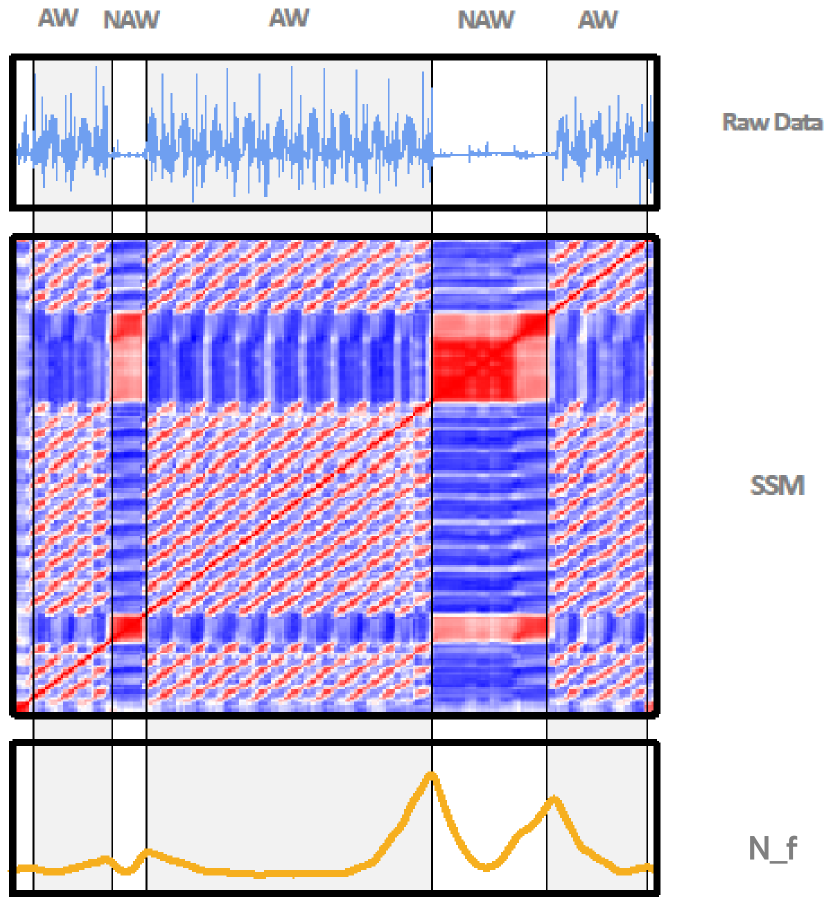

Fig. 4. Schematic of active working period detection, by means of a novelty function. Each image represents a peak detection process. From the top to the bottom, the first step represents the x coordinate of the hand accelerometer during the (Opr4 Wkst1) acquisition. The following images represent the respective SSM and N_f, as labeled. The *active* and *non active* work periods are accordingly represented with the acronyms of AW and NAW at the top of the image [3]

Considering the example provided by Figure 5.1. This signal can visibly highlight the time instances where the *SSM* transitions from high valued, squared, and homogeneous blocks, which represent periods of *non-active work*, to blocks with several parallel sub-diagonals, equally spaced, which represent the periods of *active work*.

The measurements of the algorithm's performance are summarized in table 1. This analysis was made to understand how close were the events identified by the algorithm to the manually annotated labels. For this, we used Precision, Recall and the F_n-score (with the TP annotations having a tolerance of 50s), and how much time distance existed between the identified event and those labelled annotations (Mean absolute error measured in seconds). From table 1, it is possible to see that 4 of the 7 samples had the perfect score measurements, detecting every single intended event without any false classification. From the remaining 3 samples, the false classifications were all False Positives, which, in this context, are more affordable when compared to False Negative classifications. There is given a greater significance on the measurement of R against P, because a presence of FP could be noticed in a further analysis of the results, as, after all, the motivation of this work is to serve as a support for the analysis of ergonomic data.

Moreover, the identified FP events just tended to divide the *non active work* states into various sub-states. Something understandable as there might be a more complex motion description within these time instances, which wasn't considered for the context of this work. However, this is a manageable error, as long as there is still a clear segmentation between the *active work* time periods from the remaining time series, the segmentation of the *non-active work* is irrelevant for the context of this problem.

