António Luís Marim dos Santos

Bachelor of Science in Biomedical Engineering

Musculoskeletal Load Exposure Estimation by Non-supervised Annotation of Events on Motion Data

Dissertation submitted in partial fulfillment
of the requirements for the degree of

Master of Science in
Biomedical Engineering

Adviser: Hugo Filipe Silveira Gamboa, Full Professor, NOVA School of Science and Technology, NOVA University of Lisbon

FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

December, 2021

Musculoskeletal Load Exposure Estimation by Non-supervised Annotation of Events on Motion Data

This document was created using the (pdf)LaTeX processor, based on the NOVAthesis template, developed at the Dep. Informática of FCT-NOVA by João M. Lourenço.

For my loving parents and brother,
For my dearest friends,
and For my supportive love.

Acknowledgements

Com o finalizar deste trabalho, que se tem revelado um grande marco tanto a nivel académico como pessoal, há um conjunto de pessoas a quem gostaria de agradecer.

Para começar, a minha gratidão ao Professor Hugo Gamboa, por me ter facultado esta oportunidade de explorar, pela sua orientação, apoio, e compreensão. Um grande obrigado ao João Rodrigues que esteve sempre presente ao longo de todo o trabalho, tendo sido essencial para a conclusão do mesmo. Também o meu mais sincero agradecimento a todos os membros do LibPhys que me receberam.

A nível pessoal, tenho de agradecer primeiramete aos meus pais e ao meu irmão, pela paciência e apoio incondicionais, não só no decorrer desta dissertação mas também ao longo de toda a minha vida. Pela ajuda pessoal e motivação para continuar.

Aos colegas que criei nesta faculdade e que me acompanharam ao longo destes 5 anos letivos, sem os quais não teria conseguido alcançar esta etapa, um grande obrigado. E a todos os meus amigos, que de alguma forma ou de outra me empurraram sempre na direção de ultrapassar os desafios que foram aparecendo durante este tempo. Seria incomportável referir-vos a todos, mas espero que saibam que o vosso valor é reconhecido.

Há, contudo, um conjunto de pessoas que terei de mencionar pelo suporte incondicional e constante preocupação. A minha sincera gratidão para Francisca Andrade, Daniel Zagalo, Diogo Santos, Rita Bento, Sara Piedade e Miguel Rodrigues. Finalmente, um muito obrigado à Telma Esteves pela sustentação que me deu e sem a qual não teria sido capaz de encerrar este capítulo.

"Give me six hours to chop down a tree
and I will spend the first four sharpening the axe."
- Abraham Lincoln

Abstract

There is a significant number of work pressures that promote the incidence of musculoskeletal disorders in industrial environments. As, unfortunately, many workplace conditions are subject to these biomechanical hazards, this has become an extensively common health disorder. To properly adjust intervention strategies, an ergonomic assessment through surveillance measurements is required. However, most measurements still depend on subjective assessment tools like self-reporting and expert observation.

The ideal approach for this scenario would be to use direct measurements that use sensors to retrieve more precise/accurate information of how workers interact with their work environment. Following this approach, one of the major constraints would be that a systematic retrieval of data from a labor environment would require a tiresome process of analysis and manual annotation, deviating resources and requiring data analysts.

Hence, this work proposes an unsupervised methodology able to automatically annotate relevant events from direct acquisitions, with the final intent of promoting this type of analysis. The event detection methodology proposes to detect three different event types: 1) work period transition; 2) work cycle transition; and 3) sub-sequence matching by query. To achieve this, the multivariate time series are represented as a Self-Similarity matrix built with the features extracted. This matrix is analysed for each event needed to be searched.

The results were successful in the segmentation of Active and Non-active working periods and in the detection of points of transition between repetitive human motions, i.e. work cycles. A method of search-by-example is also presented, being that it allows for the user to detect specific motions of interest. Although this method could still be further optimized in future work, this approach has a very promising prospect as it proposes a strategy of similarity analysis that has not yet been deeply explored in the context of ergonomic acquisition. These advances are also significant given that the summarization of ergonomic data is still a subject in expansion.

Resumo

Num contexto industrial, são várias as tensões que promovem a incidência de distúrbios musculosqueléticos. Uma vez que a maioria das condições laborais estão sujeitas a estas propensões do foro biomecânico, os distúrbios musculosqueléticos tornaram-se patologias amplamente diagnosticadas na população ativa. Para desenhar estratégias de intervenção eficientes, é necessário proceder a uma avaliação ergonómica baseada em metododologias de vigilância. Não obstante o reconhecimento desta necessidade, a maioria das medidas ainda depende de ferramentas subjetivas como a auto-avaliação e a observação externa por parte de especialistas.

A abordagem preferencial para esta problemática passaria pela aplicação de medições diretas que recorressem a sensores com vista a extrair informação exata e fidedigna do ambiente laboral. Uma das maiores limitações deste leque de soluções consiste no facto de um sistema de recolha de dados neste ambiente implicar um processo exaustivo de análise e anotação manual, o que consome recursos e requer os serviços de analistas de dados.

Assim, este trabalho propõe uma metodologia capaz de anotar automaticamente eventos relevantes provenientes de aquisições diretas, com o objetivo final de promover este tipo de análises mais eficientes. A metodologia de deteção de eventos proposta foca-se em três diferentes tipos de eventos: 1) transições entre tarefas; 2) transições entre ciclos de trabalho; e 3) procura de movimentos-exemplo em amostras segmentadas. Para concretizar este trabalho, realizou-se um estudo de matrizes de auto-semelhança.

Os resultados provaram-se, na sua maioria, bem-sucedidos no caso da segmentação de períodos Ativos e Não-ativos e na deteção de momentos de transição entre movimentos repetitivos, isto é, ciclos de trabalho. É ainda apresentado um método de procura-por-exemplo que permite ao utilizador detetar movimentos-exemplo do seu interesse. Embora este método possa ainda ser otimizado em trabalhos futuros, reflete uma abordagem promissora uma vez que propõe uma estratégia de análise de similaridade que não foi ainda especialmente explorada no contexto dos estudos ergonómicos. Estes avanços são ainda significantes na perspetiva de que a sumarização de dados ergonómicos é uma linha de investigação ainda em expansão.

Contents

List of Figures

List of Tables

Abbreviations

ABPM	American Board of Preventive Medicine
Acc	accelerometer
AW	Active Work
DALY	daily activity life years
DCT	discrete cosine transformation
DFT	discrete fourier transformation
DM	Data Mining
DoF	degrees of freedom
DTW	dynamic time warping
DWT	discrete wavelet transformation
ECG	eletrocardiography
EMG	eletromiography
ESAW	European Statistics on Accidents at work
EU	European Union
EU-28	European Union's 28 member states (still including the UK)
FFT	fast Fourier transform
GBD	Global Burden of Disease
Gyro	gyroscope
HAR	human activity recognition
HMM	hidden Markov model
ICD	International Classification of Diseases
ILO	International Labor Organization
IMU	inertial measurement unit
KNN	K-Nearest Neighbors

MAE	Mean Absolute Error
Mag	magnetometer
MEM	microelectromechanical
MFCC	multi mel frequency cepstrual coefficients
MP	Matrix Profile
MSD	musculoskeletal disorders
NAW	Non-active Work
Opr	operators
PCA	principal component analysis
SAX	symbolic aggregate approximation
SSLM	self-similarity lag matrix
SSM	self-similarity matrix
STFT	short-time Fourier transform
SVM	Support Vector Machine
USA	United States of America
UT	unifying time
WHO	World Health Organization
Wkst	workstations
WMSDs	work-related musculoskeletal disorders
YLD	years lived with disability

Symbols

A	Accuracy
b^{Est}	point event detected
B_{Est}	Set of all point event detected
b^{Ref}	point ground annotation
B_{Ref}	Set of all ground annotations
$Conv_f$	Convolution function
D_e	Absolute Mean Duration Error
$Dist_f$	Distance function
F	F_n-score
$Overlap_{frac}$	Fraction of window overlapping applied during the representation of the data
Fs	Aquisition Rate
$Kernel_{len}$	Length of the Kernel convoluted in the detection of type event 1 (Work Period Transition)
$Overlap_{len}$	Fraction of window overlapping applied during the representation of the data
$SmoothWind_{len}$	Length of the window applied for the smoothing of the data, during the peak detection phase
$SubWind_{len}$	Length of the window applied for the subtraction of the mean, during the peak detection phase
$Wind_{len}$	Length of window function applied during the representation of the data
P	Precision
R	Recall

Sim_f	Similarity function
$Thresh_{frac}$	Threshold fraction, applied during the peak detection phase

Chapter 1

Introduction

1.1 Context

Currently, industrial environments have been increasingly promoting the investment and inclusion of state of the art technology as well as data analysis techniques to improve, automate and optimize processes. This not only includes processes for the improvement of productivity but also for the improvement of workers' health and well-being. The latter topic has an increasing relevance since such processes promote the prevention of occupational disorders and their future implications.

1.1.1 Occupational Medical Health

Occupational Medicine is the branch of medicine that focuses on people's interaction over their work environment and how it affects their health. It consequently requires the expertise of multiple medical specialities to investigate possible physical or chemical stresses imposed over workers[1, 2]. Additionally, it's important to highlight that this health discipline is intrinsically linked with the subject of preventive medicine, being considered by the American Board of Preventive Medicine (ABPM) one of its primary specialties[3]. The focus on prevention in the workplace, especially primary prevention, has received more attention in recent times, thanks to a large sum of studies that have proven its economic benefits through positive analysis into cost-efficiency, cost-benefits and cost-effectiveness [4]. This given attention is also on the account of various political and medical efforts to ensure healthy labour environments, which have come as far back as 1950, with the Joint International Labor Organization (ILO)/World Health Organization (WHO) Committee on Occupational Health[4], and with the most recent development being the Global Plan of Action on Workers Health[5].

The medical branch of occupational health tends to tackle a wide variety of workplace exposures like chemical or radiation toxins in the workplace, which consequently may lead to various health concerns such as lung, renal, hematologic, neurologic and gastrointestinal disorders, as well as cancer. We will nonetheless concentrate on musculoskeletal disorders (MSD), another serious concern of occupational medicine which is mostly resultant from physical (biomechanical) stresses induced in the workplace. [6, 7].

1.1.2 Musculoskeletal Disorders

MSD are a broad variety of health conditions that affect the locomotor system, also named the musculoskeletal system. This organ system is responsible for the movement and stability of the human body, being composed, by the skeletal and muscular systems, but also by connective tissues: like cartilage, ligaments and tendons. In the International Classification of Diseases (ICD), MSDs compose an entire chapter that lists over 150 diagnoses. These conditions may vary in intensity and duration but being overall characterized 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 [8].

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[9, 10], which on its turn, precludes active participation in all types of social activities and may seriously limit one's work capabilities. Furthermore, pain and disability may also lead to various types of comorbidities, which can be both physical and/or mental conditions [11–13]. Beyond the health burden, there is also a significant economic load, as health costs from MSDs can become very expensive[14, 15]. The sum of all these afflictions, unfortunately makes MSDs a condition that can lead to a higher mortality rate, while making comparisons with the general population[16–20].

This is an unfortunate condition that is extensively common and very widespread. Globally, the best-provided estimation comes from the *Global Burden of Disease (GBD)* studies. Where the latest recordings of 2019 pointing MSDs as the cause for 15.91% (14.13% - 17.78% as it has 95% of uncertainty) of the total number of years lived with disability (YLD)[21], with back pain, in specific, being the single greatest cause for a time living with a disability since 1990[22]. Other conditions that are also described with significantly high values of disability include neck pain, osteoarthritis and rheumatoid arthritis[23]. Despite the significantly great values, this estimation is still greatly underestimated as injury, fracture, sprain or strain (other than hip fracture) are classified under the injury category instead of MSDs.

Between 1990 to 2010 the MSDs value of daily activity life years (DALY) has increased about 60% in developing countries. This trend is due mostly to the rapid demographic shift, which has seen rapid growth and ageing of its population. Adding to that, higher rates of obesity, physical demanding work in subsistence communities, the extension of

urban areas and industrial growth has also contributed to the prevalence of MSDs in low and middle-income countries. Unlike higher-income countries, where such demographic changes have occurred gradually, developing populations have been presented with a too rapid transition, which made them especially susceptible to these conditions[24]. In the United States of America (USA) and European Union (EU), where multiple surveys regarding the health of its population are made, more precise information can be obtained. In the case of the USA, about 50.1% of the adult population (124.6 million people age 18 years and older) self reported one or more conditions of MSDs, in 2015. During 2014 the total direct and indirect costs reached about $980.1 billion, which is about 5.76% of its GDP [14]. In Europe, about three fifths (60.1%) of people reporting for work-related health problems had MSDs[25]. Despite not existing any summarized costs for MSDs from all of Europe, the total cost of only rheumatoid arthritis has reached €45.3 million, with the estimated annual average health cost, per patient, being approximately €13.4 thousand[26].

1.1.3 Occupational Exposure

Concerning the epidemiological protocols on the prevention of disorders, there is an important focus given over identifying the possible causes of the condition, as well as providing quantitative measures over risk exposure[27].

MSDs is, after all, a term that describes a broad range of conditions. As such, to distinguish the variety of possible causes of these disorders, some essays have grouped them into some generic categorizations. In this sense, there are 2 types of MSDs: 1) the acute ones, being usually caused by some type of non-fatal accident, which can become disabling or completely recoverable depending, of course, on the gravity of the incident (e.g. muscle fiber rupturing, as a result of heavy lifting); or 2) the chronic ones, being usually caused by a repetitive and persistent source of physical stress, which might be very light but it's additive effect over time, eventually, causes disabling consequences (e.g. tendon inflammation). This last definition of MSDs is sometimes also described as cumulative trauma disorders or repetitive strain injuries. Additionally, MSDs can also be vaguely classified into 1) initial somatically caused, as some conditions are most of the times caused by a specific origin strain (e.g.carpal tunnel syndrome, caused repetitive wrist movement, or having frequently a flexed wrist) or 2) unspecific trigger cause.

The focus of the present thesis will be on work-related musculoskeletal disorders (WMSDs). This is, MSDs which are induced or aggravated by work and the circumstances of its performance[23]. The work environment is a great contributor to MSDs[28], with the added advantage of being also a controlled environment that can be adapted towards health policies concentrated on primary prevention.

The main physical (biomechanical) risk factors that promote WMSDs include the performance of tasks which may be:

1. more strength intense (intensity), especially if they are being applied during long consecutive periods (duration)

2. tasks applied on a periodic repetitive motion (frequency), with little time of rest

3. poor posture (e.g. unfavourable handling of devices, awkward joint posture, or prolonged restrictive position)

4. application of vibration, which may also lead to another serious condition called vibration white finger.[29, 30]

Beyond biomechanical risk hazards, there are also other concerns to take into account, as is the case of 1) individual and 2) psychosocial factors. Individual factors refers to intrinsic health consideration that might make a person more susceptible to MSDs, these may include age[31, 32], one's gender (with the variations depending on specific conditions, although women are more commonly disabled by MSDs than men)[31, 33], prior medical condition (smoking[34–36], alcohol consumption[37, 38], obesity[39–41]), and genetic predisposition[42]. In contrast, psychosocial factors concern the organization and social support systems existing in the workplace, some examples include job variety, job identity, workload and time pressure[43, 44].

MSDs can be seen as the body's response to a misfitting between the load to which a worker is exposed and its capacity. As such, interventions over WMSDs need to either diminish the exposure to those external loads or to enhance the workforce's capacity. The literature on the subject has pointed measures like engineering redesign, changes in work methods, administrative controls, training, organized exercise, work hardening, personal protective equipment, and medical management to reduce exposure[43, 44].

The discipline of ergonomics is concerned with the study of the interaction between humans and their working environment, promoting that work and workers are in a symbiotic relationship, such that it optimizes both work production and worker's safety[45]. As such, the subject of ergonomics in this context takes a central role as almost every proposed intervention in the literature needs to consider some of its principles. In practice, this usually involves tailoring each workplace intervention towards specific problems where it's required to study the individual worker and the specific work environment. Moreover, it also calls for recurrent evaluations over time, to understand if the implemented intervention is still effective in a changing work environment and if it becomes necessary to propose new intervention methods. As such, surveillance and job analysis become essential in this subject, as when regarding WMSDs it's impossible to tailor a "one size fits all"approach to every work situation. [43]

1.1.4 Measures of Workplace Exposure

As previously mentioned, it is very important to evaluate each individual job-environment relationship to project a proper intervention. In this sense, there have

been developed several ergonomic occupational exposure assessments, falling mostly into three main categories, which differ from each other by their method of data acquisition. These are listed in the following order, by their increased precision and invasiveness on the work performance.

1. Self-report from workers
2. Observational methods (with both screening tools and Video-based)
3. Direct Measurements

The self-reporting protocol enables workers to individually provide the data about their personal work experience. This can be achieved by means of questionnaires, diaries, personal interviews or checklist surveys so that they might be further evaluated by experts. Some common examples include the Nordic Musculoskeletal Questionnaire [46, 47], or the various checklists from National Institute for Occupational Safety and Health[47]. This methodology has firstly the advantage of allowing the possibility to acquire data that would otherwise be impossible to know. This may include information about their past work experiences, as well as individual and psychosocial related data. Moreover, it is usually considered a straightforward, low-cost and versatile tool, which can be applied to a large number of workers.[43, 48, 49] Nevertheless, due to its subjective nature, it can be a flawed system, as workers perception of their symptoms has proven to be imprecise or even totally unreliable[50, 51].

The observational techniques consist of visual surveillance of the work routine by trained experts. The most conventional systems usually retrieve their metrics from pre-designed ergonomic risk assessment sheets. Some examples include: Ovako Working Posture Analysis System (OWAS)[52], Rapid Upper Limb Assessment (RULA)[53], Rapid Entire Body Assessment (REBA)[54], Ergonomic Assessment Work-Sheet (EAWS)[55], Strain Index (SI)[56]. This methodology, like the previous one, has the advantage of being an inexpensive option, which can be applied in various workplace contexts and has a minimum disturbance effect in the workers' activities. Nonetheless, the score based risk evaluation system is not completely precise, leaving yet space for incorrect measurements. The greatest issue of this methodology is its subjectivity as it is subject to intra- and inter-observer variability, being considered more suitable for static and/or repetitive jobs[57]. Moreover, the metrics are based on global indicators that do not fit or take into account the variability of characteristics among the working population, namely anthropometric variation, age, working experience, among others. Beyond this subjectivity related with the analyser and subjects analysed, there is still an issue when yielding only a single score to represent the ergonomic risk of a workstation as it is insufficient to properly explain all risk factors. In addition, although costs are reduced for this kind of assessment, it requires that ergonomic experts spend a considerable amount of time performing visual assessments.

As a response to these limitations, more complex observational techniques may resort to video surveillance to acquire more objective results. This type of approach usually is assisted by software systems, to facilitate the retrieval of specific information. The unfortunate disadvantage of this video-based approach regards the substantial costs of this system as well as the need to have extensive technical support.

The final option of direct measurements, proposes 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 (this will also be extensively detailed in the next chapter). Direct measurements systems, depending on the sensors, provide information about the repetition, force, posture and vibration stressed upon a worker, usually by electrophysiological data or by a biomechanical model. Instead of completely replacing expert observation, it can be used as a support system from where can later be applied ergonomic assessments from specialists, much like the video-based approaches. The limitations of this framework regard the greater costs for its implementations and the narrow success that it had in a practical environment, being most successful in controlled simulated laboratories. Nonetheless, it provides a solution to counteract the limitations of the previous assessment method.

1.1.5 Overall

The work environment has become a place where people spent a lot of their lifetime[58–60], as such, while designing these places in a well-developed and technological structure, it's essential to ensure a level of health quality, as well as the investment over research and monitorization effort towards the minimization of all health stressors. MSDs have revealed themselves as a health crisis with a great prevalence/incidence in both developed and developing countries and the capacity of seriously distorting people's ability to have a fulfilled and productive life. Ergonomics is an essential subject while discussing which are the most relevant interventions. Despite the multitude of developed techniques to properly support job analysis, the current techniques are still mostly supported on subjective methods of self-reporting or expert observation.

1.2 Motivation

Given the previously explained context of a significant health problem largely unresolved but still very prevalent globally in various levels of society, this work will make scientific contributions to the surveillance and measurement of risk factors in the work environment. The algorithm which will be proposed in this work will be used to retrieve relevant information from large manufacturing scenario data sets. More specifically, it will resort to event detection methodologies to summarize the data and identify moments of ergonomic relevance. The intended final purpose will be to facilitate data analysis, so that directly retrieved data may become more accessible to ergonomic experts.

The interest in developing algorithms to detect relevant knowledge from large datasets automatically and preferably unsupervised is not unique to work manufacturing as the modern data collection rate in various areas of society has made manual annotation by humans too tiresome to be feasible. Other real-world scenarios like marketing, investment, fraud detection, telecommunications, and internet agents have also had significant interests in the development of this type of tools[61]. This technological demand has consequently led to the development of the field of Knowledge Discovery in Databases, which in its most abstract concept, addresses this problem by mapping low-level data of significant volume into other forms that may be more compact, more abstract in its description, with the primary intent of retrieving useful information[61]. The focus of this work will be to harness the developments made in this scientific field, specifically in event detection and data mining, and apply it to this context of study, as there is still some lacking research on the topic of event detection on work-related manufacturing scenarios.

Event detection algorithms take advantage of various machine learning methodologies for the detection of relevant events for further pattern analysis. Not restringing itself only to the ergonomic assessment of the workspace, event pattern detection targets various applications, such as supply chain management for Radio Frequency Identification (RFI) tagged products, real-time stock trading, monitoring of large computing systems to detect malfunctioning or attacks, and monitoring of sensor networks for surveillance purposes[62].