The increase in the MAE value is associated with the several smoothing processes applied over the methodology, that distanced the events from their ground annotations. However, when considering that the *active work* period tended to be about 1169.11 s and even the smaller states tended to be of 128.52s. These MAE values (within 8.38 - 11.87 s), although significant, correspond to a small percentage of the time.

Table 1. Results of type event 1 (work period transition), discriminated per time serie samples. Measurements of Precision (P), Recall (R), F_n-score(F) and mean absolute error(MAE), of each according sample.

TS sample	P	R	F	MAE (s)
Opr 1 Wkst1	0,78	1	0,88	11,87
Opr1 Wkst2	1	1	1	34,77
Opr2 Wkst1&2	0,86	1	0,92	10,17
Opr3 Wkst1	0,80	1	0,89	3,21
Opr4 Wkst1	1	1	1	8,83
Opr5 Wkst1	1	1	1	8,54
Opr5 Wkst2	1	1	1	8,38

Overall, this process was successful without the necessity of a very intense search, and by performing a simple manual selection of the parameters, most events were able to be detected.

5.2 Working Cycles Segmentation

Much like the work on [3] the best methodology for the segmentation between cycles is the retrieval of the S_f followed by a smoothing and valley detection operation of this signal.

Considering the example provided in figure 5.2, we can see that the similarity function will consistently segment the dataset into positions very close to the annotated ones. Despite usually having a minor delay, the events tend to maintain that delay constantly across the entire time series. The reasoning for this is because the algorithm is unsupervised, it does not have a reference of where the cycle has the "real" start. As such the algorithm will take into consideration the beginning of the data as a reference, which might not be the same precise instant when the operator starts the work cycle. This does not make the detection necessarily incorrect, as long as the following cycles are also segmented in the same transition point, it can be considered a work cycle motif. This consideration means that the assessment will insist more on the question of the consistency of the detections made rather than on how close they are to the labeled positions.

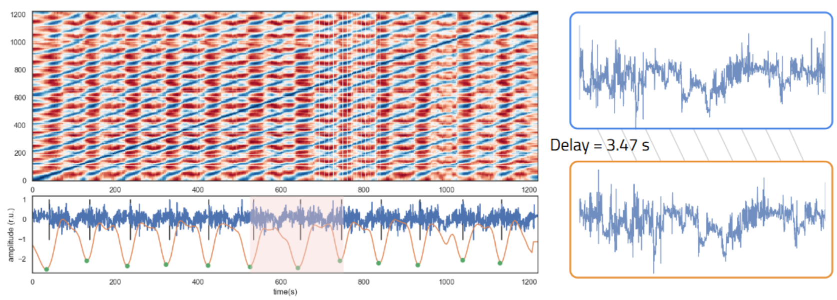

Fig. 5. Schematic of active working period detection, by means of a novelty function. Each image represents a peak detection process. From the top to the bottom, the first step represents the x coordinate of the hand accelerometer during the (Opr4 Wkst1) acquisition. The following images represent the respective SSM and N_f, as labeled. The *active* and *non active* work periods are accordingly represented with the acronyms of AW and NAW at the top of the image [3]

As such, the following metrics will insist on two points: 1) Number of cycles detected 2) consistency of duration of the detected cycles. The last point is described by the mean absolute error between the duration of the ground truth segmentation cycles and the algorithms' segmentation cycles, described by DE (Duration). In the work of 2 the results of the proposed algorithm were also compared with a more well established technique proven to work in similar problems, the Matrix Profile.

Overall, the measures considered for this analysis demonstrate the ability to identify working cycles with good accuracy. Almost all the work cycles of interest were detected, as, within the 157 ground annotation cycles, 154 cycles

Table 2. Detected cycles and DE results of the detection of working period events detection, under an occupational context, with the results being separated according with two different methodologies with the SSM being the technique based on the analysis of the similarity function S_f, presented in this work, while the other is the Matrix Profile algorithm

Signal	*SSM*		*Matrix Profile*	
	Detected Cycles	Duration Error	Detected Cycles	Duration Error
Opr1 Wkst1	11/11	3.26s (3.04%)	11/11	11.08s (10.34%)
Opr1 Wkst2	14/15	16.97s (15.83%)	14/15	8.09s (7.55%)
Opr2 Wkst1	14/14	6.45s (6.40%)	14/14	6.74s (6.70%)
Opr2 Wkst2	11/11	8.48s (8.62%)	11/11	11.2s (11.39%)
Opr3 Wkst1	16/16	12.35s (11.79%)	16/16	7.39s (7.05%)
Opr3 Wkst2	13/13	8.81s (8.25%)	12/13	11.41s (10.68%)
Opr4 Wkst1	14/14	1.05s (0.4%)	14/14	8.72s (8.24%)
Opr4 Wkst2	11/11	3.42s (3.32%)	10/11	4.9s (4.75%)
Opr5 Wkst1	12/12	2.83s (2.85%)	11/12	5.39s (5.43%)
Opr5 Wkst2	10/11	3.47s (3.45%)	10/11	6.7s (6.69%)
Opr6 Wkst1	14/15	3.79s (3.74%)	15/15	7.25s(7.15%)
Opr6 Wkst2	14/15	5.79s (5.73%)	15/15	6.13s (6.06%)
Total	154/157	6.12%	153/157	7.6%

were detected. The duration error was mostly good with an average value of 6.12% of the working cycle, but still significantly high in some cases (2 and 5). When compared with the *MP*, the results are comparable. To be clear, we are not trying to say if our algorithm is better or worse than the *MP*, but simply to have a standard measure of reference to compare with.

5.3 Towards multi time-scale Segmentation

In the two previous sections, we highlighted the ability to perform the segmentation of time series based on novelty, while also being able to segment subsequences of cyclic nature. Now, we intended to extend the usage of the *SSM* to show how it could be used in an iterative segmentation over multiple timescales, successively over smaller previously segmented sub-sequences. Considering the previously detection of AW as the *high level*, and the segmented working cycles as the *middle level*, this next section devotes its attention to analyse in-cycle detection.

Considering the scenario of occupational health, we are dealing with a multidimensional dataset with some variability in the worker's motion, which means that each working cycle might be relatively different. Still, the sequence of in-cycle activities are the same, with small variations. Using the novelty function with a smaller time scale, we segmented several working cycles of one subject to present as an exploratory example of a lower-scale segmentation. The *SSM* was calculated for each previously segmented working cycle, with a time scale of 2.5 seconds. In Figure 6, we present the corresponding novelty functions for

3 working cycles (*C1, C2 and C3*), highlighting their peaks as the segmentation instants.

Firstly, the image shows us that the inertial signals are mostly consistent over working cycles. There are small variations, but the significant changes occur in the same sequence. As the novelty function highlights instants where a change is significant, we expected its peaks to be related to a significant change in the posture/motion of the worker.

In a first inspection, we can see that the peaks of the novelty functions for all cycles match a significant change in the group of inertial signals from which the *SSM* was built. Matching the peaks with a video inspection, we were able to associate the changes with what happened in the video. The caption of Figure 6 indicates the list of activities. In general, these changes are related to a transition between homogeneous blocks of posture/motion, that is, sub-activities that have a certain pattern and are shifted to another sub-activity by a change of motion/posture. For instance, block C indicates a quick motion to a new position to perform a set of tasks with *tool1*. There are also obvious symmetric behaviors that are segmented, such as *I, J and K*, where the subject positions *piece1 (I)*, works on a static posture *(J)* and unfits *piece1*. These actions are well separated in all exemplified cycles.

These are still very preliminary results in a first exploratory experiment on motion data, but we believe that this method, already used in audio thumbnailling (*technically summarization*), is worth exploring in other types of data for the problematic of time series segmentation and summarization. In addition, we also show that this method is worth exploring in a multi time-scale segmentation process, with simple ECG data (from Figure 1), but also in more complex data, such as motion and posture.

The importance of summarizing the data is in the way we can then represent the signal in a higher levelled representation. As an example, Figure 6.*right* shows the length of each sub-activity segmented for each cycle. This type of visual summary gives the analyst a quicker form of getting feedback from the data. Besides, if integrated into an interactive platform, it can provide valuable interactive power.

5.4 Labeling with query-by-example

Finally, the last scenario focuses on demonstrating the possibility of performing labeling by querying an example along the *SSM*. Figure 7 exemplifies a case where we are trying to label 3 different segments from a working cycle. The segments are highlighted as *A, B and C*. Applying the distance measure across the signal results in the signals $D_A, D_B and D_C$. These distances show how well the valleys (minimal distance) indicate the match with the example. Performing the match along the sub-diagonals of the *SSM*, it is not *distracted* by similar shapes, but rather by the exact moment the pattern occurs in the working cycle. This method can be used assigned with a label, to semi-automatically label segments of a working cycle, such as *A, B and C*.