Tasks like change point detection, motif identification, and search by example/query have revealed themselves helpful in solving various problems. By analyzing simple dynamic behaviors of periodicity or recurrence, they can also retrieve a lot of information about what is happening during the acquisition. Under a health concern, anomaly detection has been used to make early detection of critical health episodes[63], while motif discovery was used in multidimensional data to facilitate the process of monitorization and to assist the analysis as a conformity indicator[64, 65]. The difficulty related to designing an algorithm of event detection is very linked with the quality of the data, as well as the subjectivity of the event specification. Problems with events of interest that are not well understood or that involve some level of irregularities will require detection processes with a high level of specification, as its definition is considerably hard. However, this may lead to another difficulty because if the methodology becomes very overfitted for a set of trained scenarios, the proposed solution will unlikely be suitable for the detection of other event types in other acquisition settings. This means that the methodology will require a certain level of subject independence as well.

In the occupational context, events are related to relevant moments during the working period. These events can be associated with periodic working tasks, anomalies during the working period, or specific instances of the working cycle that require special concerns. In that sense, the proposed work will provide solutions to detect three main types of events:

- ***Work Period Transitions*** - A working day is not composed of active working during all the time. Workers in manufacturing need to occasionally interact with each other or wait for the next assembly section, and these will also be reflected in the direct acquisition of data from the assembly line. By detecting transition between different types of motion, it is possible to segment from the larger database the motion made in different workstations, the active working period, or alternatively the non-active working periods.

  Advantages: The identification of events that indicate the separation between moments of active working and non-active working will allow a more clean dataset for the ergonomist specialists to work with by focusing only on the interesting temporal interval. Furthermore, by separating different workstations motions, it is possible to analyze the work strains in a specific workstation in isolation or by comparing different workstations.

- ***Work Cycle Transitions*** - If we assume that the data being analyzed presents a periodic behavior, we can further divide working periods into sections of work cycles. This assumption is possible in this working context as it is composed of short activities that are recurrently performed. Then, the detection of this event type will try to identify the frontiers between these periodic work motions.

  Advantages: It allows for an ergonomic analysis over a single motif to generalize the entire work period when the subject is on the same workstation. This is commonly made in the observational methodologies, where single postures (usually the ones that present more considerable ergonomic strains) are analyzed to assess the entire working performance[66]. Conversely, it also facilitates a comparison between the different work cycles over the entire working period to understand if the ergonomic risk changes considerably over the extension of the working period.

- ***Sub segment example search***: The last events this work will study are the identification of specific previously predefined motion postures in a working cycle by means of a query example.

  Advantages: This targets a more interactive methodology and has the advantage of studying the risk of specific segments of the working cycle, as well as understanding how these segments change in behavior, dynamics, and risk factors over the entire working period.

In summary, this thesis will try to add a novel proposal for detecting events in an ergonomic context. Occupational medicine has great importance in the quality of life of the larger part of society. This work will intend to tackle the lack of studies proposed towards the annotation of significant events in a work-related context, a relevant assist for the analysis of larger data. Moreover, if the intent is to make this type of data accessible, especially to experts of this field, the data needs to be translated into a higher dimension,

more compatible with the terminology used in ergonomy. To solve these requirements, this thesis will focus on detecting three different types of events that try to first structure the entire working period acquisition.

1.3 Objectives

The main objective of this work will be the development of a non-supervised algorithm that identifies different events on time series data in an efficient matter. This method is supposed to be applicable to data directly retrieved by means of inertial sensors in a working manufacturing environment. These are the minimum practical circumstances in which the present algorithm needs to be valid. Adding to that, the algorithm will also need to demonstrate a set of important characteristics to be an accessible alternative to be used. As such, the proposed algorithm will be developed taking into consideration the following:

Agnostic to time series type: As the data acquired in occupational scenarios can be of multiple types (motion, angular, electrophysiological), the algorithm should be fitted to analyze all types of time series.

Multidimensionality: Time series datasets in occupational scenarios are multivariate, and all dimensions contribute with relevant information. The algorithm should be able to cope with multiple time series. Adding to that, as human motion is a complex system that can be segmented in various ways depending on the interest of each motion, it is useful for the returned values to be parameterized in the same way. This way, it allows for the user to adjust the sensibility of the analysis procedure. As such, the methodology needs to be able to use and return multidimensional structures of data efficiently.

Time-Efficiency: All this processing must be made in an acceptable time frame. Even though the analysis will be done in an offline procedure, the amount of data can become very extensive as a proper analysis may require taking into account various workers performing different types of tasks along different periods. Moreover, working schedules can be very long, and multiple time series may require to be analyzed simultaneously. Despite these impositions, the algorithm is still required to be fast or at least to allow processing in a reasonable amount of time so that it may be further extended into more complex analysis contexts.

Interpretability: This algorithm will try to provide information that may facilitate the access of data analysis to professionals, who may not have a computer science background, like ergonomist technicians, but who nonetheless should have the means to work with the direct retrieval of data to reach more accurate conclusions. This will transpire through an analysis of the three types of events previously described in the motivation section 1.2, which will create a structure to the data acquired in this acquisition context, and as such, allow the analysis to be done over pre-segmented periods of interest.

1.4 Structural Summary

The purpose of the current chapter was to focus on explaining the practical medical context, along with the more theoretical computational context in which this work will try to insert itself by explaining the context motivation and objectives of this thesis. The main body of this thesis will be further divided into six different chapters of interest.

The chapter 2 "Theoretical Background"will further expand on the context of this work by explaining concepts that are pre-required to understand the remaining thesis properly. Subsequently, the chapter 3 "Literature Review"will give an overview of the more current developments made in the scientific literature in the interest of this work.

The description of the algorithm, as well as the processes of testing its validity, will be done in the chapters 4 "Database Description"and 5 "Methodology". The first will describe the databases which were used and their acquisition setting. Then, the chapter 5 will explain what were the various event detection techniques used in this work, and how they were fitted to the database described in chapter 4.

The results retrieved from the application of these various methodologies on the described databases will be presented and then simultaneously discussed in the chapter 6 "Results and Discussion". After that point, the conclusions regarding this entire work will be summarized in the last chapter 7 "Conclusions and Future Work". Also, in here, there will be given some suggestions on how to further expand over the work of this thesis.

In the end, there will also be provided appendices to serve as support material for the results and methodologies described over this thesis.

Chapter 2

Theoretical Background

A collection of observations made chronologically during some time are usually stored in time series objects. It is then predictable that time series are vastly used in various scientific circumstances, with multiple types of applications. In this chapter, the fundamentals of all necessary terms and concepts will be described. It will start by explaining the definitions associated with (1) what represents a time series and further related definitions; (2) what is the definition of an event, the concept of event detection in time series and how it can be seen as a data mining task for the pattern analysis, further including a description of what are the typical methodologies usually used; and (3) the definition of the SSM, what it is, what represents, and how can it be used for the detection of time series' events.

2.1 Time Series

This chapter will serve to define and explain some required terms about time series definitions, and data mining techniques applied to time series.

Definition 2.1.1 (Univariate Time Series)**.** A Time Series is an ordered sequence of n real-valued variables, indexed by time.

$$X = \{x_t\}_{t \in T \subseteq \mathbb{Z}^+}$$

, or

$$X = (x_1, \cdots, x_n)$$

Under this definition x_t is a *point*, or *observation* collected at a particular time t. Meanwhile, a subsequence is a smaller time series $S = x_p, x_{p+1}, \cdots, x_{p+m-1}$ of length $m <= n$ and that starts in a position $p \in \{1, \cdots, n-m+1\}$, from a larger time series X.

A time series is usually the result of an observation analysis made at uniformly spaced time instants with a specific sampling rate. Otherwise, if not uniformly spaced, it requires an array of length n, to label the time instance corresponding to each point observation x_t.

Definition 2.1.2 (Multivariate Time Series)**.** A multivariate time series is an ordered array of k sequences, each one with real-valued variables in its positions and all indexed by time.

$$X = \{x_t\}_{t \in T \subseteq \mathbb{Z}^+}$$

, where

$$x_t = (x_{1t}, \cdots, x_{kt})$$

, or, because all the sequences have the same sampling rate and the same length size n, it can be represented as

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{k1} & x_{k2} & \cdots & x_{kn} \end{bmatrix}$$

Multivariate time series are the result of k observation analysis occurring at the same time, at specific sampling rates. One example could be the acquisition and storing of multiple sensors that acquire physiological data from a human subject at the same time. In here, x_t is also a *sample*, or observation at time t, with the difference that instead of being a real-valued variable, it is an array of length k. The same applies for the definition of subsequences, which remains as $S = x_p, x_{p+1}, \cdots, x_{p+m-1}$, with $m \leqslant n$ and $p \in \{1, \cdots, n - m + 1\}$, however it adds k dimensions to the previously defined univariate subsequences.

Definition 2.1.3 (Time Series Database)**.** A Time Series Database is an unordered set of time series.

A database can contain multiple time series, either univariate or multivariate. Moreover, sample observations of different time series in the same position are not required to correspond to the same time instances.

Definition 2.1.4 (Feature Matrix)**.** Given a time series X with a length n and in the case of being multivariate, also with dimension k, then a "representation"or "feature matrix"of X is a model $\overline{X}$ that has extracted attributes about the time series, and usually has a different dimension or volume of data, represented by $n*$ and $k*$, respectively.

To obtain a Feature matrix, it is necessary to consider a specific feature retrieval function F, such that

$$F\colon \mathbb{R}^{n\times k} \to \mathbb{R}^{n^*\times k^*}$$
$$X \mapsto \overline{X}$$

This type of function is usually characterized by:

- **Applying simultaneously a specific set of feature functions to X**

  Consider the example of three simple feature functions $mean()$, $var()$, and $ptp()$, which return the mean, variation and peak to peak distance values, respectively. All these functions take as input an array of values and return only a single value number. Then, if it is also considered that F is composed by these three functions, in that case, each one of these minor functions will be applied over the k dimensions of the multivariate time series X.

  To clarify the explanation, if the operation is described in its simpler form as

  $$\overline{X} = F(X)$$

  then, reciprocally this means that

  $$\overline{X} = \begin{bmatrix} mean(X) \\ std(X) \\ ptp(X) \end{bmatrix} = \begin{bmatrix} mean(x^1) \\ std(x^1) \\ ptp(x^1) \\ \vdots \\ mean(x^k) \\ std(x^k) \\ ptp(x^k) \end{bmatrix}$$

  This way, the value of $k*$ of the resulting $\overline{X}$ representation will be equal to the number of features in F (which in the case of this example were 3), times the dimension k of the time series X.

- **Being performed over a windowing function**

  The feature representation of a time series $\overline{X}$ is obtained by sliding a window over the length of the time series X. For each window, a subsequence of the time series is selected, and the F function is applied over it. Considering an example where the time series X is divided into $n*$ subsequences by a window function of constant length L. These multivariate subsequences will be represented by $(x_1, x_2, \cdots, x_{n*})$, and each one will have a fixed dimension of $(L \times k)$. Then, applying a function F, it is equivalent to

  $$\overline{X} = F(X)$$
  $$\overline{X} = \begin{bmatrix} F(x_1) & F(x_2) & \cdots & F(x_{n*}) \end{bmatrix}$$

by also joining the conclusions retrieved in the previous point, then the final result will be

$$\overline{X} = \begin{bmatrix} mean(x_1^1) & mean(x_2^1) & \cdots & mean(x_{n*}^1) \\ std(x_1^1) & std(x_2^1) & \cdots & std(x_{n*}^1) \\ ptp(x_1^1) & ptp(x_2^1) & \cdots & ptp(x_{n*}^1) \\ \vdots & \vdots & \ddots & \vdots \\ mean(x_1^k) & mean(x_2^k) & \cdots & mean(x_{n*}^k) \\ std(x_1^k) & std(x_2^k) & \cdots & std(x_{n*}^k) \\ ptp(x_1^k) & ptp(x_2^k) & \cdots & ptp(x_{n*}^k) \end{bmatrix}$$

,where the symbols of x_j^i, represent the dimension i of the subsequence j. This final feature matrix $\overline{X}$ has the dimension of ($n*\times k*$). The value of $n*$ will be dependent on the length of the window L and, if used, on the overlap percentage of the windowing process.

2.2 Time Series' Events

As previously mentioned, this work will propose a technique to find and annotate relevant instances within a large time series, which can provide useful information regarding an ergonomic analysis. This analysis will be general and unsupervised, and as such, it will have the difficulty of analyzing long, complex, and multivariate time series datasets without any previous assumptions.

The following chapter will first define what is an event in the field of time series data mining, associating this definition with pattern analysis. This first part will conclude by defining the events of interest for this work and explaining their selection. The second half of this chapter will present the theoretical background of algorithms used for the intent of event detection, with these being further divided into three tasks of importance in the field of data mining. This final section will have the objective of introducing terms of interest for this work while at the same time providing some state of art examples regarding the topic of event detection in time series.

2.2.1 Event Definition

As it was briefly approached in Chapter 1.3, this work's main purpose is the development of an algorithm for event detection applied to the specific context of occupational data and ergonomics.

An event is the state or change in state of the system under study

This definition of events comes from [67], and its intent was to link it with the process of pattern analysis. Through this basic description of events, patterns could be defined as a combination of events correlated over time[62]. Some easy to understand examples of events that could be retrieved in biosignals include heartbeats, EMG onsets, or specifically

directed thoughts from brain-computer interfaces (BCI) [68]. The search for these events has to take into account the type of event as well as the dimensions of its search.

The subject of events has already been widely explored in literature, by focusing on various social settings, much progress has been made in the development of event driven architecture systems and detection methodologies for events specific to each contextual problem. Events are, in various social fields, a well established tool to describe and analyze data. The existence of event driven architecture systems allows for tools which may detect (usually by means of a query language), process and act upon events. These have already existed for decades being initially targeted towards the field of military command and control systems and SCADA(supervisory control and data acquisition) systems, with more recent decades having seen a shift towards the investment by way of companies.[69, 70]

The focus of this thesis will be on the actual process of detection, developing a methodology able to find the predefined events of interest, with recourse to machine learning methodologies. Machine Learning has already been proved as a valid way to tackle the task of event detection. With this said, there are usually some particular issues of event detection, which are very contrary to the standard machine learning algorithms. As for starters, the events of interest typically are just a minuscule fraction of the data, which introduces a class imbalance problem. Moreover, the data is not normally sufficient, especially for detecting anomalous events of interest, which means that there is usually required previous information regarding the domain being analyzed. This means that if an algorithm is applied under a different context, the algorithm may require to be significantly re-engineered. These will be issues which will be present tougouy the work and will try to be minimized as much as possible to try to create a tool that requires as little information regarding the dataset, with as little optimization interventions as possible by the user. [71]

With this said, an event, as seen by the definition of [67] which was presented at the beginning of this section, is a subjective event that will depend on the database used and on the interest of the analyzer. To further specify an event in a more objective, it is necessary also to refer how and where the search processes will be applied. This can be done in several ways, with the three main points being:

1. Search Dimensionality: This point is, of course, only pertinent for multivariate time series. With this said, events do not necessarily occur globally in all dimensions of a time series, as such, the analysis has the option of taking into consideration only one (uni-signal), some, or all the dimensions (multi-signal) of that time serie. Considering the analysis of EMG signals retrieved from multiple muscles of the arm, and with each acquisition being represented in different dimensions. During hand movement, the consideration of the *biceps brachii* and of the *triceps brachii*, are useless for the analysis, worst, they might be providing noise associated with detached movements.

2. Search Window: The temporal length of the window function at which events are searched is relevant, as different events will be observed depending on the time scale of the search. For instance, consider the human visual analysis of an ECG signal. If we consider a signal with 2 hours, we will be able to see events related with global fluctuations with time during this 2 hours period, but when zooming in, we would be able to find events that are related to the QRS complex (if in a time scale of around 10 seconds). As such, the analysis may be on a single specified time length or over (uni-scale) or along multiple time scales (multi-scale)

3. Search Domain: As time series can be converted into other domains, so can events, and instead of these being detected and represented only by a time stamp, they will also be defined by the other domain of concern. This, naturally, can be made in immense different ways as there are multiple methods of feature representation. However, excluding the standard time domain, two common domains of analysis would be amplitude and frequency.

This process is further schematized in figure 2.1

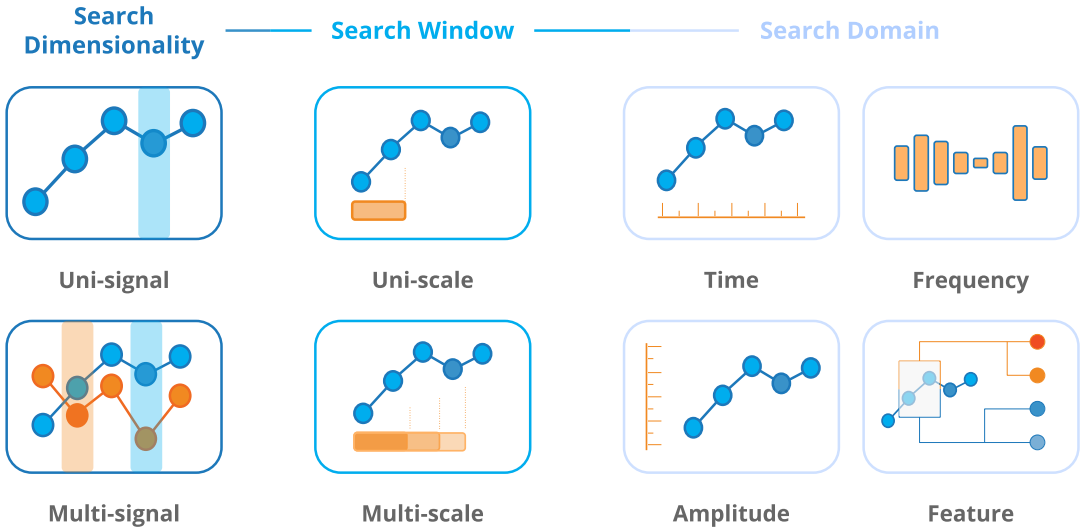


Figure 2.1: Schematic representation of the different types of search which can be made for event detection

Considering the aforementioned definitions found in the literature for what represents an event, we are able to associate it with the type of events that are searched in this context. The search will be over multiple dimensions, uni-scale, and be applied in a multivariate feature domain. Based on what was mentioned in Chapter 1, the events that will be searched will need to follow different strategies of search for detection of the different types of events.

- **Work Period Transition**
  Events which describe points of transition between different work activities.
  Labor context acquisitions, as a rule, do not return uniform signals. Instead, they usually represent multiple sequences of time periods with consistent behaviors, but still all very diverse between themselves, depending on the motion being performed. For example, the case of a manufacturing context as the one which will be analyzed in this thesis has to consider periods of repetitive work in different work stations and inactive middle periods of rest. As the repetitive tasks create a signal with periodic pattern structures, changes occurring in these patterns can be seen as breaks in the natural behavior of that work, and consequently, as possible transition to another motion state with a different pattern behavior. By detecting the changes occurring in a constant dynamic behavior, it is expected that the entire signal can be structured into smaller motion states which represent the type of motion being performed.

- **Work Cycle Transition**
  Events that describe points of transition between the work cycle motion of a given active period.
  As mentioned in the previous task, active working periods are composed of working cycles that are repetitive in nature. By only focusing on the section of the signal where these work cycle structures are occurring periodically, then it is expected that by looking for structures which are considerably similar with themselves, it will be possible to detect the transition points between these work cycles.

- **Sub-segment Example Search**
  Events that describe points occurrences where a pre-specified motion happens.
  These events are associated with an example provided by the user, and to be found require an insistent search over the various subsequences of the time serie

2.2.2 Event Detection Methodologies

Various fields have offered methodologies for event detection, with them usually fitting on the label of traditional methodologies of Data Mining (DM), despite some adaptations for the requirements of the contextual problem. This makes sense as DM is very related with the notion of finding useful patterns within large amounts of data and the definition of events, was initially referred as a way to denote patterns, by decomposing them in smaller constituent characteristics. In a formal description, DM is a field of computer science that uses expertises from artificial Intelligence and statistics, with the purpose of retrieving meaningful knowledge from large databases [1]. DM has also become one of the most well known and revised of these terms, being used by statisticians, data analysts, the management information systems communities and being popularized in the fields of artificial intelligence and machine learning[61]. A better understanding of

the relationship between these two subjects is that event detection can be seen as a multi task approach of data mining, which usually tends to englobe more traditional tasks, like change point detection, motif identification and search by query, between others. As such it's relevant to understand and study the vast literature on the subject of DM to better understand what methodologies can be harnessed towards our intended objective of event detection.

The literature regarding the process of event detection in time series is fairly consistent and describes the existence of two diferent event types [72, 73], differentiated by their detection methodology:

- **change point events**. Events which represent an instant point where there was a sudden and significant change in the properties of the time serie. These points samples may represent a change in the dynamic state of the signal, or the application of an instant operation in the examined system. The event type 1 (work period transitions), will fall under this description

- **similarity based events**. Event which represents an instant points where an predefined signal complex is repeated in the signal. Unlike the previous point, this is more focused on identifying time instants with greater similarity, either to predefined queries or examples, being an event type commonly related with the field of pattern matching. Under this description there will be explained the two remaining events of type 3 (Sub Segment Event Search) and type 2 (Work Period Transitions). The strategy for these last two types of events differs considerably. Both event types do search for the time instants of greater similarity, but while the type 3 event (Sub Segment Event Search) compares the subsegments of the time serie with a specific model example, the type 2 event compares all the subsegments of the time serie with themselves to highlight motifs of interest. These events intersect with the disciplines of search-by-example/query and motif identification, respectively.

Because in the start of the proposed work the computing was not narrowed to any specific computer methodology, and there is a vast and varied literature on the subject of event detection, it was necessary an extensive component of literature research and experimentation of various techniques before actually compromising with a specific path. In this regard, the following chapter will serve to give an overall description of the standard procedures to really find these events. As this is a large subject, only some highlights will be provided. The structure of the following section will be divided in two, according to the event definition previously introduced.

2.2.2.1 Change Point Events

Change points are abrupt variations in the time series data, representing a transition between states where the properties of the time series change in a significant way[74]. Multiple methods able to solve CP problem have been harnessed for the objective of event

detection in general. The most common approaches use statistical tools like the maximum likelihood function [75–78], where the principle follows that if sub sequences separated by a change point will have probability distributions significantly different, and through monitoring the logarithm of the likelihood ratio between two consecutive intervals, these events can be identified. Another popular method is the fitting of distribution models like Poisson model[79], Bayesian Model[80] or other methods, such as Monte Carlo[81], to determine events of interest. Most of these methods follow the principle that if a time series belongs to a certain state, then it follows a specific statistical model, and in the time instants where that rule is less stable, there is a greater probability of existing a change point that divides the time series into two separate states[74]. Adding to this, genetic algorithms[73], clustering and graph based methodologies are also found to be used for this purpose [82].

These are all unsupervised methodologies of greater interest for this dissertation, notwithstanding there is also a significant amount of work on supervised approaches of change point detection. Supervised methodologies can be divided into 2 strategies, depending whether it makes a binary or a multi-class classification[74]. In the case of a binary classification, the objective is to identify the time instants of transition between states, using methods like Support vector machines[83, 84], logistic regression[83], AdaBoost[83], decision tree[83] or random Forest[85]. On the other hand, multi-class classifiers have all interesting states specified, considering each separation between consequent states as a change point. This technique has a much wider literature on the subject being mostly related with the fields of classification, segmentation and summarization.

2.2.2.2 Similarity-based Events

As already mentioned, the interest of this work will insist on the process of motif identification and Search-by-example. Both these methodologies involve assessing the similarity between subsegments of time and the remaining time series. Where they differ is that Search-by-example compares a specific subsequence selected as an example with the remaining time series under analysis, with the final objective of understanding when and if a specific dynamic behaviour happens. However, Motif identification will compare all the subsequences present in a time series with themselves, with the final objective of highlighting which of these subsequences present a significant prevalence within the signal. Both of these methodologies can be very useful to analyse the recurrence behaviour of a time series.

Both of these types of similarity based events follow a procedure with the steps of "data representation" and "similarity measurement", being very crucial for the well performance of the algorithm. Data representation underline the transformation of the raw data into a different domain of features. This process has the advantages of turning the data into an higher level of abstraction, which might be more relevant for the process of detection, and of enhancing the speed of processing of the algorithm. With the most

commonly used in motif identification being symbolic aggregate approximation (SAX), discrete fourier transformation (DFT) and Random Projection [86], meanwhile Search-by-query algorithms tend to focus on DFT, and SAX as well, but also on discrete wavelet transformation (DWT)[87], DWT (more precisely Haar wavelet[88]), single value decomposition (SVD) and discrete cosine transformation (DCT)[89]. Regarding the similarity measurement, this will always need to be defined in these types of algorithms as it is essential for the comparison of the various segments. Some of the most used similarity measures used in both algorithms are Euclidean distance, correlation coefficient and dynamic time warping (DTW) [86, 90]. In the case of comparing strings due to a symbolic representation of the data, Edit distance can also be used[91].

The term motif originally came from the field of genetics, with a DNA motif being considered a nucleic acid sequence pattern with some sort of genetic significance, being usually fairly short and known by recurring in different genes or several times within the same gene [92]. In resemblance with this description, motifs in a temporal database are defined as approximately similar subsequences, which tend to also be repeated within a larger time series. These patterns usually carry important information about the dynamic behaviour of that time series, being events of great importance.The case for motifs in time series was introduced only in 2002 by the work of [93, 94] which did propose an identification methodology for k-frequent type motifs. This method tries to identify which candidate segment has the most number of neighbours within an R threshold of similarity by reducing the data through the use of SAX and then by comparing the resulting data with Euclidean distance metric with recourse to a MINDSIST function. As this methodology has a high time complexity, the work in [95] surged to answer this problem of poor scalability by introducing the use of random projection in time series, resulting in equally good results with a significantly smaller time demand. Another important contribution came from [96] which proposed an algorithm of detection of exact motifs, with the use of early abandonment to optimize the performance of the algorithm for large amounts of data.

Query by content or indexing is a data mining task, which, within the interest of this work, is the process of finding a set of subsequences within a time series that are more similar to a query or example provided by the user. This might be one of the most active areas of research in time series data mining, as it is also the oldest theme to be explored in this context. The paper of [87] is usually pointed as the research which laid the foundations for this type of techniques applied on time series [87, 97]. In this process, the data was firstly reduced using the discrete Fourier transform(DFT). Later, an a R*-tree was constructed as an indexing structure to retrieve the sequences that were at most of a threshold distance to the query sequence, then in a post processing step, false hits were removed by computing the euclidean distance within the time domain. Advances in research studies also tried to further optimize this methodology [98].

2.3 Self Similarity matrix

Within the previous section 2.2 there were very briefly summarized several classical approaches used for the problem of event detection. Despite the wide literature on this subject, the greatest difficulty relates with trying to unify the processing in some consistent methodology. As the intent in practice will be to try to detect three different types of events, there is an interest in providing a methodology relatively similar for all of them to optimize the entire process. As seen in the preceding section 2.2, there is a great significance over studying the degree of similarity within subsequences of the same time series. As such, this work opted for a study based on self similarity matrices.

A self similarity matrix is a graphical representation of the similarity between each data point and all the other points in the rest of the same time series.

Definition 2.3.1 (Self Similarity Matrix)**.** Given a time series $X = (x_1, \cdots, x_n)$. The self similarity matrix S, from the time series X, will have a size $n \times n$, where each of its position is defined by

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

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

From this definition, it is also important to note that the distance function can take various forms, with some examples being displayed further in the section 3.3.3. Another consideration is that the time series X can be either univariate or multivariate, with the latter being more frequent, as what is usually computed is not the raw time series data but a multidimensional feature matrix of the signal. These intakes that s needs to be a method that is able to compare vectors, position by position, with the prerequisite of still remaining undemanding in its computation. It is expected for the matrix to display illustrative structures according to the recurrent behavior of the signal. These structures are described in the following subsections.

2.3.1 Main Diagonal and SSM Scheme

Some behaviors are predictable to happen in every SSM. Starting by the Main Diagonal, it is expected that there would always be an intense diagonal across the matrix, as these will be the positions where the points are compared with themselves, which means that for a given $x_i \in X$, it is expected that $s(x_i, x_i)$ will acquire a maximum value, independently of the coordinate i.

Another important characteristic is that the SSM is symmetric along the Main Diagonal. In other words, for any two points x_i, $x_j \in X$, then $s(x_i, x_j) = s(x_j, x_i)$, which is equivalent to $S(i, j) = S(j, i)$.

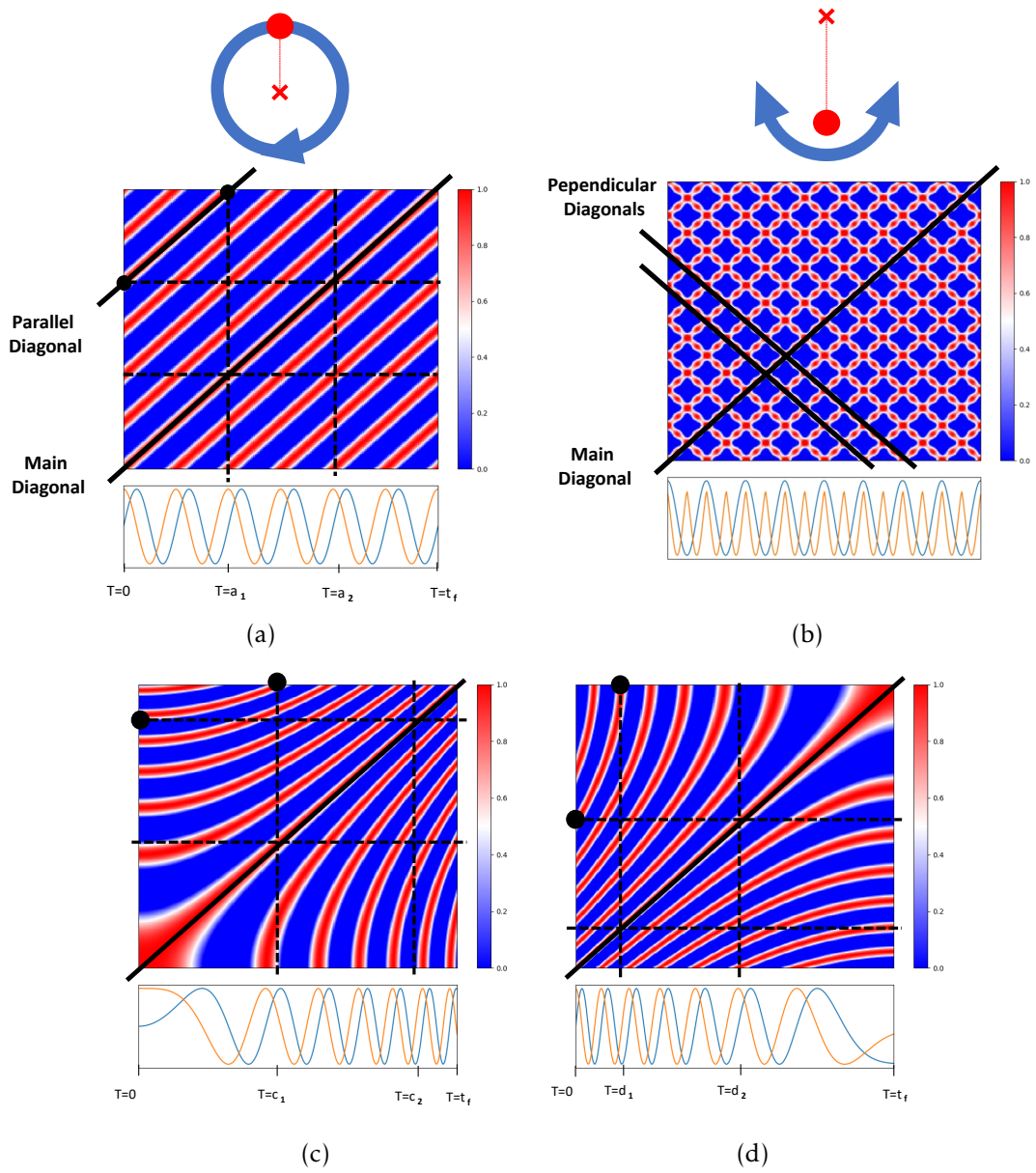


Figure 2.2: Schematic examples of simple motions of a point mass and the respective SSM from the signals of position of that mass. The activities performed, in relation to the subfigure presented are : (a) circular motion around a center point at the same velocity; (b) pendulum motion; (c) circular motion around a center point at accelerated velocity; (d) circular motion around a center point at decelerated velocity. The signal in orange represents the position of the mass along a vertical axis, across time; while the signal in blue represent the position of the mass along an horizontal axis, across time

2.3.2 Parallel diagonals

The existence of diagonals parallel to the main Diagonal is very relevant, as this means that two subsequences of similar lengths are very similar with each other. In subfigure 2.2a, there is highlighted a parallel diagonal, which means that within the periods of $[0,a_1]$ and $[a_t,t_f]$, there are two equal subsequences. Diagonals with a higher or lower slope reveal that also a pair of subsequences are very similar but with different time spans. The joint accumulation of various subsequences reveals a cyclic pattern of recurrent subsequences. This cyclic pattern can also vary in its slope across time, which reveals a

shortening or widening of the cyclic events. For example in subfigure 2.2c, the SSM has a diagonal indicating that the periods of $[0,c_1]$ and $[c_2,t_f]$ are equal. However, as the last subsequence is much shorter than the first subsequence, it produces a distorted diagonal. The opposite is observed in subfigure 2.2d, between the $[0,d_1]$ and $[d_2,t_f]$ subsegments.

2.3.3 Perpendicular diagonals

Parallel diagonals represent segments of the time series with similar behavior, so, predictably, perpendicular diagonals to the main diagonal represent a point instance where there was an inversion of the cyclic behavior. This can be more easily observe in subfigure 2.2b, as the various time instances where the perpendicular diagonal crosses the main diagonal, there is a repetition of the motion but in reverse, with the previous one.

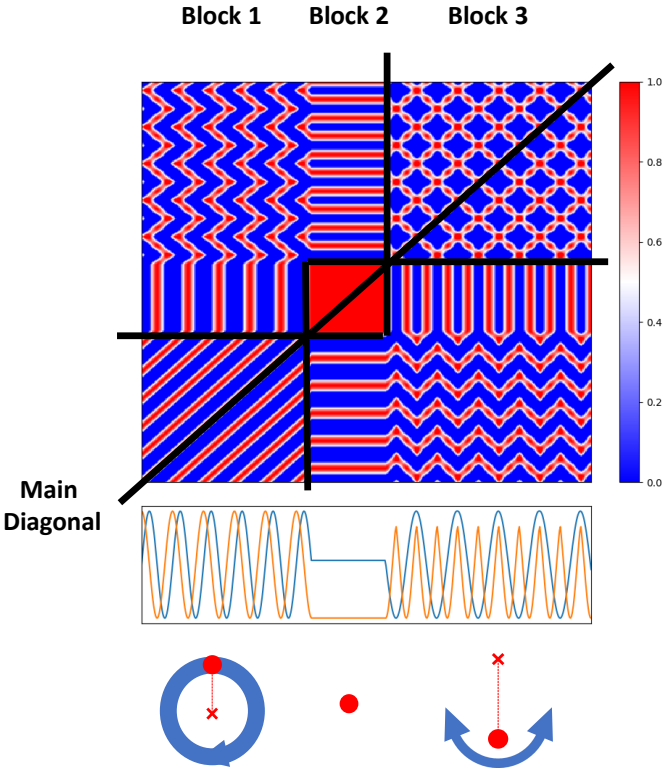


Figure 2.3: Schematic examples of varying motions of a point mass and the respective SSM from the signals of position of that mass. The signal represents a motion which starts by being cyclical (Block 1), stops in position for some time (Block 2), and then restarts its motion, but now in a pendulum form (Block 3). The signal in orange represents the position of the mass along a vertical axis, across time; while the signal in blue represent the position of the mass along an horizontal axis, across time

2.3.4 Blocks

Blocks describe the squared structures developed in the intersection with the main diagonal. With figure 2.3, highlighting 3 such blocks. Blocks are usually visible distinguishable as they possess structures with a consistent texture. The prevalence of these blocks usually means that the signal displays a consistently recurrent behavior during that time period. The transition between different blocks is also of special relevance as it indicates a significant variation regarding the dynamic behavior of the signal.

All the described structures are relevant for the detection of the mentioned events in the *Introduction*. Further, a deeper explanation of the state-of-the-art in strategies for event detection will be presented, namely with the usage of the SSM, more specifically in the audio field, where this method is found to be very used and was the inspiration for the problem in matter.

Chapter 3

Literature Review

In this chapter, the interest will be to expand on the previous chapter and explain essential theoretical concepts from this field of study, but distinguishing itself from Chapter 2 by instead providing a more detailed state of art for concepts of greater importance.

The first interest of this chapter will be to describe the current direct sensing techniques applied in an occupational analysis scenario, a topic in continuation with section 1.1.4. Followed by this, it will go into greater detail over what are inertial measurement unit (IMU)s, the sensing devices chosen for the data acquisition of this work, as well as their advantages against other sensing techniques. Detailing also about the standard procedures used regarding IMU devices and what the literature more prominently incited upon with this acquisition technique.

After providing this initial presentation of the field of interest in section 3.2, the focus will change to present and explain the methodologies which came closer to actually detecting events of ergonomic interest as this work is trying to do. Followed by a description of how SSM can be used for the detection of events in general, detailing more with examples of application on audio signal analysis, where this method is highly used for several tasks.

3.1 Direct Sensing Techniques

Resuming from a previous section (section 1.1.4), there are multiple ways to assess the risk hazards in a work environment. Among all options, Direct Measurement is a technique mostly trustworthy and considerably more objective than the alternatives of self-reporting or sole expert observation. Direct Measurement is characterized by retrieving its values from measurements by sensors. Due to the extensive array of occupational considerations, there are also multiple occupational surveillance technologies possible

to be used. This section will serve to introduce what are the types of surveillance technologies that exist to assess the labor environment, followed by an explanation for the selection of IMU sensors concerning this work.

The most trivial technology which has been used in this context is pressure sensors. Giving information about force exertion, they have been tools more directed towards office work. For instance, the analyses of strength, applied into computer mouses[99] and keyboards[100], have been used to measure the risk for upper body MSDs. To monitor the posture of a subject and identify possible hazards, some classical biomechanical sensor approaches would be to use 3-axial accelerometers as inclinometers[101, 102]. A repetitively referred example is the lower motion monitor. This tridimensional electronic goniometer is connected to the pelvis and thorax and can measure the trunk's position, velocity, and acceleration[103], as well as spine loading while being assisted with an EMG system[104].

A physical demand monitoring system has been proposed to assess physiological parameters (heart rate, eletrocardiography (ECG), breathing rate, and skin temperature) along with environmental conditions (temperature and humidity)[105]. An exciting proposal, which despite not being implemented, highlights the relevance of using electrophysiological signals to properly assess the external pressure to which a subject is being submitted. In this sense, eletromiography (EMG) is a uniquely relevant tool that is very extensively used in occupational surveillance. While measuring the muscular activity, EMG allows for an objective analysis of muscular stress, muscular fatigue[106, 107] and to report about muscle recruitment for a given action. Another relevant detail is how EMG gives information about the intensity of a given activity while also considering the subject's physiological capacity for that action [108–111].

However, to obtain a full posture assessment, there needs to be the application of several markers placed in strategic points of the human body. Through the position of these markers, computers could then design an anatomical model to analyze the subjects' posture. These sensors usually fall into three different types of technologies: a) optical scanning systems[112, 113]; b) sonic systems[114–116]; and c) electromagnetic systems[117, 118]. Optical scanning systems are associated with a framework where the markers are light-reflecting, and their displacement is then detected by light scanning devices (like mirror scanners or electric image detectors) [112, 113].

Sonic Systems analyses the frequency of ultrasonic waves produced by an oscilloscope. This method employs the Doppler effect to measure a body's segment's displacements, velocities, and accelerations. Even though it has shown promising results, there are many limitations concerning the recording speed and the environment requirements as, depending on the transmitter used, it takes from 0.23 to 3.3 s to acquire a single posture measurement. Moreover, environmental noise can be produced by temperature, air density, humidity, and atmospheric pressure, requiring careful calibration.

In an electromagnetic system, a pulsed magnetic field is created by a transmitter, and this signal is then read by receivers attached at specific body points. After, the

receiver sends data about their position and orientation to a computer. This model has an advantage regarding the previously mentioned methods as it provides information along 6 degrees of freedom, allowing for the interpretations of rotation movements. [119]

However, one of the most promising direct measuring techniques has by far been the IMU sensors. Constituted by a set of multiple inertial sensors merged in a single small device, it also has been a sensor used as a marker unit to be attached to the human body. Nevertheless, due to the concentration of various sensors in the same device, these can provide information regarding each body point's position and their movement (namely their force acceleration, angular velocity, and orientation). They are very discrete devices, which is very important as the surveillance systems need to be as less invasive as possible, so they do not affect the work performance.

3.1.1 Inertial Measurement Units

An IMU is an electronic device composed of three 3-axial sensors: a) Accelerometers; b) Gyroscopes; and c) Magnetometers; which allows it to register an object's specific force, angular momentum, and position simultaneously. Therefore, when attached to strategic points of the human body, these sensors can provide much information about the movement and posture of the subject by retrieving the position of body segments and the joint angles between them. Other wearable sensors (like optical scanning systems, sonic systems, or electromagnetic systems) retrieve this information. However, in this sense, IMUs have the advantage of also providing the orientation of each sensor, which can give the added information of the rotation movement of body parts. Another advantage relates to their size, as more recent advances in sensor miniaturization have enabled the development of small compact microelectromechanical (MEM) sensors [120].

The study of these wearable sensors in an occupational environment has increased in popularity in recent years, with a revision from [121] noticing a surge by the year 2010, which has constantly grown. This tendency can be explained partly due to technological advances that have helped develop smaller, less expensive, and equally precise inertial sensors, with currently the greatest trends aligning into MEM systems[120]. These sensors are mostly placed on specific positions of biomechanical interest for exposure monitoring and further analysis. As such, the most studied body regions are usually the upper arm (more specifically the dominant one) followed by the thoracic area (between the T1 and T2 vertebrae), pelvis, and thigh[121]. The most developed examples have extended to entire body suits of IMU sensors. [122]

Regarding the analysis component of most IMU studies, firstly, it is essential to understand that despite the preferred objective being the analysis in a real-work environment, most of the current studies are still practiced in a controlled laboratory[121]. The analysis of many studies tried to focus on objectively rating the three main characteristics: Intensity, Duration, and Frequency of Activities. These were briefly touched in section 1.1.3, and are the main biomechanical risk factors that describe the ergonomic risk of a specific

human motion. This type of analysis can be made under a relative perspective where work station, workers and tasks are compared against each other, or under an absolute perspective where ergonomic assessment sheets are used to retrieve a risk score of a given activity[123]. Other studies are found to explain the occupational activity by focusing on developing predictive models for the classification of specific actions, such as sitting, crouching, reaching, picking, and picking stuff up [105, 124, 125].