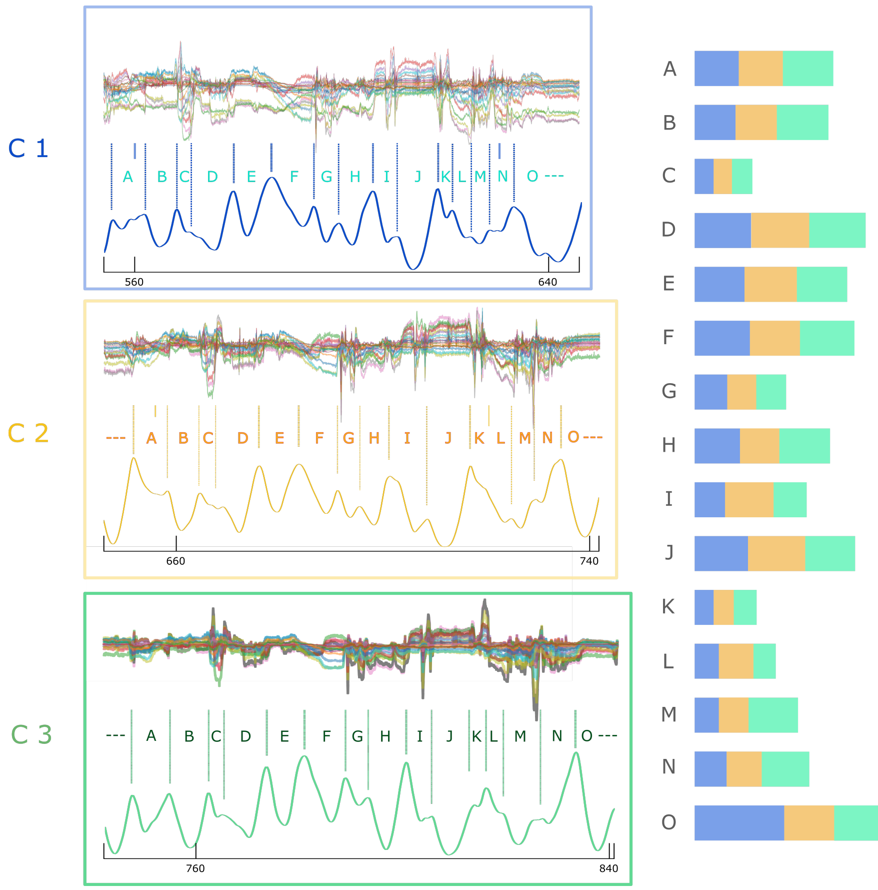

Fig. 6. Example of an in-cycle segmentation of sub-activities of a worker. There are 3 cycles (*C1, C2 and C3*. For each, the corresponding set of inertial signals and the resulting novelty function are presented. The list of activities are labelled from A to O, as follows: A - stretching arms to the top and slowly pulling an object; B - adjusting the object to be fixed on the car; C - picking tool1 and using it on the side of the car; D - moving to the front of the car and use tool1; E - move back to the side and leave tool1; F - perform inverse sub-activity of A; G - walk away from the car; H - walk back to the front of the car; I - Fit piece1 on the car; J - use tool2; K - unfit piece1 from the car; L - walks away from the car; M - drops tools and pieces; N - moves back to the initial position; O - waits for the next cycle.

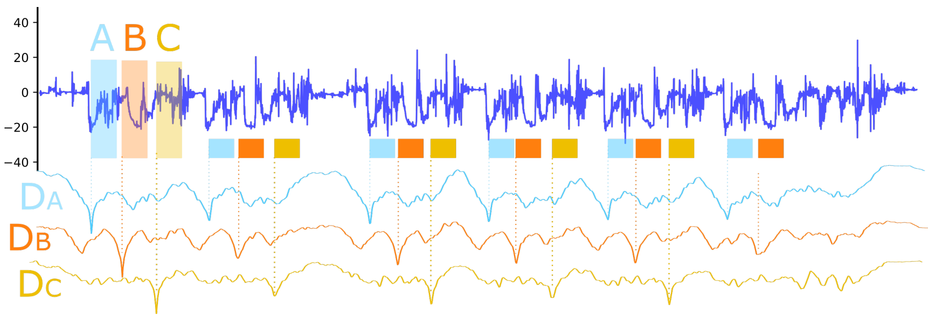

Fig. 7. Query-by-example on the *SSM*. Segments A, B and C are used as examples by being selected on the *SSM* columns. The resulting distance functions ($D_A, D_B and D_C$) show the ability of this method to find the starting point. The signal is only illustrative, as all the signals available were used to build the *SSM*.