Interestingly, a variety of studies[126–131] explored the possibility of a real-time acquisition to further expand on the utility of wearable IMUs as an intervention technique to mitigate ergonomic pressures. This experience is done by wireless inertial sensing technology connected to cloud-based web services that can compute and communicate the information retrieved by the sensors in near real-time. After that, the worker will be subject to auditory and vibrotactile feedback to adjust his posture instantly. [121]

IMU sensors are an attractive tool to assess the occupational risk because they are low cost, have high accuracy, and are small in size. All this allows them to be integrated into the work environment without significantly affecting the work performance, which is very important as most of the limitations refereeing to direct sensing techniques are usually about their invasiveness, costs, and portability. Moreover, regarding the accuracy of IMU sensors, these have a specificity which makes them very trustworthy, adding to that as they are also a complex system of multiple sensors, there are methods of internal correction of its errors like sensor fusion, which have been developed [132]. The mentioned advantage makes this the chosen system to perform the direct occupational risk evaluation for workers while performing their activities.

Unlike the current trend in studies that use IMU devices in a labor context acquisition, this thesis will not serve to assess the work strain applied at every movement, nor will it try to classify the human motions that are being done. Instead, this work will serve to structure the data and discover patterns of interest with the final objective of facilitating the ergonomic analysis.

3.2 Event Detection Under an Ergonomic Context

Expanding on the previous section 3.1.1, the type of analysis most common for the data of human motion is the classification of specific simple postures of activities, being more related to the field of human activity recognition.

High-level annotation in ergonomic data is still a more scarcely studied problem. The following review will explain the various approaches that tried to identify and annotate events representative of human action. Throughout this section, one of the intended objectives was to find methods that required as minimal interaction with the human user as possible while at the same time focusing on the analyses of time series. With this, it means that it is preferable to search over sensory data instead of video imaging. Because there is a considerable variation within all studies, this review will be presented on a more point-to-point basis.

The work of [133, 134], analysed multidimensional angular joint motion time series data. The objective was to summarize a work cycle into smaller motion segments, each one represented by dynamic models approximations. Interestingly, this work proposed a high-dimensional motion segmentation based on univariate active forgetting segmentation methods. A two-step recursive least square algorithm predicts change points of the system's dynamic behavior. The motion data segments are then represented by parameters derivative from a dynamic model fit. The advantage of this process is that these model parameters are features insensitive to small-time variation, which are very common in this type of data, and because they form a time-invariant representation of each segment's motion. In the end, each segment is compared based on a "kinetic energy-like"measure.

Matrix profile is an extensively explored algorithm, having proven itself as a valuable tool of change point detection in various time series contexts. Unlike the others presented in this section, this method was not explicitly designed towards the processing of human motion. Instead, it was developed to extend the field of all-pair similarity search (also known as similarity join), whose objective is to discover the nearest neighbor for a given collection of data objects. Despite this, it has shown promising potential in human activity data.

The actual definition of the algorithm is relatively simple. The matrix profile is a signal where the value in each point coordinate has the similarity value between the sequence represented by that point and its nearest neighbor sequence. This construction starts by computing the distance of all pairs of subsequences and then taking the minimum value at each location. From here, the analysis of the motif profile can retrieve from the time series various pieces of information, as for instances, maximums and minimums in this signal may represent the presence of anomalies and motifs in the time series. Moreover, the variance can also be used as a measure of complexity in a given time interval. Matrix profile is a helpful tool for motif discovery, anomaly discovery, shapelet discovery, and semantic segmentation. There has also been an important focus concerning optimizing this algorithm, with matrix profile being a very fast algorithm appropriate to handle large data dimensions, partly because of the construction of a z-normalized euclidean distance measure that resorts to fast Fourier transformation and the dot product.

Wang *et al.* [135] proposed a method for the temporal segmentation of human repetitive actions. The data here is retrieved by two sensor sources: optical motion capture system and Microsoft Kinect[136]. With the last one being, in practice, an infrared sensor system, supported by an RGB camera, which differs from the optical motion capture system that, like the previous study, returns motion data, which like the previous studies of [134, 137] will then be translated into time series with where the joint angles are represented. The motion data was initially converted into a generic full-body kinematic model using an unscented Kalman filter. This transformation has the advantage of being a unified representation robust to noise. Then, by performing a primary frequency analysis, where most representative kinematic features are then retrieved. Based on the zero crossing of these retrieved features, the data can then be segmented, followed by

an adaptive k-means clustering to identify which segments are repetitions of each other, allowing for a strong identification.

The work from [138], much like the current thesis, proposes an algorithm that uses the properties of self-similarity for the detection of events in human motion data. Firstly the algorithm proposes a segmentation of the algorithm into primitive segments, where a motion sequence is given as a collection of various poses, with each pose being described by a feature vector. Then creates a nearest neighborhood graph, where similar poses are searched within a radius of r, with the construction of a kd-tree where the similarity measurement is the euclidean distance between the feature vectors. Then, by studying the relationship between SSM and the neighborhood graph, it allows the possibility to remove neighbors that belong to the same connected component as a given frame.

For the segmentation of different activities, this algorithm pursues region growing as a tool to solve this problem, with it considering three types of motions possibly repeated in the matrix: 1) short, intermediate action; 2) structures parallel to the main diagonal; and 3) diagonal part connecting two parallel structures, reflects a transition. Regarding region growing, the general idea is to start growing a squared connected region from seed in the upper left corner of the neighborhood representation matrix. The region is gradually extended to adjacent rows and columns as long as the number of nearest neighbors in the updated region increases. This will be perpetuated until there are no new neighbors found. Then a new region is started from the upper left region of the remaining matrix, where every entry left below the block already found is removed. Another critical factor process is the removal of the main diagonal from both the SSM and the graph. The seeding process is performed once in a forward step as previously described and then in a backward step.

The data is then subdivided to find motion primitives. The basic idea of the algorithm is to use minimal cost warping paths in the matrix whose start and end positions are associated with the start and end positions of the diagonal paths, which were found and examined according to the criteria of symmetry coherence and projection coverage. In the end, the motion primitives are clustered accordingly.

The work of [139] searched to facilitate the analysis of automatic monitoring over activities of daily living. It proposes a method to identify significant changes in the gestures of human motion retrieved from wearable inertial sensors. Firstly it has a step of pre-processing where it applies a bandpass filter. Then it does a hierarchical search for significant changes where it applies a sliding window and retrieves the dominant axis. In the end, it applies an algorithm to spot significant changes.

The dominant axis classifier uses a mathematical formula to identify which axis presents a greater level of difference. It is assumed that the axis with a more significant level of variability is the direction where the motion is more prominent.

The hierarchical search first identifies all instances where the signal transitions from a threshold value, and this threshold value is defined as the average of sensing data in the dominant axis. After detecting these instants of transition, it detects the previous local

maximum and following local minimum for each one of these moment transitions.

After extracting these events, a feature extraction process is applied, followed by a machine-learning-based classification. For the feature calculation, two different sets of features are used: the first one with statistical features which can be reliably calculated like Maximum, Minimum, Mean, Standard Deviation, Root Mean Square or range, and the second with physical features. Physical Features can provide accurate and objective information about the movement quality and smoothness of human motion. In the context of this work, there were used movement time (a measure of the time required over the significant series), peak number (a gradient analysis of the signal over the significant series), and jerk metric (a quality measure of smoothness over the significant series). The selection feature process for each dominant axis is testing the feature vector to determine which one displays a better performance. Regarding the classification process, three machine learning algorithms are tested: Naive Bayes, K-Nearest Neighbors (KNN), and Support Vector Machine (SVM), to determine which of the candidate events have significant relevance. Overall, the KNN and SVM performed better than the Naive Bayes.

3.3 Event Detection with the Self-Similarity Matrix

Self-similarity matrices are a tool largely used in various scientific areas, with most of the found studies for this review being fairly recent from the decades 2000 and 2010. The study of [140] is usually recognized as the first that initiated this type of analysis. This study, which was constructed over previous advances from statistics, described a recurrence plot as a method able to easily display and retrieve information regarding the dynamic behavior of time series[141]. SSMs are a data representation technique that has mostly been applied in two specific types of data: 1) sound, more specifically audio signals, and 2) moving video images. This review will concentrate exclusively on studying sound databases as these were the greatest inspiration for the proposed final algorithm.

3.3.1 Inspiration from Audio Signal Analysis

Music Structure Analysis is a field with a vast literature on the subject of SSM, which makes sense considering that a repetitive/cyclical behavior usually defines this type of data. This can be easily seen in various music genres: with 19th-century European compositions use of one or more motives which are then repeated and transformed; modern popular music has the use of verse/chorus structure and instrument solo for instrumental introduction and coda; as well as "classic jazz"from ”New Orleans” style to ”Be-bop” is based on the repeated exposition and improvisation around a specific theme[142]. An additional convenience of music data structure is that its boundaries are usually well delineated by variations in specific features, unlike continuous speech data. The favoring of one type of audio signal to another relates to the difficulty of considering what the human comprehension of speech an implicit requirement in automatic speech recognition is.

Words themselves are not necessarily sequences if there is not considered some grammatical structure to the data. A simple example would be the phrase "that's Steven". If we were to apply a simple algorithm of self-similarity without considerations about the grammar proprieties, the signal would probably be segmented into "that's-S"and "teven"[143].

The first development made in this context came from [144], where SSM was proposed as a tool for better visualization of the time structures of music datasets, with possible applications on content based analysis and segmentation[145].

Most of the audio analysis goes through a relatively identical sequence of steps towards a final result. From where there is usually a step for

1. Feature Representation, which may also be accompanied by a pre-procesing step over the raw data;

2. The Calculation of the SSM from the feature data obtained in the previous point, by using a given distance metric;

3. Analysis process, which varies a lot and usually depends on the studies' objective.

3.3.2 Feature Representation

This first subsection describes a standard process of parameterization of the raw data time series. The actual processing can either be by tapering the raw data with a Hamming window or by making a previous segmentation of the entire dataset and only then applying some transformation function in the resulting subsegments. This segmentation is usually done in overlapping steps as it can optimize the quality of the data signal [146].

The choice of a feature needs to consider how the signal is intended to be divided. In the context of music data, the structural boundaries are usually defined by variations in the proprieties of 1) timbre, 2) tonality, or the 3) rhythm of sound [147]. Specific features are extracted to describe each of these properties, such as Mel-frequency Cepstral Coefficients (MFCC) and Chroma features (or Chromagram).

Still adding to the theme of feature representation, some other methods used include the 1) direct transformation of each window into the frequency domain using the fast Fourier transform (FFT) [143]. 2) spectral envelope estimation, by means of Cepstrum Coefficients[142], 3) short-time Fourier transform (STFT) [146] 4) monophonic pitch estimation, Wakefield chroma (spectral) representation, and polyphonic transcription followed by harmonic analysis[148], 5) 13 Mel Frequency Cepstral Coefficient, 12 spectral contrast coefficients, based on [14], Pitch Class Profile Coefficients[149]

Other interesting features are used to describe dynamic properties[150, 151]. These have the advantage of properly highlighting diagonals or blocks in the structure, depending on whether it does a long or short duration modeling. Firstly the audio signal is passed by a bank o Mel filters, and then each output of each filter is analyzed by STFT, with a variation in the windows' length, defining a different long or short duration modeling.

The features which are maintained are two coefficients: i) the angular frequency returned from the STFT analysis and ii) an identifier of what was the Mel filter used.

3.3.2.1 Post Processing of the Feature Vector

The calculation of the SSM did not necessarily use a single one of these features in isolation. Instead, the signal was usually represented by multiple types of features simultaneously. Each point coordinate was represented by vectors developed from multiple types of features, which were then compared with themselves on the calculation of a single resulting SSM. Interestingly, some examples[149, 152, 153] diverged from this process and instead calculated various similarity matrices from different features, normalized them, and later combined all matrices into a single SSM. With the combination being a simple sum of matrices. This method allows for the introduction of different weights for the different features. Even though the specific cases of the two refereed examples were given the same weight for all matrices.

When multiple features are used, these tend to be normalized to mean zero and variance one.[153]. Another processing step used after retrieving the features is a principal component analysis (PCA) reduction[149, 154, 155].

Other separate variations make a previous segmentation and classification of the data, in this context, the SSM becomes closer to the recurrence plot, as every position is binary depending if there is a pair that is the same or not[142]. An example of such could be [142] which, before calculating the SSM matrix, transforms the data into a class-based representation, which it defines as texture score representation. For this, it has a first step, which, much like others, start by retrieving a feature representation, in which it estimates the spectral envelope using spectrum coefficients (focusing as such on the timbre evolution of the music), but after that, the frames are classified with a hidden Markov model (HMM). More specifically, for learning, it used a Baum_Welsh algorithm, with each state of the HMM accounting for a specific texture and through a Viterbi decoding, labeled each frame with a corresponding texture.

3.3.3 SSM Calculation

Advancing to the second point of the actual SSM construction. As seen by the definition of section 2.3, an important requirement for the calculation of SSM is the definition of the similarity measurement. The distances used in these studies were:

1. Euclidean distance
2. Cosine between angles
3. Kullback Leibler distance.

Euclidean distance is a simple and usually successful option, which is mathematically described by the expression of

$$s(x_i, x_j) = \sqrt{\sum_{n=1}^{N} (x_i(n) - x_j(n))^2}$$

With N being the dimension of each x point vector. If this point vector had only one dimension, the formula could be reduced into

$$s(x_i, x_j) = \mid x_i - x_j \mid$$

.

The cosine of the angle between the vectors x_i, x_j is defined by the expression

$$s(x_i, x_j) = \frac{x_i \cdot x_j}{\|x_i\| \times \|x_j\|} = \frac{\sum_{n=1}^{N} (x_i(n) x_j(n))}{\sqrt{\sum_{n=1}^{N} (x_i(n)^2)} \sqrt{\sum_{n=1}^{N} (x_j(n)^2)}}$$

with N being the dimension of each x point vector. If this point vector had only one dimension, the formula can be reduced into the second term of the previous expression. Cosine distance has the advantageous property of returning a large value of similarity, even if the feature arrays are smaller in magnitude. A variation of this similarity measurement uses the dot product i.e. the convolution of the feature vectors with each others

$$s(x_i, x_j) = x_i \cdot x_j$$

. Despite not being normalized, it is also refereed multiple times as a reliable possibility, with this being the measurement selected for this work as it will be further presented.

Another interesting variation proposed by [153] proposes a variation on this formula with a exponential weighting

$$s(x_i, x_j) = \exp\left(\frac{x_i \cdot x_j}{\|x_i\| \times \|x_j\|} - 1 \right)$$

with this equation limiting the similarity range to $[0, 1]$ and creating an higher contrast at higher similarity values and lowering the contrast in lower similarity values. The definition of Kullback Leibler distance is

$$D_{KL}(P \parallel Q) = -\sum_{n=1}^{N} P(n) \log \frac{Q(n)}{P(n)}$$

So the similarity measurement of Kullback Leible distance becomes

$$s(x_i, x_j) = \exp(-D_{KL}(\mathbb{G}(\mu_i, \Sigma_i) \parallel \mathbb{G}(\mu_j, \Sigma_j)) - D_{KL}(\mathbb{G}(\mu_j, \Sigma_j) \parallel \mathbb{G}(\mu_i, \Sigma_i)))$$

, where $\mathbb{G}(\mu_j, \Sigma_j)$ is the multivariate normal density function, with μ_j and Σ_j being the average and the covariance matrix of the vector x_j, with the respective also being true for the vector x_i. [146]

In the case of comparison between class-based representations of time series like the case of [142], then the SSM calculation is much simpler as it is only a binary correlation, comparable to comparing "strings"instead of numbers.

$$s(x_i, x_j) = \begin{cases} 1, & \text{if } x_i = x_j \\ 0, & \text{otherwise} \end{cases}$$

This tends to be a method that still requires an optimization step to get an acceptable final result. For this, it is done a process of discarding all trivial occurrences by substituting squared moments by diagonals and then manages nontrivial patterns so that all occurrences are coherently linked with each other.

After obtaining the matrix, some studies [154, 155] applied a two-dimension median filter over the SSM.

Another interesting variation proposed by [149], proposes an higher degree of matrices that makes the final result more consistent within itself, for this, it defines a second order matrix as:

$$S_2(t_x, t_y) = \int_{t_z} S(t_x, t_z) S(t_z, t_y) dt_z$$

and a third order matrix as:

$$S_3(t_x, t_y) = \int_{t_{z1}} \int_{t_{z2}} S(t_x, t_{z1}) S(t_{z1}, t_{z2}) S(t_{z2}, t_y) dt_{z2} dt_{z1}$$

These matrices have the advantage of reinforcing the diagonals corresponding to common repetitions and reducing the background noise

3.3.3.1 SSM transformation

Some studies also mentioned another important step, applied over the SSM before any analyses process. Some studies refer the transformation of the SSM into a lag matrix, which is represented in a "slanted"domain $L(i, l)$, where $l = i - j$. Also called by self-similarity lag matrix (SSLM). Computing the SSLM for small non-negative values of l has the advantage of reducing computation and storage requirements[143]. A good example is the case of [156], wherein the search for a chorus of the music only considered lags greater than one-tenth and lesser than three-fourths the length of the song.

$$L : L(t_i, lag_{ij}) = S(t_i, t_i - t_j)$$

This matrix has the property of being equivalent to the SSM with the diagonals being converted into vertical lines. This matrix is usually calculated by performing an average moving filtering along the diagonals of the SSM. [151, 156] Another way is the application of a horizontal high-pass filter and a low vertical pass-filter[149].Another advantage that is explored in these studies is that the matrix provides a much simpler way to retrieve the size of the subsequence.

Another method of analysis, involves the calculation of a similarity function (average similarity)[157], which is the sum of the SSM between a segment and the entire work, normalized by the segment length. In the work of [157], it is also refereed the possibility of applying a weighted function to optimize the final result.

3.3.4 Event Detection with the SSM

Finally, the actual methods used to analyse the SSM is where there is greatest level of variations within all the studies, with this depending mostly on the intended objective.

3.3.4.1 Change Point Detection

Regarding methods precisely of change point detection, in a much consistent way, the focus was the detection of cross check-board moments along the diagonal. As such, the method mostly consisted on the Smoothed "Checkerboard"Kernel correlation along the SSM main diagonal followed by a peak finding analysis in the resulting data[143, 146, 151, 153]. This last processing step uses a threshold value and the search for peaks in the signal which surpass this threshold. In [153], it was used an adaptive threshold algorithm [158].

In the work from [150, 151] the similarity function is calculated and then it applies a simple threshold operation to identify which moments are significantly dissimilar within each other. Interesting variations also include organizing the indexed points in a binary tree structure constructed by ranking all the index points by novelty score

In a recent study from [154], to capture the boundaries in the SSM, it was applied a sliding window that calculated the sum of the standard deviation of each vector column that was included inside the window. Then, it was applied a peak finding function on the resulting data. The clustering is performed from the highest similarity segments to the lowest similarity segments, with each segment being clustered with their highest similarity pair. Moreover each neighbouring segment belonging to the same cluster will be merged. In [155] added a new proposal and instead of analysing the similarity value of the SSM created from multi mel frequency cepstrual coefficients (MFCC) features, retrieved the chromagram features of each segment and then calculate the cross correlation to assert their similarity and then design clusters, using these values.

3.3.5 Similarity base detection

Regarding Similarity-based detection methods, the objective was the detection of highly intensive diagonals, and for these, there is a more extensive variety of methods proposed.

In order to identify the diagonals, several methods were proposed. As already refereed, the lag matrix can be a helpful tool in the detection of the more prominent diagonals. Bartsch *et al.* [156] started by filtering over the length of diagonal of interest, another

advantage provided by SSLM. As the interest of this work was the detection of a single motif that could define a music thumbnail (possibly the chorus), it was only necessary to detect the time position of maximum similarity of the SSLM. Conversely, [151] had a much more deepened analysis of the SSLM, where there was first performed a line detection, and from those, sequence representations were derived. After that, it is found which segments are repetitions of which other segments, by a proposed algorithm of segment connection, where each is clustered one by one. Then, for the most prominent motif, it simply had to select the segment that had the most time shared with all the other segments.

An alternative to this is that instead of searching for diagonals, the algorithm would search for instances of significant similarity. It is defined as a similarity function, i.e, the sum of the SSM between a segment and the entire work, normalized by the segment length. A possibility also referred was applying a weighted function, optimizing the final result. Then, it is simply necessary to detect the positions of maximum similarity value.

3.3.6 Summarization

Still, most work with SSM does not only focus on detecting change point events, anomalies, or motifs. This is a consequence of the real objective of summarizing the time series. In other words, in most cases of the studies presented in this review, the objective is actually to summarize the data properly. Summarizing procedures can be divided into

1. a representation of *states*

2. a representation of *sequences*

The difference is that a representation by *sequence(s)* tries to identify which subsequences have a recurrent behavior in the sound data. Usually being a theme more related to similarity-based detection. In the case of most of the literature on this subject, after identifying all subsequences of interest, one of them is singled out as a thumbnail to describe the entire time serie. Peeters *et al.* [149, 151] is described a new approach to identify "mother"subsegments and repetition times in such a way that it represents all the segments detected in the matrix using the smallest number of possible sequences.