6 Conclusions

One of the objectives of this work was to demonstrate the relevance of retrieving events through the usage of an unsupervised method, for posterior ergonomic analysis. In this sense, the methods were able to structure the time series under different levels of abstraction. Firstly, by detecting the sub-sequences of actual active work movement, then by segmenting the various work cycles, and in the end, by summarizing the work cycle into smaller primitive structures, different from each other. Then it also provided a tool to further detect new instances of an user-defined query.

This methodology shows promise in doing so, as when the events of "Active Working Periods Segmentation" and "Working Cycles Segmentation" were compared with manually labelled annotations they were proven to be mostly all detected. Moreover, when calculating the error duration between the events detected and the annotation these were noticed to be rather short and of little variation. On the other hand, the remaining events were proven by comparison to provide useful and valid information about the work period motion. The most significant utility provided by this algorithm is the automation of processes related to the identification of periods of relevance, segmentation of the signal, and further identification of periodic regions. Moreover, the proposed analysis by means of the *SSM* seems to be a promising approach with the potential to be expanded upon.

The empirical results reported herein should be considered in light of some limitations. The process of feature retrieval is especially time demanding, and the construction of entire *SSMs* for each time series might become too demanding for the storage capabilities. These were concerns that were not noticeable for the dimension of the provided data, which described a short work period extension, but it might become too demanding for acquisitions of entire days. These are points should be further investigated in future research works.

Acknowledgements

PREVOCUPAI (DSAIPA/AI/0105/2019), and Ph.D. grant PD/BDE/142816/2018.

References

1. Santos S., Folgado D., Rodrigues J., Mollaei N., Fujão C., Gamboa H. (2021) Exploring Inertial Sensor Fusion Methods for Direct Ergonomic Assessments. In: Ye X. et al. (eds) Biomedical Engineering Systems and Technologies. BIOSTEC 2020. Communications in Computer and Information Science, vol 1400. Springer, Cham. https://doi.org/10.1007/978-3-030-72379-8_14
2. Y. Roh, G. Heo and S. E. Whang, "A Survey on Data Collection for Machine Learning: A Big Data - AI Integration Perspective," in IEEE Transactions on Knowledge and Data Engineering, vol. 33, no. 4, pp. 1328-1347, 1 April 2021, doi: 10.1109/TKDE.2019.2946162.
3. Santos, A.; Rodrigues, J.; Folgado, D.; Santos, S.; Fujão, C. and Gamboa, H. (2021). Self-Similarity Matrix of Morphological Features for Motion Data Analysis in Manufacturing Scenarios. In Proceedings of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies - BIOSIGNALS, ISBN 978-989-758-490-9; ISSN 2184-4305, pages 80-90. DOI: 10.5220/0010252800800090
4. Paulus, Jouni & Müller, Meinard Klapuri, Anssi. (2010). Audio-based music structure analysis. Proceedings of the 11th International Society for Music Information Retrieval Conference, ISMIR 2010. 625-636.
5. Heldt T, Oefinger MB, Hoshiyama M, Mark RG. Circulatory response to passive and active changes in posture. Comput Cardiol, 30:263–266, Sept. 2003.
6. J. Song, R. W. Chang, and D. D. Dunlop. “Population impact of arthritis on disability in older adults.” In: Arthritis Care and Research 55.2 (2006), pp. 248–255. issn: 21514658. doi: 10.1002/art.21842
7. C. Palazzo, J.-F. Ravaud, A. Papelard, P. Ravaud, and S. Poiraudeau. “The Burden of Musculoskeletal Conditions.” In: PLoS ONE 9.3 (2014). Ed. by A. Chopra, e90633. issn: 1932-6203. doi: 10.1371/journal.pone.0090633.
8. C. H. Dominick, F. M. Blyth, and M. K. Nicholas. “Unpacking the burden: Understanding the relationships between chronic pain and comorbidity in the general population.” In: Pain 153.2 (2012), pp. 293–304. issn: 03043959. doi: 10.1016/j.pain.2011.09.018.
9. M. D. Antonopoulou, A. K. Alegakis, A. G. Hadjipavlou, and C. D. Lionis. “Studying the association between musculoskeletal disorders, quality of life and mental health. A primary care pilot study in rural Crete, Greece.” In: BMC Musculoskeletal Disorders 10.1 (2009), pp. 1–8. issn: 14712474. doi: 10.1186/1471-2474-10-143
10. U. T. Kadam, K. Jordan, and P. R. Croft. Clinical comorbidity in patients with osteoarthritis: A case-control study of general practice consumers in England and Wales. 2004. doi: 10.1136/ard.2003.007526.
11. A. Luttmann, M. Jager, B. Griefahn, G. Caffier, F. Liebers, and W. H. Organization. “Preventing musculoskeletal disorders in the workplace.” In: (2003). issn:924159053X.
12. A Pruss-Ustun and C Corvalan. “Preventing disease through healthy environment: A global assessment of the burden of disease from environmental risks.” In:World Health Organization (2006), p. 105

13. Santos, S., Folgado, D., Rodrigues, J., Mollaei, N., Fujao, C., Gamboa, H. (2020, February). Explaining the Ergonomic Assessment of Human Movement in Industrial Contexts. In BIOSIGNALS (pp. 79-88).
14. Lu, ChunMei, and Nicola J. Ferrier. "Repetitive motion analysis: Segmentation and event classification." IEEE transactions on pattern analysis and machine intelligence 26.2 (2004): 258-263.
15. Lu, ChunMei, and Nicola J. Ferrier. "Automated analysis of repetitive joint motion." IEEE transactions on information technology in biomedicine 7.4 (2003): 263-273.
16. Wang, Qifei, et al. "Unsupervised temporal segmentation of repetitive human actions based on kinematic modeling and frequency analysis." 2015 international conference on 3D vision. IEEE, 2015.
17. Matias P, Folgado D, Gamboa H, Carreiro A. Time Series Segmentation Using Neural Networks with Cross-Domain Transfer Learning. Electronics. 2021; 10(15):1805. https://doi.org/10.3390/electronics10151805
18. Vögele, A., Krüger, B., & Klein, R. (2014, July). Efficient Unsupervised Temporal Segmentation of Human Motion. In Symposium on Computer Animation (pp. 167-176).
19. Yeh, Chin-Chia Michael, et al. "Time series joins, motifs, discords and shapelets: a unifying view that exploits the matrix profile." Data Mining and Knowledge Discovery 32.1 (2018): 83-123.
20. Müller et al., (2021). libfmp: A Python Package for Fundamentals of Music Processing. Journal of Open Source Software, 6(63), 3326, https://doi.org/10.21105/joss.03326
21. Khusainov, Rinat, et al. "Real-time human ambulation, activity, and physiological monitoring: Taxonomy of issues, techniques, applications, challenges and limitations." Sensors 13.10 (2013): 12852-12902.
22. João Rodrigues, David Belo, Hugo Gamboa, Noise detection on ECG based on agglomerative clustering of morphological features, Computers in Biology and Medicine, Volume 87, 2017, Pages 322-334, ISSN 0010-4825, https://doi.org/10.1016/j.compbiomed.2017.06.009.
23. X. Zheng, M. Wang, and J. Ordieres-Meré. "Comparison of Data Preprocessing Approaches for Applying Deep Learning to Human Activity Recognition in the Context of Industry 4.0." In: Sensors 18 (2018), p.2146. doi: 10.3390/s18072146.
24. Bello J.P., Grosche P., Müller M., Weiss R. (2018) Content-Based Methods for Knowledge Discovery in Music. In: Bader R. (eds) Springer Handbook of Systematic Musicology. Springer Handbooks. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-55004-5_39
25. Müller, Meinard. (2021). Fundamentals of Music Processing: Using Python and Jupyter Notebooks. 10.1007/978-3-030-69808-9.
26. Müller, Meinard. (2021). An Educational Guide through the FMP Notebooks for Teaching and Learning Fundamentals of Music Processing. 245-285. 10.3390/signals2020018.
27. J.-L. Reyes-Ortiz, L. Oneto, A. Samà, X. Parra, and D. Anguita. "Transition-aware human activity recognition using smartphones." In: Neurocomputing 171 (2016), pp. 754–767. issn: 0925-2312.