On the other hand, *state* representation has an entirely different approach and tries to segment the data into multiple subsequences, which usually do not overlap and may be included in specific classes. In this perspective, the data can be seen as a progression of states, which can sometimes repeat themselves. The *state* representation has the advantage of return much summary information: class transitions, an example of each audio state, the class succession order, or an audio example of the most important class.[150, 151].

Chapter 4

Database Description

In the interest of testing how valid was the algorithm which will be purposed in [**cha:methodology**], multiple database sources were tested in the present work. Thus, the purpose of this chapter will be to explain these various databases. Firstly, there will be presented multiple databases of human motion retrieved from several online sources that have a simpler context and are used to test and validate the algorithm. Then, there will be presented the main database which was retrieved from a real working context from the manufacturing industry.

4.1 *HAR* database

For this thesis, there were allowed access to inertial acquisitions made in the context of human activity recognition. The database used was obtained on the scope of the *Arthrokinemat* project whose main objective was the development of a learning adaptive sensor-based measurement system to prevent osteoarthritis[159]. More specifically, the recording was made for the work of [160], which introduces a human activity recognition system able to recognize between a list of several daily activities.

A set of multiple Biosensors were used, with various internal characteristics. Beginning with two *8 channel PLUX hubs* a device which allows for the wireless acquisition of biosensors via Bluetooth, with them being part of the *biosignals plux Research Kits*. From both *plux hubs* there were used two 3-axial accelerometer (Acc) sensors, 4 sets of EMG sensors and an electrogoniometer, with more details about these sensors being displayed in Table 4.1.

Adding to these sensors there were also used 4 other types of biosensors: one airborne microphone, one piezoelectric microphone, two 3-axial gyroscopes and one force sensor. From here, 18 activities were performed recursively, with all of these being listed in Table

4.2

Table 4.1: *Plux Hub* sensors descriptions, for the HAR databases[160]. The sensors are divided in half by a horizontal dash line with the first group of sensors using a single *Plux Hub* device and the other sensors using the second device. The information provided by each sensor include also the positioning of the sensors as well as it's sampling frequency.

Sensor	**Position / Muscle**	**Sampling Frequency**
3 axial Upper Accelerometer	Thigh, proximal ventral	100 Hz
3 axial Lower Accelerometer	Shank, distal ventral	100 Hz
Eletrogoniometer	Knee of the right leg, lateral	100 Hz
EMG1	Musculus vastus medialis	1000 Hz
EMG2	Musculus tibialis anterior	1000 Hz
EMG3	Musculus biceps femoris	1000 Hz
EMG4	Musculus gastrocnemius	1000 Hz

Table 4.2: Activities description, of the HAR database[160]. For each activity is displayed the number of repetitive occurrences each activity is performed as well as the total time length and the minimum and maximum length of each occurrence.

Activity	**Occurrence**	**Total Length**	**Min Length**	**Max Length**
sit	47	123.75s	0.86s	4.69s
stand	46	127.70s	1.36s	4.73s
sit-to-stand	45	30.94s	0.15s	1.30s
stand-to-sit	53	72.90s	0.56s	3.10s
stair-up	55	190.45s	1.59s	4.93s
stair-down	57	181.96s	1.37s	4.86s
walk	220	554.07s	1.18s	4.78s
curve-left-step	57	143.09s	1.10	3.87s
curve-left-spin	46	109.15s	1.25s	3.39s
curve-right-step	51	67.51s	0.54s	3.19s
curve-right-spin	48	41.50s	0.28s	1.88s
run	97	151.27s	0.64s	2.74s
v-cut-left	53	43.76	0.29s	2.08s
v-cut-right	55	61.75s	0.35s	2.37s
lateral-shuffle-left	53	97.54s	0.73s	4.11s
lateral-shuffle-right	52	90.42s	0.75s	3.98s
jump-one-leg	59	61.36s	0.33s	2.85s
jump-two-leg	63	63.40s	0.51s	1.63s
Total	**1157**	**2212.52s**	**0.15s**	**4.93s**

The activities **sit-to-stand** and **stand-to-sit** were moments of transition between the activities stand and sit. Moreover the activities **curve-left** and **curve-right** were interposed between the activity of **walking**. **curve-left/right** are also divided into **step** or **spin**, depending if it is a big 90°turn with several walking steps or a fast 90°turn of the entire body like a parade command, respectively.

Another set of activities that may also require further explanation are **lateral-shuffle-left/right**, a motion usually done in sports that describes the subject's lateral movement of the left/right foot, with the other foot following along and continuing the shuffling in the same direction. **V-cut-left/right** means that the subject changes his direction by 90°at jogging speed. The remaining activities are self explanatory.

This database was used to study how the algorithm could be used to detect periodic events with different levels of motion cycle complexity.

4.2 *UCI* Database

This database was available in the UCI Machine Learning Repository[161], an online service with more than 500 databases for the machine learning community, as the "Smartphone-Based Recognition of Human Activities and Postural Transitions Data Set"[162, 163].

The database was collected from a 3-axial Acc and 3-axial gyroscope (Gyro) embedded in the smartphone, with 30 volunteers performing a set of different activities, while the device sits in their waist during the execution of the experiments. The smartphone used during the experiment was a Samsung Galaxy S II with a 50Hz acquisition rate and 30 volunteer subjects within the ages of 19 and 48 years old participated. The experiment was video recorded so that the data were manually labelled, with the protocol describing 6 activities: 3 static postures (standing, sitting, lying) and 3 dynamic activities (walking, walking downstairs and walking upstairs) for the subjects to perform, as well as all the possible transitions between the different postures (stand-to-sit, sit-to-stand, sit-to-lie, lie-to-sit, stand-to-lie, and lie-to-stand).

This database was used to study the ability of the algorithm to perform change point event detection along the time series.

4.3 *Industrial* Database

This database has greater importance in comparison with the previous ones, as it will serve to test if this methodology is indeed valid for processing data acquired from a work environment. With this said, two main requirements need to be presented. Firstly, the data should have been retrieved in a real-working setting (i.e. not in a laboratory controlled environment), and secondly, it had to be the result of a direct sensor acquisition system that could provide information about the motion of the various workers.

The data acquisition was made in a previous thesis project from [164], whose main objective was to calculate the ergonomic risk through direct measuring data. The technological setting was provided by a sensing framework named Internet of Things in Package (IoTiP), designed and provided by Fraunhofer AICOS. IoTiP is a system that intends to combine hardware, firmware and software components to promote the field "Internet of Things"[165]. For this work, the technology setting consisted on 4 9-DoF IMU sensors

(composed internally by a triaxial Acc, triaxial Gyro and triaxial magnetometer (Mag)) and an Android wireless communication system. The last one was made through an application called Recorder, also developed by Fraunhofer AICOS.[164].

The acquisition was made in a Volkswagen Industrial assembly line where 12 manufacturing workers performed their work tasks while having attached 4 IMU sensors in their bodies. During the acquisition, the subjects performed various tasks in multiple workstations. Relevantly to this thesis, the database comprehends three different workstations from *Bodyshop assembly line*, a section where cars' doors were assembled: 1) Liftgate workstation, where back doors are mounted; 2) Fender workstation, involving front door tasks and 3) Doors workstation, which demanded tasks on the front doors and in the cars' hood[164]. The acquisitions made involved a total of 6 Opr with each one performing at least 2 different Wkst. The various acquisitions were simultaneously filmed and to synchronize the ground manual annotations of the data, in the beginning and end of the acquisition the subjects were asked to stay, firstly, in a neutral anatomic position and then perform a T pose (calibration position) as represented in Figure 4.1. There were also registered some details regarding the anthropomtetric characteristics of the various subjects, as displayed in Table A.1.

The mentioned study was centered on the ergonomic assessment of the dominant arm. For this reason, the IMUs were attached in: 1) the posterior side of the hand, 2) posterior side of the forearm and 3) posterior side of the arm and a final one 4) placed in the anterior side of the thorax area. All of the devices were attached with elastic bands, in such a way that all had their Y-axis pointed up while in a neutral anatomical position, as illustrated in Figure 4.2. The last device is unique as in its original work it allowed an assessment of the movement of the arm in relation to the entire body. Moreover, it is also different from the other since its IMU sensor was incorporated in a smartphone. Smartphones can indeed work as IMU devices, since these can sense the acceleration, angular momentum and magnetic field. Each one of these 4 IMU devices had incorporated within them 3 triaxial sensors (Acc, Gyro, Mag), with each producing a multivariate time series of 3 dimensions correspondent to the 3 coordinate orientations of x, y and z.

The data is hierarchically divided, as showed in 4.3. The Database is firstly divided into 12 groups corresponding to 6 subjects with each working at two different workstations; followed by 4 other databases: Hand, Forearm, Arm and Torso; and then with all of these databases being further divided into 3 other Time Series: Accelerometer, Gyroscope. These time series are multivariate with 3 dimensions representing the 3 orientations of x, y and z of the sensors (Figure 4.3). The first level represents the various subject acquisitions made, the second level corresponds to the data retrieved from different IMU devices and the third level represents the data retrieved from different sensor types. Considering this grouping method is relevant due to the data from different databases having different time properties as it will be seen in Section 5.1.

This database will help in studying the application of the algorithm to detect (1) Working Periods, (2) Periodic Working Cycles and (3) Search by example.

Figure 4.1: T pose for synchronization of the various IMU devices.

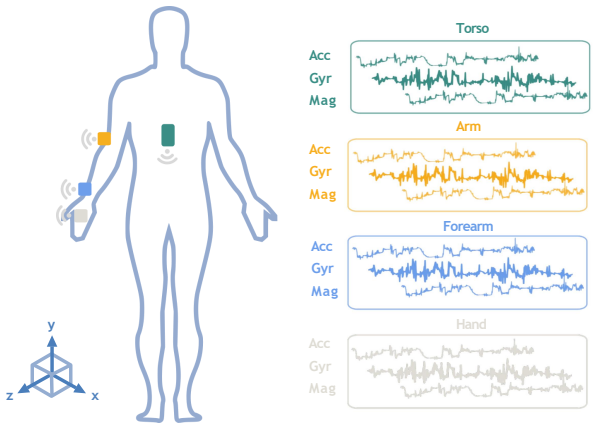


Figure 4.2: Schematic placements of IMU sensors over the acquisition of *Industrial* Database.

Some important notes:

1. Despite the actual acquisition having three sensors Acc, Gyro and Mag, only the Acc and Gyro sensors were considered for this study as this had the best behaviour, and the Mag was acting in an erratic manner.

2. The two workstation of Opr2 were made during the same acquisition, resulting in a single time serie, where the subject performed two different types of active work motions.

3. In Opr2 Wkst1&2 the torso was not considered, due to malfunction.

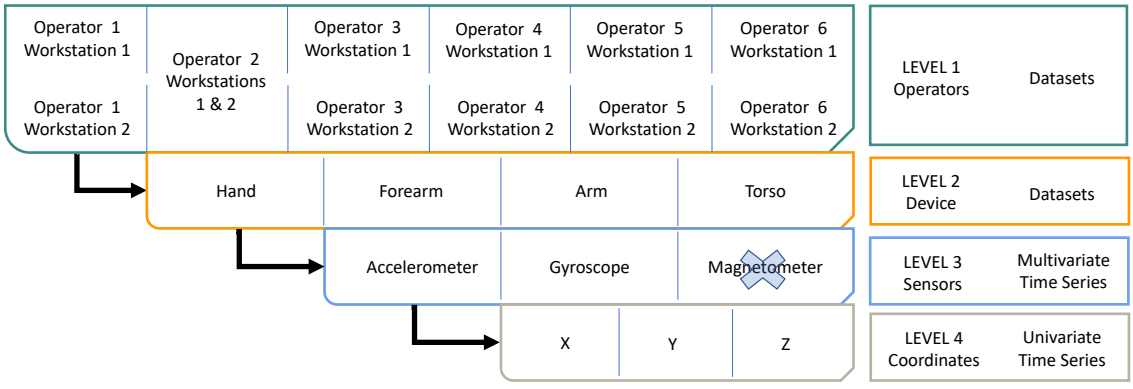


Figure 4.3: Diagram representation of the abstract levels of the *Industrial* Database, with the highest level being in the top and each cell of each level being subsequently divided into the lower level cells. The 2 right columns give information regarding the level identity and the type of data as described in Section 2.1.

Chapter 5

Methodology

Data mining and event detection techniques over time series usually have a standard sequence of processing steps, which overall remain unchanged. This chapter will describe the proposed methodology, and it will be divided into various sequential steps of processing.

1. Pre-processing
2. Representation
3. SSM calculation
4. Event detection

The first 3 methods are somewhat self-explanatory. The last point differentiates itself due to being divided into 3 different subsections. The reason for this is because this thesis proposed to find multiple events: (1) Work period transition, (2) Work cycle transitions, and (3) Sub-segment example search. As it has been extensively discussed, events are somewhat subjective concepts that require a specific analysis depending on the intended event to be detected. As such, each section will describe three different methodologies of analysis of the SSM depending on the intended event.

5.1 Pre-processing

Pre-processing is an essential step in any data analysis. This section will focus on two points: 1) Temporal Synchronisation and Alignment, as it is important to ensure that the time series retrieved from the sensors are in shape suitable to be processed; and 2) Noise reduction, as inertial acquisitions are very prone towards various mistakes.

5.1.1 Temporal Synchronisation and Alignment

All IMU sensors suffer from integration drift, where small errors appear during the measurement of the sensors, which will progressively be integrated into larger errors in the measurements [166]. To tackle this problem, before any type of analysis being done, it is necessary to guaranty that:

1. all time series have the same acquisition rate, as well as the same time length;

2. all the point observations are temporally aligned.

While the *Industrial* database required to take both these considerations into account, the *HAR* and *UCI* databases were already aligned, and they just required the synchronization over the different sampling acquisition of each of their dimensions. As such, the following section will focus only on the analysis of *Industrial* database as this was the most demanding. However, the *HAR* and *UCI* databases also passed by the same exact process of synchronization, despite not requiring the alignment phases.

Regarding the first task, the main obstacle relates to the fact that all time series from a subject's database have a different and complex time form. As the sensors are not synchronized, the resulting data develops different time lengths and acquisition rates between different time series. This is displayed in Table A.2, where there are registered important characteristics about the temporal information of each time series. Focusing on a single example of the "Opr1 Wkst1", we can see that the variability of the average acquisition rate depends on the sensors used, with it even being different within the same devices. In this specific case, the data acquisitions variate between 50 samples/s, 100 samples/s, or 200 samples/s. Moreover, there are also observed slight differences in the time lengths and on the initial time offset values when comparing different devices. These are all scrutinized in the appendix A.3. But considering the specific example of the "Opr1 Wkst1"time serie sample, the time lengths of the signals retrieved from the hand, forearm, arm, and torso IMUs were ≈ 1593, ≈ 1552, ≈ 1583 and ≈ 1602, respectively. Meanwhile, the time offsets also had some considerably large divergences, being ≈ 1994, ≈ 1999, ≈ 2004, ≈ 0, hand, forearm, arm, and torso.

Notwithstanding, it is also apparent that the acquisition rate is not constant over the same time series, as all of the signals acquire some disturbances somewhere in time, with most of them having a significant reduction in their mean value. This variation in the acquisition rate can be more easily observed in the example of Figure 5.1, where during the time intervals of the close-ups 1 and 2 there are various disturbances in the rate of acquisition, leading to a small but significant delay on the sampling process.

To solve these problems of (de)synchronization between different and alongside the same time series, the main objective of this step will be to design an "artificial"synchronized time axis, which could be shared among all IMU devices. For the sake of simplicity, this time vector will be designated as unifying time (UT). The first characteristic of UT vector

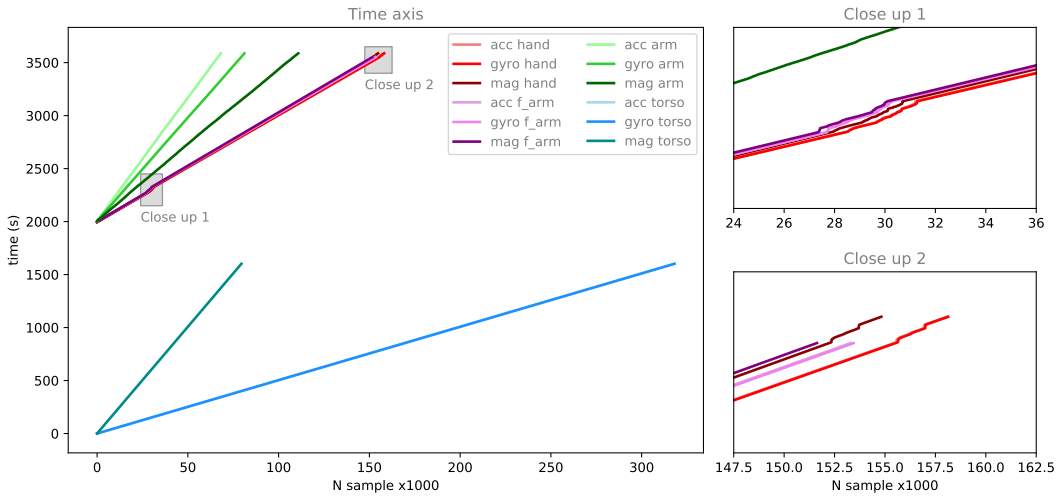


Figure 5.1: Plot of the time in function of the number of time samples, for each time serie from the *Industrial* database. On the left, there are presented all the times series in a color-coded manner. On the right, there are presented amplified plots of the squared positions signaled in the left plot.

is that it will have a constant acquisition rate Fs, which will be the same for all time series. The selection of Fs will be by selecting the maximum value between the considered sampling rates. Returning to the example of "glsOpr1 Wkst1"between the 3 values of 50 samples/s, 100 samples/s, or 200 samples/s, Fs would be 200N/s. This is because if another smaller value was chosen, it would be possible to register some unwanted aliasing behavior in the time series. Since the torso device is a Smartphone with a higher acquisition rate, it usually is this device that defines the value of Fs.

With the definition of Fs settled, another consideration relates with the temporal length of UT. As the intent is to lose as little information as possible without corrupting the data, the time axis will be the same length as the shortest time series (t_{fmin}). To give a specific example, if it is considered the subject "Opr1 Wkst1"as provided by A.2, the new time axis would have a time length of 1552.36 s, as the shortest time series are the Acc and Gyro sensors from the Forearm device.

Regarding the second step of temporal alignment, the main problem is related to the different instants of acquisition in which each IMU sensor starts acquiring. This means that even if the first step is solved, the events occurring at a specific time in a sensor do not necessarily occur in the remaining other sensors. To answer this problem, a specific calibration event performed during the acquisition phase is registered in all time series. For this, it is used the T pose for calibration described in Section 4.3, as it is easily identifiable with the naked eye in the data signals.

Considering four example devices as represented in schematic of the Figure 5.2, with 4 time vectors represented by tD_1, tD_2, tD_3, tD_4. Then the first step will be to manually identify and annotate the time instants where an event occurs (σ_1, σ_2, σ_3, σ_4). Then to simplify the coding process, it was identified which of the devices had the event starting

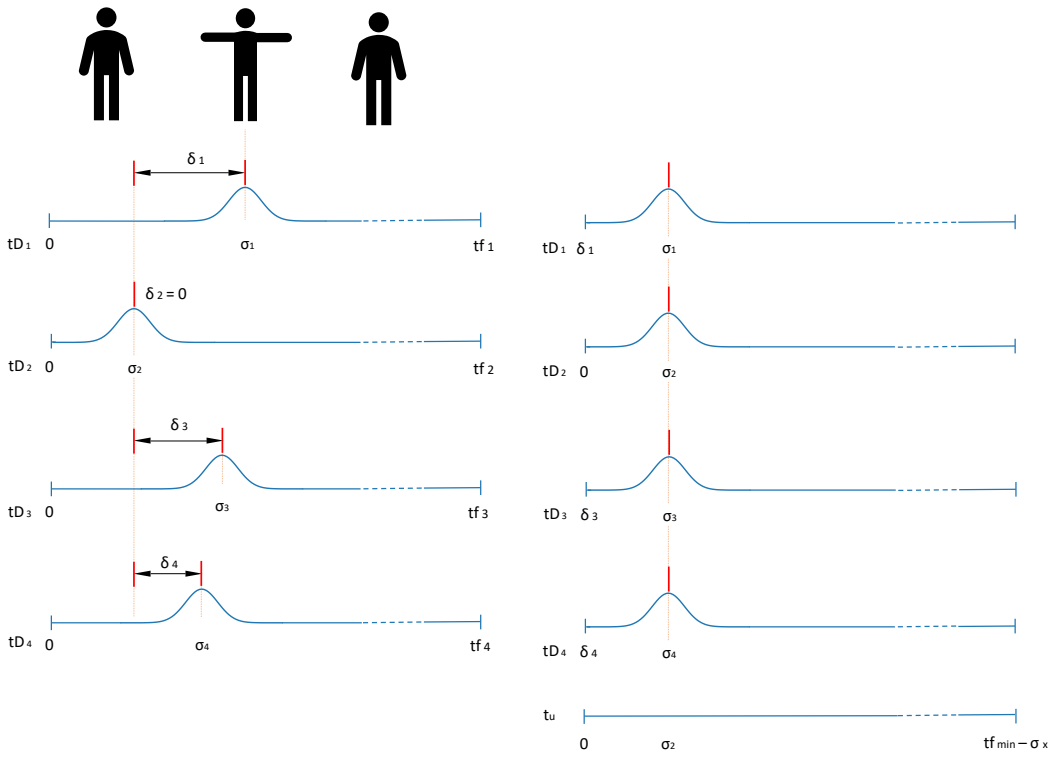


Figure 5.2: Time Alignment Schematic

first $\sigma_{min} = min(\sigma_1, \sigma_2, \sigma_3, \sigma_4)$. With this information, the time differences between the σ_{min} point and the other ones, could be calculated.

$$\delta_j = \sigma_j - \sigma_{min}$$

, with $j = [1, 2, 3, 4]$. Then the alignment of the various time vector devices is done by subtracting these δ values over the respective devices, as represented in Figure 5.2.

$$new\_tD_j = tD_j - \sigma_j$$

, for $j = [1, 2, 3, 4]$. Now with this consideration in mind, the final time axis size will also need to take into consideration this small delay when defining its time length, as such, the time length of UT will be of

$$t_{fmin} - \sigma_x$$

, with t_{fmin} being the time length, and σ_x the σ value (all of them are further described in appendix A.2) of the shortest time series.

After following these two steps, with the various time axis devices aligned and the definition of UT, there is a second cycle pass along the data, where all the time series with the aligned time axis of new_tD will be interpolated into UT. This step will fix the misalignment caused by sporadic variations in the rate of acquisition as previously described.

All this process which was described during this section, is further summarized in Algorithm 1, in a generic code syntax. This is composed of 2 cycle passes along the entire database, represented by the colors pink and blue. During the first pass, the information necessary to calculate UT is retrieved. Then in the yellow section UT is calculated, being then followed by a second pass where the interpolation of the time series is indeed performed.

Algorithm 1: Time Synchronization and Alignment

```
int F_S = 200 ;
database D, time series database of a pre selected subject acquisition ;
dict events_pad, dictionary with the time deviations of each sensor;
database hand, forearm, arm, torso ← retrieve_devices(D) ;
float last_t = 0 ;
tseries synchTS;
foreach device ∈ [hand, forearm, arm, torso] do
    tseries acc, gyro, mag ← retrieve_sensors(device) ;
    array t_acc, t_gyro, t_mag ← retrieve_time (acc, gyro, mag) ;
    float d_last_t ← min( t_acc[-1], t_gyro[-1], t_mag[-1]);
    if last_t ≤ d_last_t then
        last_t = d_last_t;
        last_t_pos = device;
    end
end
f = last_t - events_pad[last_t_pos] ;
time ← arange (init=0, finit = f, sep = 1/ F_S) ;
foreach device ∈ [hand, forearm, arm, torso] do
    tseries acc, gyro, mag ← retrieve_sensors(device) ;
    array t_acc, t_gyro, t_mag ← retrieve_time (acc, gyro, mag) ;
    tseries acc_synch ← interpolate (time+events_pad[device], t_acc, acc) ;
    tseries gyro_synch ← interpolate (time+events_pad[device], t_gyro, gyro) ;
    tseries mag_synch ← interpolate (time+events_pad[device], t_mag, mag) ;
    tseries synchTS ← addTS (acc_synch, gyro_synch, mag_synch);
end
return time, synchTS
```

5.1.2 Filtering

The removal of noise from the signal is a common process that is preferably done before the analysis process. Noisy related problems are slightly intensified in contexts outside of a controlled laboratory environment [167, 168], as it is the case for the *Industrial* database, with portable devices being susceptible to artifacts associated with increased movement and, in the case of IMU sensors, to small drifts in the placement of the devices.

Regarding this problem, considering the context of acquisitions made with both Acc and Gyro, specifically for the context of HAR, the common solution is the application of filtering mechanisms. The most used methods would include low pass Butterworth,

Kalman and moving average filters[167]. With this said, the method chosen to be applied in this work was a second-order low pass Butterworth filter of 40 Hz. The choice of this cut frequency was because experimental studies[169] have already proved that human physical activity could be well represented by a frequency up to 20Hz, and according to the Nyquist theorem, a signal must be sampled at more than twice the highest frequency component of the signal.

5.2 Representation

In this section, the database is transformed, before its actual analysis, into a set of representative features. A standard process of time series analysis and machine learning, already introduced in Section 2.1.

For this work, the data was tapered with a window of fixed length over the time series. This window function applied a feature retrieval process to obtain a time series representation $\overline{X}$. The tapering of the window was applied over the k dimensions of the multivariate time series X and it implemented a variety of feature operations alongside the data, all agglomerated inside function F. This process will change the rate of acquisition (new_F_s), as well as the volume and dimension of the multivariate time series data ($k*$, $n*$).

The new point coordinates of the samples of the time series representation will be defined by $k*$ dimensions, which is equal to the product of k and the number of features used. Moreover, the values will describe the application of the F function along a window centered in this point coordinate. This is more easily seen in the schematic figure 5.3, where a window employed in X during the time interval between Tpi and Tpf generates a point coordinate in the time instant of $(Tpi + Tpf)/2$, in $\overline{X}$. The passing of the window will be defined by the length of this window and the overlapping function, producing a new multivariate time series with an acquisition rate defined by

$$new\_Fs = \frac{Fs}{(Wind_{len}(1 - Overlap_{frac}))}$$

, with $Wind_{len}$ representing the time length of sliding window and $Overlap_{frac}$ the fraction of the window which is overlapped, i.e. the fraction between $Overlap_{len}$ and $Wind_{len}$.

Regarding the margins of the data, much like is done in image processing, the data is replied and mirrored in the borders, so it does not lose relevant information. This reveals another important characteristic of this process. It requires a set of parameters to be previously defined. For this step, it was identified three parameters required for the definition of this function:

- Feature set Selection
- Window time length $Wind_{len}$
- Overaping coefficient $Overlap_{frac}$

The relevance of each one of these parameters will be further described in the following sections. After this, a final section will describe the actual coding methodology which will be applied.

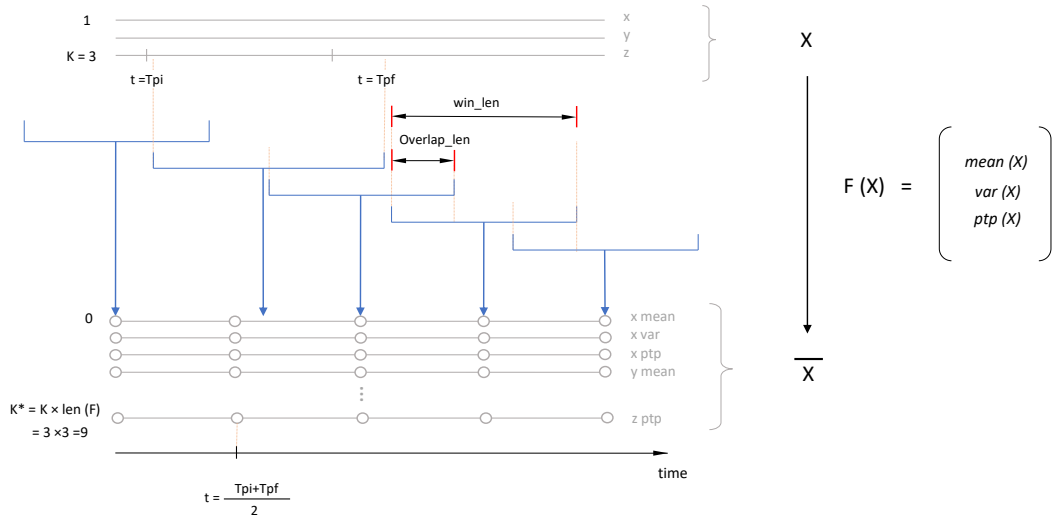


Figure 5.3: Sliding window Schematic. Representation where a time series X displayed in the upper corner with three dimensions (x, y, and z) transformed into the representation $\overline{X}$ with 9 dimensions. These 9 dimension are the result of applying the feature operations $mean(X)$ $var(X)$ and $peak\_to\_peak\_dist(X)$ in each x, y and z axis.

5.2.1 Feature Retrieval

The selection of features is of extreme importance as, depending on what is the dynamic behavior of the data, it may highlight significant changes in the signal and smooth the information for easier processing, or it may have the opposite effect and add significant noise to the resulting signal. The main objective in this step is to define a set of features, such that, when applied to the process previously described, they return a different value vector for significantly different time instants.

A good example is expressed in the following figure 5.4 images, that considers 3 simple but very different biosignals (Acc, EMG and ECG). In this context, the red and green vertical lines represent the various relevant events that would be interesting to be detected. From here, the same process of representation was applied for the three signals in the same way, with feature function applying mean, variance and peak to peak distance operations. Through this example, it can be easily observed that the feature selection needs to be adapted towards the signal and to events being analyzed. For example, in the case of Acc, in general, all features display good behavior, however **peak to peak distance** can introduce some noise as it is susceptible to oscillations in the signal displaying some unwanted peaks. In the case of EMG, the **mean** feature is visibly noisy, developing a random behavior along with the contractions. Regarding the final ECG signal, despite

all features being useful in the detection of the QRS complex, the **variance** seems to be especially bad for the detection of the T wave.

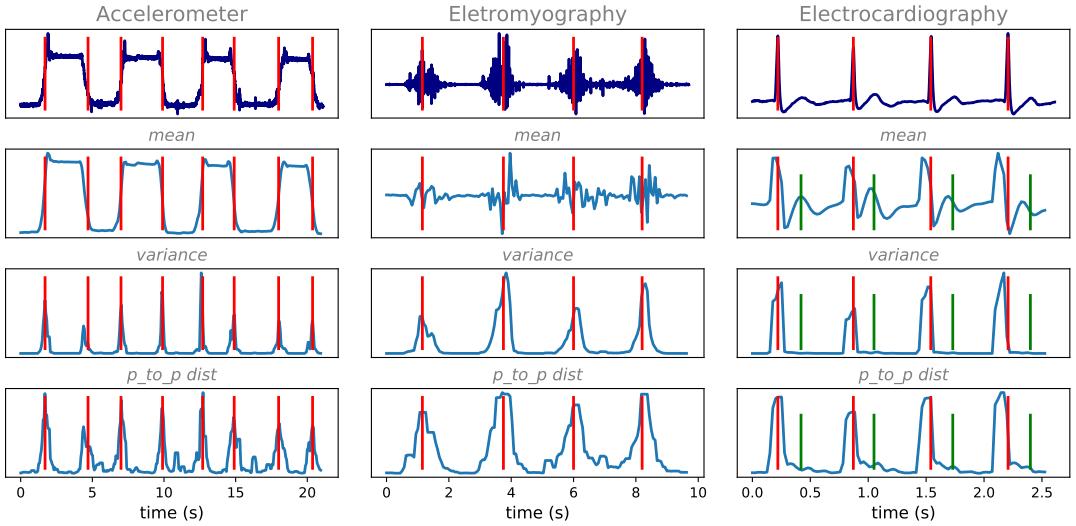


Figure 5.4: Feature selection schematic example of Acc, EMG, ECG[170, 171] signals. The upper, dark coloured signals represent the raw aquisitions and the following 3 × 3 signals below are the representations after applying 3 different feature operations (mean, variance, peak to peak distance). The red and green vertical lines serve to annotate point events of interest for each column.

This work focused on the use of shallow features over the temporal and statistical domain, and for an easier event detection, features from the existing TSFEL python library were used[172]. In the following table, there are listed all 19 features which were used in this work, as well as a brief description of the operations which have a more complex definition.

5.2.2 Window Length

The window's time length describes the size of the window, which remains constant throughout the time series. The event detection algorithm will search for events with a sliding window with a length that has to be as similar to the length of the periods of interest.

In the context of HAR, most studies consider the use of windows from 1 to 10 seconds [167, 173], as the main intent in this analysis is also to have described within these time intervals, signals that allow for the recognition of human activity. As in ergonomic data, there are multiple events of varying sizes, and without any previous information, there cannot be made any assumptions. This fact may present a great problem, as the approach made for this work was first to try to focus on the most fundamental repetitive event, which would theoretically be the working period (event type 1).

These would, as such, require a window with a size similar to the length of the working period, which will never be constant. Following that detection, the analysis may

Table 5.1: Feature list with the respective descriptions of each operation. The descriptions which are blank are considered self-explanatory.

Features	**Description**
Energy	
Median Absolute Difference	median of the absolute value of the derivative of the input signal
Sum Absolute Difference	sum of the absolute value of the derivative of the input signal
Centroid	centroid along the time axis
Centroid cum	absolute cumulative sum of the centroid along the time axis
Max peaks	number of maximum peaks of the signal
Min peaks	number of minimum peaks of the signal
Mean difference	mean of the derivative of the input signal
Mean absolute difference	sum of the absolute value of the derivative of the input signal
Distance	the total distance travelled by the signal
Area under the curve	
Max	
Min	
Mean	
Median	
RMS	Root Mean Square of the signal
Interquartile range	Difference between the first and third quartil
Variation	
total : 18	

further focus on the sub-events in each work period. With this logic, by repeating the methodology, it is theoretically possible to detect a hierarchically deep expansion of possible events by shorting the window length into smaller database segments. This, however, will not be made in this work.

The risk associated with retrieving features in a windowed manner is that it requires to be correctly adjusted. In the case of being too large, the feature time series may lose information by averaging too much the signal. In the case of being too short, the events of interest have the risk of not being as well represented, as it will also increase the computation processing of the algorithm due to the resulting greater dimension of the time series.

5.2.3 Window Overlap

The overlapping size describes how much the windows overlap with each other in their representation. This process was also approached in the literature review (Section 3.3.1), as it is a standard process to optimize the process of representation.

The definition of overlap fraction will affect the time of the algorithm processing and the amount of information retrieved. Under this, if the overlapping is too great, we might expect a heavy and slow process from part of the algorithm, which might also be pointless, as large overlapping windows usually tend to reveal an accuracy as great as

a total overlap. On the other extreme, an overlap that tends to zero might be a risky representation, as it can tend towards a reduction of the events, to the point where it is hardly noticeable.

5.2.4 Processing

Given all this theoretic description of the methodology, we can now describe the algorithm in practice. A firstly important characteristic of our method was the vectorization of the entire process to make it as efficient as possible. A summarized code of this application is displayed in algorithm 2.

Algorithm 2: Feature representation

```
int win_len, window length parameter;
float overlap_len, overlapping window length parameter;
list features, list of features to be retrieved parameter;
tseries s, raw signal time series;
int tf, l_dim = len(s), volume and dimention of the time series, respectively;
tseries output = None;
for dim in range(l_dim) do
    int w_range = int(win_len/2);
    tseries s_mirror ← r_concatenate (s[w_range:0:-1, dim], s, s[-1:tf-w_range:-1,
     dim]);
    int tf = len(s);
    int n_wind = int( (tf - win_len) / (win_len-overlap_len) +1 ) ;
    int overhang = (n_wind × win_len - (n_wind-1) × overlap_len) - tf;
    if overhang!=0 then
        tseries pad = zeros (overhang);
        tseries s_final ← r_concatenate( s_mirror, pad );
    end
    matrix chunked_s ← chunk_data (s_final, n_wind, win_len-overlap_len);
    tseries s_representation = None ;
    foreach feature ∈ features do
        f_r ← feature_operation (feature,chunk_data);
        s_representation ← r_concatenate(f_r);
    end
    output ← c_concatenate (output, s_representation)
end
return output
```

The process iterates along each dimension of the multivariate time series X. For each of the arrays, first, the time series is transformed. For this, it calculates and concatenates the mirroring of the borders of the signal. Then it checks if the time series does not require some more data points for it to be synchronized with N sliding window. In the case that it does, an array of zero points are added at the end of the time series. With this done, a matrix of N windows of chunked subsequences is created, each one of win_len time length and separated within each other by $win\_len - overlap\_len$. After this, the various features

are applied over the matrix, making the operation along the window simultaneously to all chunked subsequences. After all of these, the chunks are concatenated, with the process being iterated along all dimensions of the database until it obtains the final $output$.

5.3 SSM Calculation

As it was previously described, the main method which will be used for the detection of events will be the analysis of the resulting SSM. This theme was introduced in Section 2.3 and a deeper study within the literature of the subject was made in Section 3.3. As such, at this point, the SSM can be seen as a tool that is able to facilitate the visualization and description of dynamic behaviors prevalent in a specific signal. This is a technique that requires some contextual conditions to be used, or otherwise, it will not signal much relevant information. Overall, the data needs to 1) describe a sequence of patterns that are recurrent, and repetitive throughout the time series 2) there needs to exist some time intervals which are significantly dissimilar from the overall behaviour of the remaining time series. These are all conditions that can be seen in a work motion acquisition. The calculation of the SSM was not made with a typical distance measure, but rather, because we followed the process of audio signal analysis, we represented the time series into a feature matrix and used the dot product between the transposed matrix and itself. This process enables the calculation of the similarity between a window of the signal, described by all features extracted, and all the remaining windows.

Not entering again in further details, as it was already previously described, there is still a consideration regarding the calculation of the SSM of time series representation, which is the similarity distance that is used. In the literature, there are various measurements that are used, but this work will only use the dot product, as it displays a behaviour approximately as good as most complex measurements with the benefit of having a simpler and faster processing methodology. The facility of this method is that it only requires the inner product of the signal with transverse of itself.

Considering a representation of time series as a matrix $\overline{X}$, of $(k* \times n*)$ dimensions, then calculation of the SSM can be defined by

$$SSM = \overline{X}^T \cdot \overline{X}$$

, with $\overline{X}^T$ being the transposed of $\overline{X}$, a matrix of dimensions $(n* \times k*)$.

5.4 Event Detection

The following section describes the processes to search for different events. With this said, all these processes do share some characteristics in common:

1. Transform the SSM into a different, more simplified measurement where the intended characteristic is highlighted in 1 dimensional signal.

2. Perform a change point detection methodology along this new simplified signal to identify points that represent the intended events.

After explaining the actual detection, it will also be listed a set of variable parameters for each section. Then it will be described the method of selection used for each database.

5.4.1 Work Period Transition Events

In a working environment acquisition, the subjects' activity can firstly be divided into two moments: periods of Active Work (AW) where the subject is doing a specific repetitive and intensive job, and Non-active Work (NAW) where the worker is at the workstation, but not doing that specific AW, despite also not being necessarily at rest. Examples of NAW would be moments where they are instructing a partner, breaks in the car chain-supply, waiting for the next car to enter the manufacturing line, or moving themselves to a new workstation.

The kernel convolution intends to highlight moments of transition within blocks. The simplest definition of this kernel would, as such, be (2×2) unit checkerboard kernel, defined as

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

This definition considers the kernel as the difference between a "coherence"kernel and an "anti-coherence"kernel. The "coherence"kernel measures the self-similarity on either side of the center point, highlighting it whenever these two regions are more homogeneous. Meanwhile, the second kernel of "anti-coherence"measures the cross similarity between those two regions, being highlighted whenever these two regions are similar within each other. As this last component will be negative, the opposite will be expected. With this said, K_{Box} will be used to highlight the regions with a high level of self-similarity, but yet different between each other. As the interest of this work will be to detect changes between big-time intervals and not only two-point positions, these kernels will have to be further extended. 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 be either be 1 or -1 according to the same pattern as previously seen in K_{Box}. For example, if L=2, then

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

To avoid edge effects, this kernel is smoothed with a radially symmetric Gaussian function. This function is defined as

$$\phi(s,t) = \exp(-\varepsilon^2(s^2 + t^2))$$

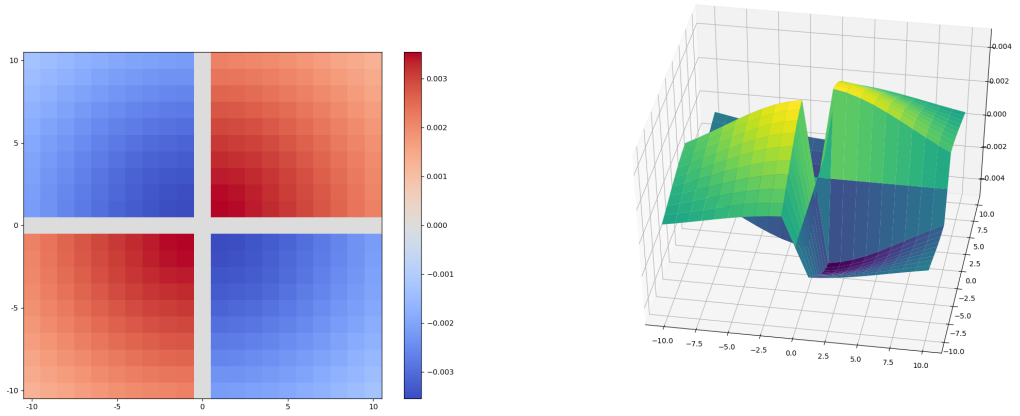
, with the parameter $\varepsilon >= 0$ adjusting the level of smoothing applied in the kernel. The resulting Kernel, will as such be described as :

$$K_{Gauss}(k,l) = \phi(s,t) \cdot K_{Box}(s,t)$$

Furthermore, the kernel is also normalized to compensate for the influence of the actual kernel size and of the tapering. The normalization is done by dividing the kernel by the sum of all absolute values of the kernel.

$$K_{Norm}(k,l) = \frac{K_{Gauss}(k,l)}{\sum_{k,l \in [-L,L]} | K_{Gauss}(k,l) |}$$

Figure 5.5 illustrates an example of this type of Kernel, for L = 10. This process was mostly inspired by the work made in [145], more specifically over the online notebook of [174].



(a) (b)

Figure 5.5: 2D (a) and 3D (b) representation of the kernel K_{Norm}, with dimension of 10 units.

This kernel K_{Norm} is then convoluted along the main diagonal of the SSM. This process will result in a convolution function $Conv_f$, a signal with greater values in the points with more contrast in between blocks. This methodology will compare at each point of the main diagonal the close neighborhood at the left and at the right of that point, and then compare it. In the case of them representing two blocks in the main diagonal of the SSM with a consistent texture, but still with a considerable difference between each other, then there will be a significant high peak on the $Conv_f$. In figure fig:schem convf, it is displayed an example where it is possible to observe the application of $Conv_f$. This signal can visibly highlight the time instances where the SSM transitions from high valued, squared, and homogenous blocks, which represent periods of NAW, to blocks with several parallel diagonals equally spaced, which represent the periods AW. From here, it's only necessary to use a peak detection methodology to identify the points of interest.

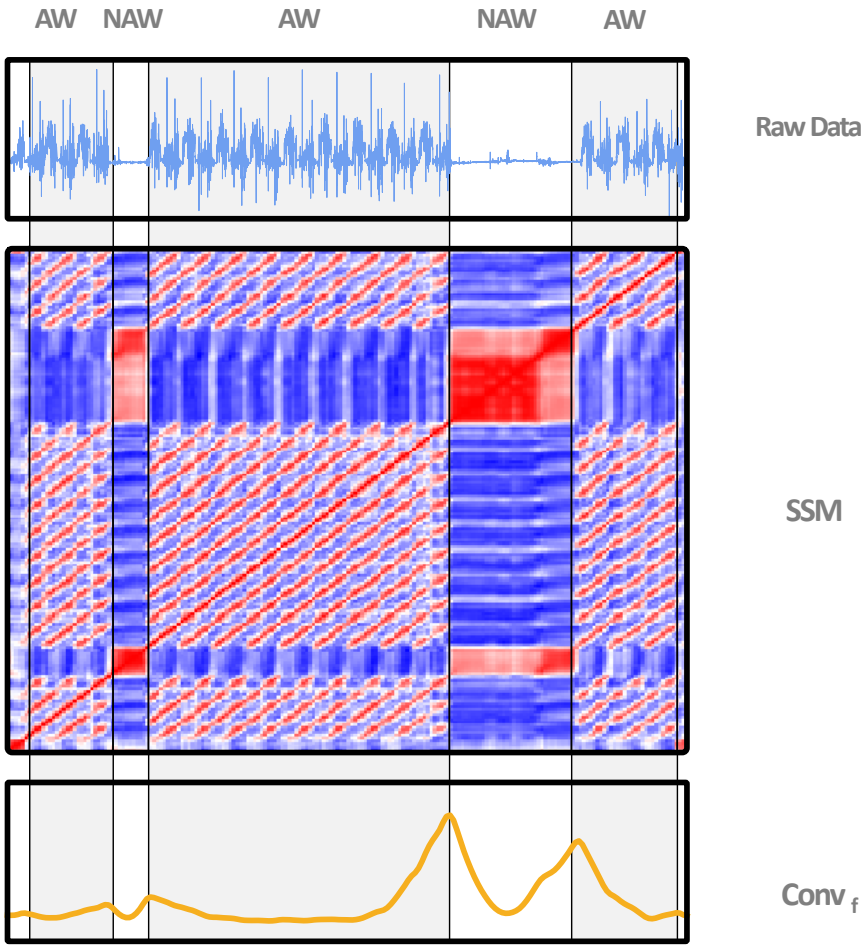


Figure 5.6: Schematic of construction of the Convolution Function $Conv_f$. Each image row represents a peak detection process. The first step represents the x coordinate of the hand accelerometer durig a working period, where the subject has several periods of active periodic work and periods on non active work, accordingly represented with the acronyms of AW and NAW at the top of the image.

In this following section, the peak detection methodology was structured in 3 processing phases. Firstly, it was subtracted to the signal its own averaged signal. This method highlights the peaks of interest present in the signal. Followed by this, it was applied a smoothing window to diminish or even eliminate the smaller peaks. Then the actual simple peak finding function was applied, with it only searching for time instants that had a positive derivative before that point and a negative derivative after that point. The final process was to select the events with a higher intensity than a given threshold, so the smaller peaks were not selected as correct points.

There are six variable hyperparameters that will be used for the detection of this event with each one being related with specific processing steps. Firstly, the $Wind_{len}$ and $Overlap_{frac}$, which were already introduced in section 5.2 will give information regarding

the process of data representation. Also introduced in this section there is the $Kernel_{len}$ which will regard the process of constructing the convolution function $Conv_f$. Regarding the previously described process of peak detection, there are three parameters to describe each one of those steps: $SubWind_{len}$, which describes the smoothing window length that will be subtracted in the first step; $SmoothWind_{len}$, which describes the length of the smoothing window of the second step; and $Thresh_{frac}$ for the final filtering of that signal.

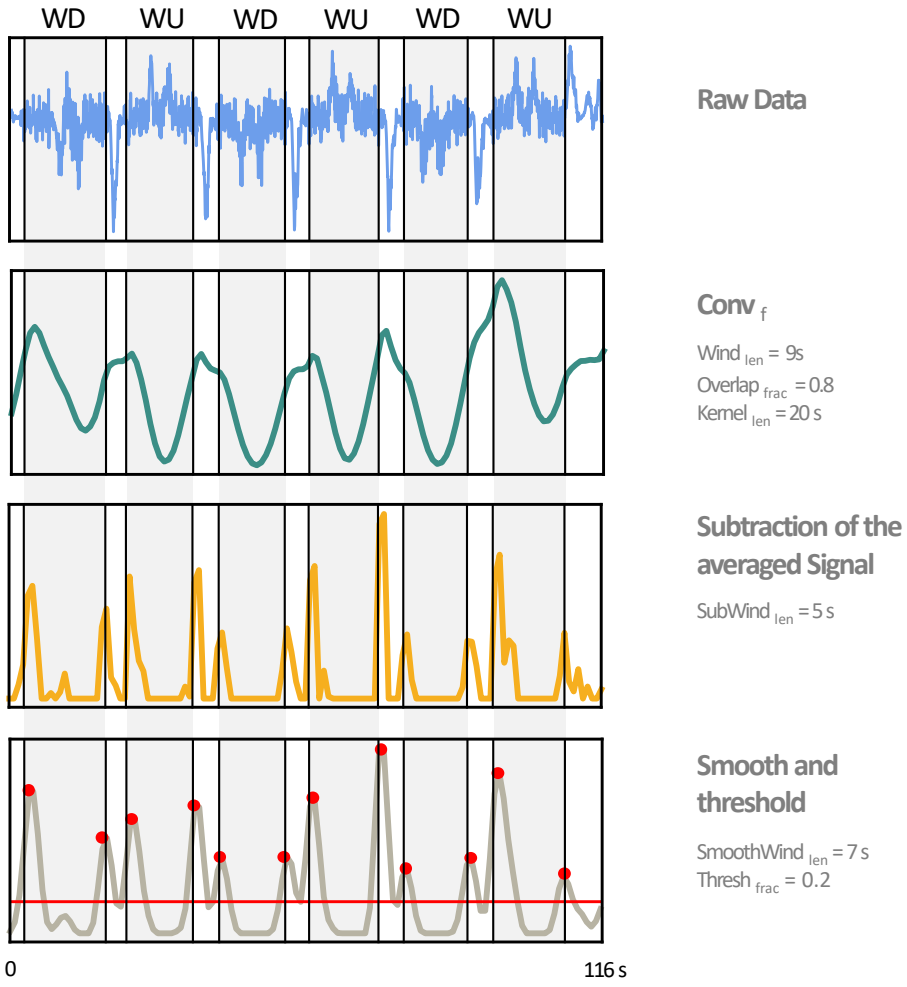


Figure 5.7: Schematic of Peak Detection. The represented example originates from a cropped sample (number 2) of the *UCI* database. Each image row represents a peak detection process, with the right column giving information regarding the process and the parameters required for that step. The first step represents the z coordinate of the accelerometer in a cycled process where the subject is asked to walk up and walk down some stairs, accordingly represented with the acronyms of WU and WD at the top of the image.

Figure 5.7 serves as example to further clarify the explanation of the peak detection and parameter selection. Each one of the schematic's rows depict processing steps to detect the events of interest. The first phase considers the raw data, with the time series

representing a transition between walking up (WU) and down (WD) some stairs. Importantly these steps are intercalated with a small transition time interval. During the second step, it is represented the $Conv_f$, which calls for the stipulation of the parameters $Wind_{len}$, $Overlap_{frac}$ and $Kernel_{len}$. For this example, the values of $Wind_{len}$ and $Overlap_{frac}$ were very purposively defined as very high, when compared to the shorter transition *states*, which on average, lasted only about 4,8 s. This resulted in a $Conv_f$ which only considered the main *states* of WU and WD, "filtering out"these smaller transitional *states*. Despite this being a helpful tool when it is intended to remove noise or to disregard some small changes present in different instances of the same type of *state*, in this context, it is prejudicial. As such, the third phase will subtract the averaged signal (created with a smoothing window of length $SubWind_{len}$) to highlight these smaller transitions. Finally, after further smoothing (with a window of length $SmoothWind_{len}$), the peak detection can actually be performed in this pre-processed signal. The only requirement in this final process is that the peaks must be over a threshold described by the parameter $Thresh_{frac}$ to assure that the events found only represent significant changes in the signal.

This methodology was tested in two database: *UCI* database, and then the *Industrial* database. The *UCI* database contains about 61 samples, each performing and transitioning between several diverse human motions. Due to the large number of time serie samples of this database, a hyperparameter optimization techniques was used to set the best possible parameters for each sample. The optimization of hyperparameters is a well-documented process that can be implemented in various manners[175], with the chosen process being a Grid Search. Grid Search is the most basic of these methodologies, being also known as full factorial design. It describes a process where every possible combination of a predefined collection of parameters is tested and assessed. The predefined collection of parameters for this section is presented in appendix B.1. As Grid Search suffers from the curse of dimensionality with the time complexity growing significantly with an increase on the resolution of the analysis[175], there were only selected some few values for each parameter. The set of parameters with the highest accuracy value, obtained with the Grid Search, was then saved, and the measurements of interest were calculated.

The *Industrial* database is composed of 11 samples where six operators perform their work in 2 different workstations. From this database, only seven samples were selected as some were very homogeneous, being only centered on the AW period and not having any instances of pause or deviation, which might be of interest to be identified. The parameter selection process did not include any hyperparameter optimization technique. Instead, the parameter selection was made manually, since there were very few samples which facilitated this process. These had also a large volume of data point, which slowed the algorithm and consequently made the optimization process harder to execute.

5.4.2 Work Cycle Transition Events

The next logical process is to focus on each one of the detected AW periods and to further detect instants of relevant interest, now on this subsegment. The initial assumption made in the previous chapter was that the work would be recurrent, then to properly organize the human activity, the most relevant analysis would be to segment the database into these periodic work cycles. It will be expected that the AW signal will only be composed of a single fundamental motif repeated recurrently over time.

During this section, there will be analyzed each block of AW individually after the previous detection has been made, only the subsegment of interest will be transformed into a SSM. This has the purpose of both simplifying the overall computing process and taking off the noise related with other time instants not relevant for the analysis. Ideally, what would be expected to have is an SSM composed only of multiple parallel diagonals (like exemplified in figure 5.8). These diagonals will represent each instant of the work cycle motif, giving information regarding their transition points and their length. With this said, the main purpose of this section will be to identify the time instants where a diagonal ends and another begins, to annotate the moments of transition between these motifs.

To detect this type of events, a different variation of the method of similarity function was used. To briefly review, a similarity function is a univariate time series where each value represents how different is a specific time instant with the rest of the time series. With this said, we expect that the points of transition will display a significant valley, in some way much like the method of matrix profile, which will also later be tested.

To calculate the similarity function Sim_f, it was made the sum throughout one of the axes of the SSM.

$$s_f(x) = \sum_{i=0}^{N} SSM_{ix}$$

with, N being the size of the SSM. Due to the property of symmetry of the SSM, this function can either be applied over the lines or the columns, with the final result being the same. Figure 5.8 presents an example of the application of a similarity functionSim_f. This signal has an almost sinusoidal behavior with the valley positions synchronizing perfectly with the points where the parallel diagonals end in the upper corner of the SSM, and another one starts in the lower corner of the SSM. These points indicate the time instances when a work cycle ends and another restart. As such, it is only necessary to further apply a valley detection methodology, to identify the points of interest.

Also, like the previous section, the valley detection will first pass by a simpler but much more extensive database, in this case, the *HAR* database, before analyzing the *Industrial* database. The main requirement of choice for the *HAR* database is that the sample data should only display a specific human motion repeated periodically over time, like what is excepted to be studied under a work-related ergonomic context acquisition.

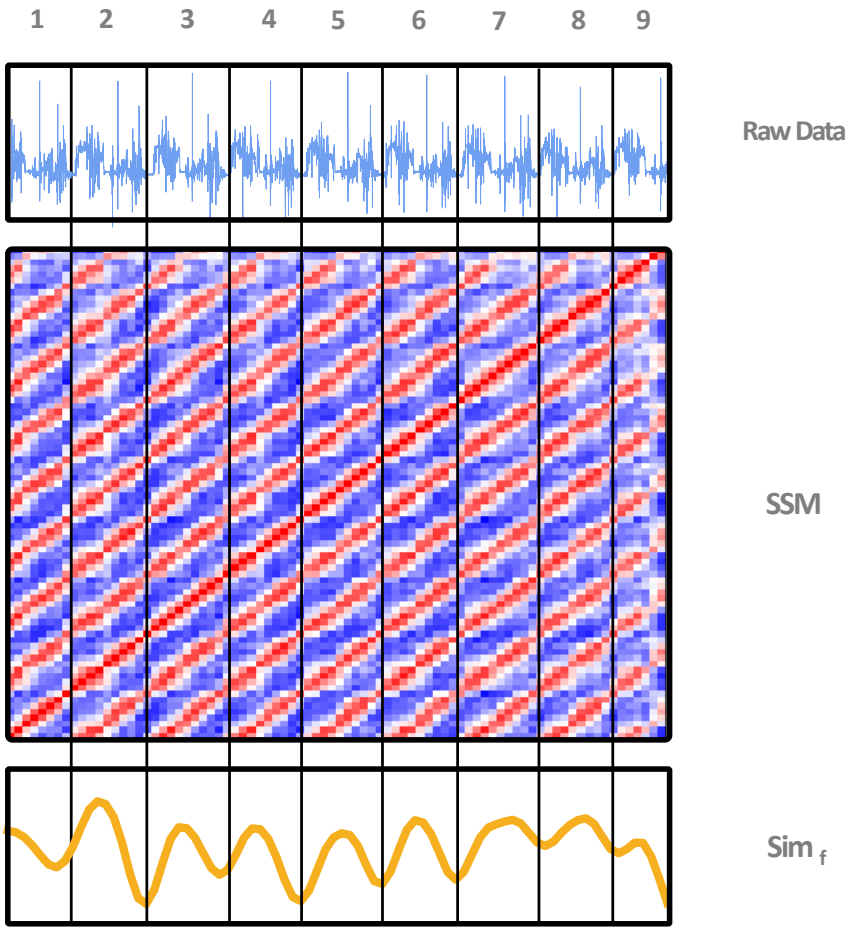


Figure 5.8: Schematic of construction of similarity function Sim_f. Each image row represents a process step. The first step represents the an accelerometer axis signal a cycled process where the subject is doing an active work exercise. This signal also corresponds to the second AW period represented in Figure 5.6

Another change in this chapter will be the peak/valley detection process. As the distance between event annotations does not tend to variate considerably, there is no need to make the subtraction process, as this was mainly intended to intensify smaller peaks that were being absorbed into more prominent peaks. Instead, the detection comprises only the smoothing of the similarity function Sim_f, and the sequent valley detection, also without the application of a threshold. This results in a process with only three parameters: $Wind_{len}$, $Overlap_{frac}$, which like the previous section, indicate the length and overlap fraction of the window function, which takes the features along the time series; and $SmoothWind_{len}$ which indicates the length of the smoothing window function applied to the Sim_f before getting the valley points of interest.

Much like section **??**, there was also applied a grid search, over the various time series samples, with the parameters of interest being listed in appendix C.1. The reasoning for this procedure was the same as that previous section, as the database was significantly

extensive, it is simpler to automatize the adjustment of those values.

The parameter combinations selected were those able to detect all the cycles of interest, with the smallest value of duration error D_e possible. The measurement of duration errorD_e will be further explained in section [**subsec:evaluationmetrics**]. Once again, the *Industrial* database had a manual selection of the parameters.

5.4.3 Sub-Segment Example Search

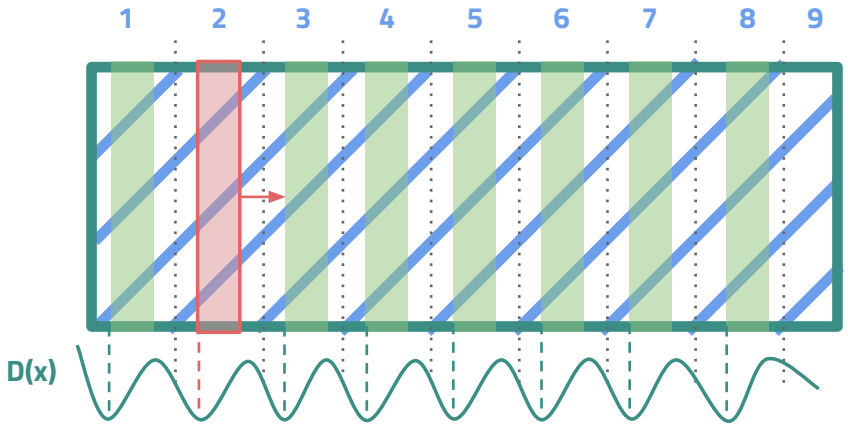


Figure 5.9: Schematic of construction of distance function $Dist_f$. The figure schematizes a simplified representation of an SSM with several repeating parallel diagonals represented in blue. The strides in green represent the various repetitions of the subsegment in red. In the bottom there is a representation of the expected behaviour of $Dist_f$ with the valleys intersecting with the starting points of the repeated sub sequences

The two previous detection methods were mainly concerned with the identification of the structure of the working schedule. However, this subsection will focus on identifying moments that may have some relevance for the analyzer through a method of search by example. There is great importance in detecting some specific human action throughout a working database. For example, if a more concerning motion is performed and detected, an ergonomist may want to see how recurrently that action is performed, allowing for a deeper analysis of the ergonomic risk. This proposal will nonetheless require first identifying a specific motion in the time series, then the method will search for moments of greater similarity in the rest of the SSM.

After identified the subsegment of interest SSM_{t0}, and considering it has an origin in the point t_0 and a time length of L, then a distance function is performed between this segment and the rest of the time series. To calculate this distance function $Dist_f$ a sliding window is applied throughout the matrix, where the sum of absolute differences between the windowed matrix and the sub-segment of interest is calculated.

$$Dist_f(t) = \sum_{t=0}^{x=N} \sqrt{(SSM(t) - SSM_{t0})^2}$$

, where SSM(t) is a subsegment of the matrix which starts in point t and has a length of L (Figure 5.9). After this, the detection of these events requires only a valley detection methodology along the distance function. The identified instant will represent the time instants where the subsegment of length L starts.

The peak detection methodology will be the same as the one made in the previous section 5.4.2, with the distance function $Dist_f$ being passed by a simple smoothing window before being identified the points of interest. The parameter set was only composed by three variables: $Wind_{len}$, $Overlap_{frac}$, $SmoothWind_{len}$. The first two parameters would serve to codify the length and overlap fraction of the window function representation, while the $SmoothWind_{len}$ would inform about the window's length for the smoothing during the peak detection.

Chapter 6

Results and Discussion

In this chapter will be presented the main results achieved with the proposed methods. These methodologies will be tested in various databases to evaluate their applicability to detect the three event types of interest in a human motion data context: 1) work period transition; 2) work cycle transition; and 3) sub-sequence match with query. There will also be a joint discussion of these results along with their presentation. For each database, the results will be presented based on the ground truth of events, being the accuracy of the method displayed based on metrics explained in section 5.4. Regarding type event 2 (Work Cycle Transition), the results will also be compared with the Matrix Profile.

The structure of this chapter is divided into three sections depending on what type of event is meant to be detected, as was made in section 5.4. Each section will start by describing the used score metrics, the peak detection techniques, the parameter variability, and other methods of interest to understand the results acquired. This explanation step will then be followed by the presentation of results of these metrics when applied to two databases, with the final event of "sub-sequence match with query"only being tested in the *Industrial* database. For the first two sections, each method will be tested with the standard and lab-controlled databases (*HAR* and *UCI* databases) before showing its applicability to the real-environment database from the automotive industry. The results will be discussed afterwards.

6.1 Change Point Detection

The problematic of the change point detection is described in section 5.4.1. Shortly summarised, a kernel convolution is used to create a convolution function $Conv_f$, which is a time series where each sample point represents a value of how the data is dissimilar from behind and beyond that time instant. The peak points are then identified as events,

as these point instants represent moments of significant dissimilarity when compared with their neighboring area. Before moving to test the *Industrial* database, this section will start by providing an analysis of the *UCI* database, as it affords a continuous sequence of various simple human motions at different time lengths and repetition cycles. The intent of this is to test if the algorithm can indeed segment efficiently different types of human activities. Over this section, it will be used the term *state* to describe a homogeneous type of human motion that is expected to be identified. To clarify, this means that after all change point events are detected over a time series, it will be segmented into various non overlapping*states* with each one representing some different dynamic behaviour of the signal.

6.1.1 Evaluation metrics

To assess how well the algorithm identifies the transitions, the events identified by the algorithm B_{Est} will need to be compared with the ground truth annotations B_{Ref} of the data. These are obtained *a priori*, with the case of the *Industrial* database being labeled manually, while the *UCI* database had its annotations already provided in the online source. With this setting, the definition of a correct event detection should not be seen simply as determining if the time instant of the found event is equal to the respective ground truth annotation, as in essence, this precision would be close to impossible. The ground truth is not necessarily the most precise value, as for instances in the case of the *Industrial* database, the annotations were done manually by analyzing the video recording done during the acquisition. Despite being a good method, it still is susceptible to small divergences when transferred to the actual signal time series. Furthermore, human motion signals are very complex and noisy, besides, a repeated human motion tends to return mildly different signals each time, and a specific motion might also be divided into hierarchically less specific motions. An example would be the ECG or Gait Signal, which, despite having an easily identifiable structure, can be further sub-divided into other minor recurrent waveform structures (ECG into P wave, QRS complex, and T wave; Gait into Heel Strike, midstance and toe off). These substructures might incorrectly be seen as segmentable options by an unsupervised algorithm. This concern regarding the various forms that the same signal might be divided was solved by applying several smoothing steps throughout the data processing and by creating adjustable parameters which codified, in part, the degree of smoothing applied. This way there can be delineated a resolution of events of interest to be detected by the analyzer while at the same type removing unnecessary information. Nonetheless, this usually involves a loss of precision of the detection process. As such, it will be argued that the main interest will not be to identify the exact time instant of annotation, but instead that the detection of the event should be close enough to the event annotation that it is concluded that what is being identified is indeed the dynamic behavior of interest. The easiest method to achieve this will be to define a tolerance distance τ centered in this ground truth points, such that the

events are only considered as truth if it fits within this close time interval.

To further detail on this subject, let us consider a ground truth annotation of an event $b^{Ref} \in B_{Ref}$ and the respective event estimation made by the algorithm $b^{Est} \in B_{Est}$. Then, the transition event is correctly identified if

$$| b^{Est} - b^{Ref} | \leq \tau$$

These instant occurrences are defined as True Positive (TP). On the other hand, the remaining b^{Est} which do not fulfill that condition are defined as False Positive (FP). Likewise, the remaining b^{Ref} with no events in its τ-neighborhood are considered as False Negative(FN). Now, if it is analyzed a more ambiguous case scenario, given a b^{Ref} point instant, if, within the time interval of its τ-neighborhood, there are present severalb^{Est} then one will be considered a TP point, while the others will be considered as noise and be counted as a FP points. Between these events, the one singled out to be TP is the b^{Est} instant with the closest time distance to b^{Ref}.

After defining these occurrences, it is possible to calculate the accuracy metric, defined by the expression of

$$A = \frac{nTP}{nTP + nFP + nFN} \tag{6.1}$$

where nTP, nFP, and nFN represent the number of TP, FP, and FN instances counted, respectively. This formula is a variation from the classic accuracy formula from statistics, where the True Negative (TN) number of occurrences (nTN) would also be added to the nominator and the denominator ($A = nTP + nTN/nTP + nFP + nFN$). Given the previous descriptions of the occurrences, what would be expected is for the TN points to be all the points that are correctly not altered. In other words, they would be all the time series coordinates that are not included either on B_{Est} or in B_{Ref}. However, this classic description raises a problem, as, under it, nTN is not only far superior to the sum of nTP, nFP, and nFN, but they also result in very different values depending on the sampling rate of the signal. This is an issue that is defined in the literature as an imbalanced classification. Meaning that not only, the accuracy results would constantly be very high, but that it would not be possible to compare signals of considerably different lengths or different sampling rates. The accuracy measure intends to theoretically assess how many correct predictions exist among the total number of cases examined. The reasoning behind the formulation of equation (6.1) is that if it is only considered the cases of interest to be the assessed event points, then we can eliminate the TN from the formula. Nevertheless, it will also be presented the results of precision (P), recall (R), and F_n-score (F), as these are statistic measurements that do not require to take into consideration nTN, with the last one being cited as a common substitute to be used in a binary imbalanced classification context[176], like this one. The formulation of these measurements are followed below:

$$P = \frac{nTP}{nTP + nFP}$$

$$R = \frac{nTP}{nTP + nFN}$$
$$F = \frac{2PR}{P + R}$$

This entire explanation still requires the pre-definition of the value τ. Theoretically, τ could vary within the time intervals of

$$0 < \tau < \frac{|\, {b^{Ref}}_{k+1} - {b^{Ref}}_{k} \,|}{2}$$

, with the last part of the equation meaning that the maximum value does not make sense if it is superior to half the distance between two consecutive event annotations. Because this will mean that the tolerance windows will overlap with each other and that the tolerance value is superior to the time length of the *states* under analysis. For each time series sample, the τ was defined as being the maximum value possible, as these will return the best possible values. This is the minimal value of the distances between reference events divided by 2.

To also retrieve information about how different the time records of B_{Est} and B_{Ref} are, it is also required to have approximation measurements. For this end, the selected measurement was the MAE, which is defined as the mean value of the absolute difference between every b^{Est} and b^{Ref} match. In the case of FN or FP instances existing, those will be ignored for this assessment, and only the TP will be considered. This liberty can be made due to the first analysis concerning the accuracy, with the intent of this approximation error analysis being only to understand how far away the events correctly classified are to ground annotations. The formulation of this measurement is the following

$$MAE = \frac{1}{n} \sum_{n=1}^{n} |\, b^{Est} - b^{Ref} \,| \tag{6.2}$$

, where n is the nTP, and $b^{Est} \in B_{Est}$, $b^{Ref} \in B_{Ref}$ are in this context the time coordinates for each of the matched TP instances.

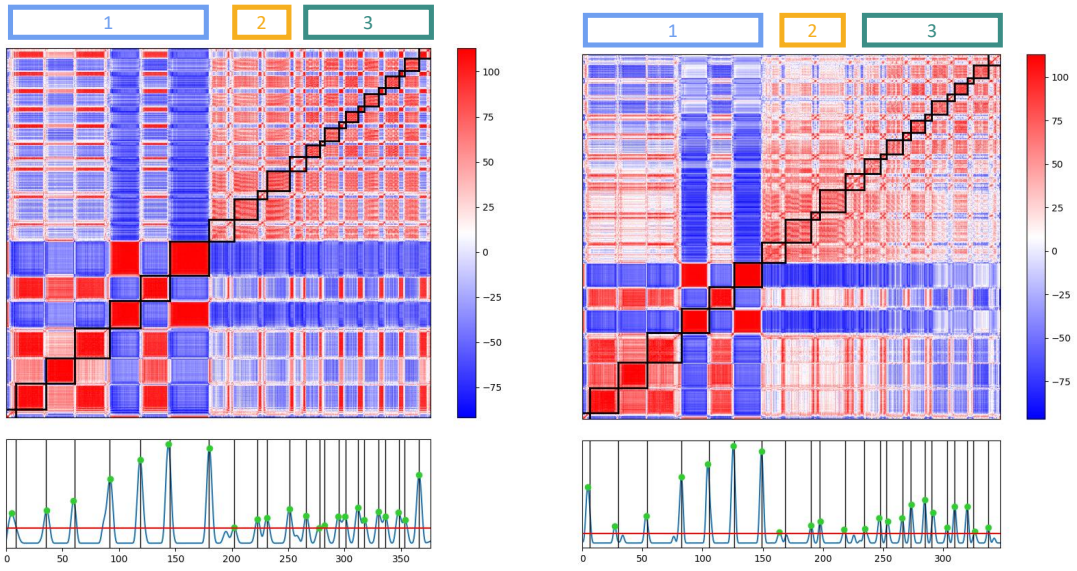
6.1.2 Human Activity Change Point Detection

The following section will serve to present and discuss the results retrieved from the *UCI* Database. As this database is too extensive, having 61 different samples, there was a need too summarize their presentation. The results of this method are all discriminated in appendix B.3, with the smaller table of 6.1, providing an overview statistical summary.

From table 6.1, the first conclusion to be taken is that as long as the parameters are properly adjusted, this method can provide very high values of A, P, R, and F. This proves the validity of this methodology for detecting moments of transition between different human motions in inertial sensor signals. The example provided in figure 6.1 illustrates two examples of a perfect classification of events, where the blocks created by the SSM, and the peaks that exist in the pre-processed $Conv_f$ are visually evident

Table 6.1: Statistical summary of the A, R, P, F and MAE results, for the detection of type 1 events, over the *UCI* Database. The discriminated results for each sample are presented in appendix B.3. This table focus on presenting the distribution of these measurements through: the first and third quartiles (1^{st} Q and 3^{rd} Q, respectively), the mean, maximum (max), and minimum values (min), for all the time series samples of *UCI* Database.

	A	R	P	F	MAE (s)
min	0.73	0.76	0.76	0.84	0.31
1^{st} Q	0.83	0.9	0.9	0.91	0.62
mean	0.88	0.94	0.93	0.93	0.89
3^{rd} Q	0.92	1	1	0.96	1.14
max	1	1	1	1	1.84



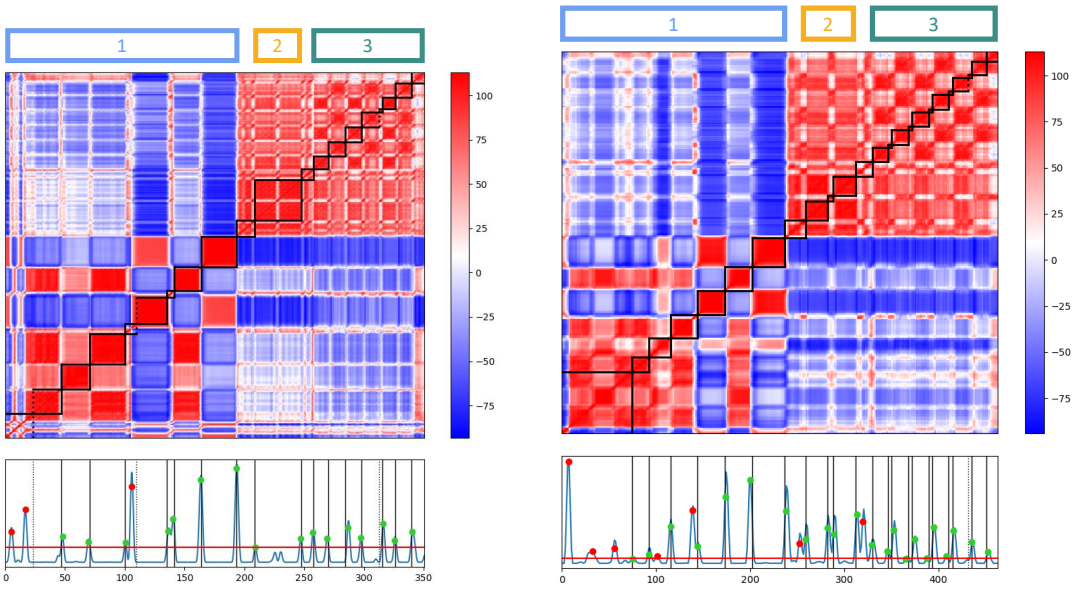
(a) (b)

Figure 6.1: The 2 best case scenario examples of a detection of type 1 events, in the *UCI* Database. The colored blocks indicate what type of motion is being performed in those time instances, with block 1: transition between sitting, standing, and lying states of motion; block 2: walking; and block 3: walking up and downstairs. The larger squared image in the center represents the SSM of the time serie sample 3 for subfigure (a) and 35 for subfigure (b). The higher values of similarity are indicated with a red color and lowest with a blue color. The signal of blue color represents the respective $Conv_f$, the dark lines represent all the B_{Ref} instants and the green points represent TP instances of B_{Est}.

to intersect perfectly with the respective ground annotations. The examples of figure 6.1 represent a perfect classification because not only all the eventsB_{Ref} were found but there was no instant of a false classification. Moreover the MAE were considerably small being 0.87s and 0.86 s, for subfigures 6.1a and 6.1b, respectively. These two examples also inform about the composition of the tested data, as most samples followed the same

structure of these examples. Most samples were about 349,19 s (varying within 151.52-460.8 s) and had about 21 *states*. Most of these *states* represented simple human motions as described in section 4.2, notwithstanding some few shorter *states* also represented transitional periods between those requested motions. The most common structure of these samples usually started by transitioning between sitting, standing, and lying, with each *states* lasting on average 2.29 s (with the dispersion of time between these three motions being reasonably similar), represented in the top corner of the figures 6.1 and 6.2, by the block 1. This was followed by a walk usually intercalated with stopping intervals and represented by block 2. The *states* of continuously walking would vary a lot between 2.8 - 28.2 s but being on average closer to 19.21 s; meanwhile, the stopping intervals tended to happen only one in each sample, with the highest number of occurrences being three, also in a much shorter period, never surpassing 10.2 s. In the end, they would perform the shorter motions of walking up and downstairs during block 3, with these being repeated 3 or 4 times (excluding an anomalous sample with 0), lasting on average 12.16s (between 4.3-17.89 s). The significance of these deeper inspections is that by being validated with this database, it means that this methodology is also able to: 1) separate significantly different temporal dimensional *states* of human motion; 2) separate between repetitive *states* (like walking or going up or downstairs) but also between static ones (like sitting, standing and lying); and 3) separating between significantly similar actions like the ones which were previously listed.

After obtaining the best possible set of parameters for each sample, it was retrieved the TP, FP, and FN occurrences, and from those, it was calculated the scores of A, P, R, and F. It was obtained an average A value of 0.88 for each sample, a considerably great value. The smaller accuracy registered was as great as 0.72 and the highest as 1, with six sample acquisitions displaying this perfect A score, which means that despite most samples still having some wrong event identification, these classifications are relatively few and usually distributed throughout the samples. Even in the cases where the event identification is not perfect, this methodology still provides a relatively good result. As expected, the f-score provides results relatively similar despite slightly higher. The samples with the highest score of 1 are still the same, but the lowest value is 0.84, and the mean is 0.93. In contrast, the P and R measurements are more tendentious to have higher values, having a skewness of -0.66 and -0.91, respectively, while the A and F-score had a skewness of -0.25 and -0.08. These values describe a considerably higher distribution of P and R towards greater values, while A and R distribution remain considerably more centralized. This is an interesting contrast, as F-score considers the values of P and R in equal proportion in its formulation. Such a different behavior from P and R, to R means that when one of the P, R scores is in its minimum, the other measurement tends to be on its maximum range values. In other words, in the samples with a lower score, there tends to either exist an oversized number o FP or FN, seldom some rare exceptions they are in approximate proportions. In this context, when possible, it is preferable to opt for a greater recall instead of precision. A presence of FP could be noticed in a further



(a) (b)

Figure 6.2: Some of the 2 worst case scenario examples of a detection of type 1 events, in the *UCI* Database, when considering the A value. The colored blocks indicate the motion type, with block 1: sitting, standing, and lyingn; block 2: walking; and block 3: walking up and downstairs. The SSMs are calculated from the time series sample 57 and 6, in the subfigures (a) and (b), respectively. The signal of blue color represents the $Conv_f$, the dark lines represent all the B_{Ref} instants, with the dashed dark lines representing an undetectedB_{Ref} instant, i.e. a FN instant. The green points represent TP instances of B_{Est}, while the red points represent FP instances.

analysis of the results, as, after all, the motivation of this work is to serve as a support for the analysis of ergonomic data. There was always assumed by the start that there would be some following interactivity with a user, as such, the identification of FP becomes a process much less demanding to solve when compared to FN. The FN assumes that there was an inability to distinguish between different motion *states*, which requires an assessment of the entire time series, a much greater effort that as such should be avoided.

To better understand the limitation of the proposed methodology, the worst scoring samples were analyzed, and the contexts where the detection fell behind were noted. To facilitate the following explanation, Figure 6.2 will provide some visual support to better understand the difficulties of this type of event detection. To simplify, the following explanation will focus on the A, but the pointed concerns are relevant in samples with lower P, and R. Overall there were pointed three types of instances that led to a lower A score:

1. In some case samples, some motion changes were performed during the database acquisition but were not annotated. This happened mostly during the first seconds of acquisition (subfigure 6.2a 0-20 s, subfigure 6.2b 0-70 s), but could also happen in between motion states(subfigure 6.2a 224-232 s, subfigure 6.2b 321 s). Despite

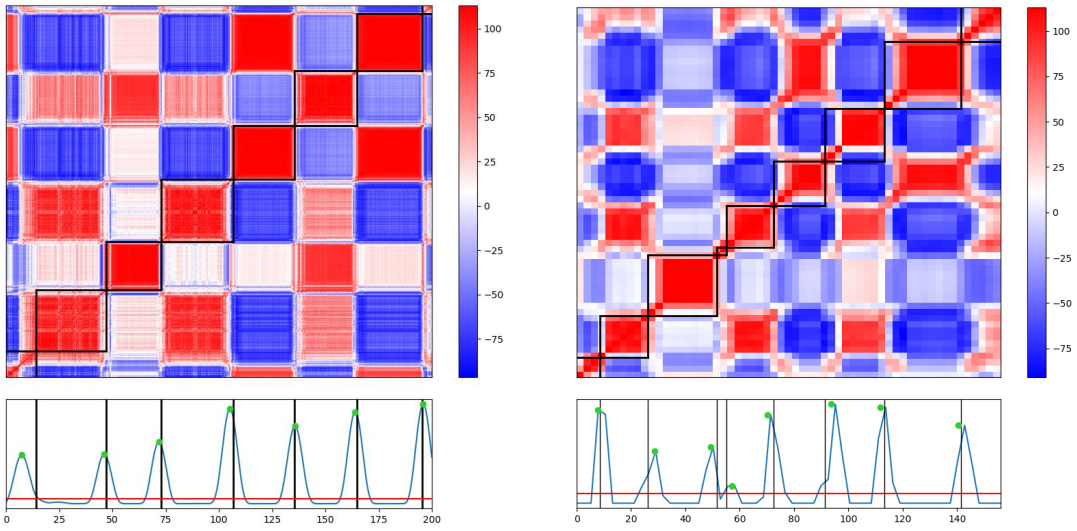
being a concern more related to the annotation instead of the methodology, this point issue was indeed a factor that was often present in these results, lowering the overall score measurements.

2. There might be some miss adjustment of the parameters across the time series. As it was already explained, the ideal parameters will be associated with the time duration of the *state* subsequences which are meant to be identified. What may happen in some time serie samples is an extreme variation in the length of different *state*s. This means that the parameters optimized for some *states* are not suitable for different *states* of the same time serie. This may lead to one of two outcomes to happen or 1) the transition representing either the beginning or the ending of a certain *state* is not identified, due to the fact of them being subsequences that are too brief in time, and the parameters used being too great. Leading to an over smoothing and the absorption of these events, falling to detect an event that was annotated and consequently increasing the nFN. To serve as an example there are subfigure 6.2a within 312-316 s, subfigure 6.2b within the 430-437 s). In both of these cases, there is a noticeable but small peak, which is unfortunately too small to pass the specified threshold; 2) Conversely, the transition between *states* may also develop multiple points when only one was expected. Contrary to the previous concern, sometimes the transition may be considered approximately instantaneous, although in practice is much slower, at least when comparing to the predefined parameters. This, as expected, will lead to an increase on nFP, with some examples including subfigure 6.2a 224-232 s, subfigure 6.2b 144 s, and between 252 - 260 s.

The first point is primarily due to a limitation provided by the database annotation, as it was not without some noise associated with unpredictable motion changes. With this said, despite the methodology used not being directly at fault, it means that the algorithm still can be sensitive to changes that were not predictable initially by the user and that care should be taken into consideration. The second issue is much harder to solve, but it seems that there needs to be some consideration relating to the dynamic nature of the time series, as the parameters which are adjusted for the identification of a specific *state* might not be appropriate for another *state* of a much shorter or longer time dimension. With figure 6.2 it was possible to see that the parameters may both be too small and too high for different *states* in the same time series. This analysis is essential to make further improvements while acknowledging the limitations of the current methodology, but overall we can conclude that FP instances are usually associated with a specific motion that is somewhat present in the data but is not intended to be annotated. Meanwhile, the FN instances are related with an over smoothing of the data making the prominence of the events too short to be efficiently detected.

As introduced before, these results are all obtained with the predefined assumption that there is a degree of tolerance τ, where the identified events B_{Est} are matched with reference points B_{Ref} independently of how distant these points are from each other, as long

as they do not go over the τ threshold. Of course, there is an interest in understanding how different the annotations made with the proposed kernel convolution and the ground truth are. To this end, there was retrieved the measurement of MAE to understand better how different is B_{Est} and B_{Ref}, with the formulation having been presented in 6.2. By calculating the MAE for each sample, an average value of about 0.89 seconds was obtained, varying between 0,31 and 1,84 seconds. Overall, these are all very small time distances, especially considering that the time series samples are on average 349.19s. Adding to that, the maximum distance of error between an event and its matched annotation is 6,7 seconds, a yet reasonably close error.



(a) (b)

Figure 6.3: Some of the 2 worst case scenario examples of a detection of type 1 events, in the *UCI* Database, when considering the MAE value. The subfigure (a), represents a zoomed sub-section of the time serie sample 36, meanwhile subfigure (b), represents the time serie sample 20. The signal of blue color represents the $Conv_f$, the dark lines represent all the B_{Ref} instants, with the green points represent TP instances of B_{Est}. The subfigure (a) presents the B_{Est}, B_{Ref} match with greatest error within all the database and subfigure (b) presents the worst MAE score, with it being 1.84s.

Once again, by analyzing the worst-case scenarios, it is possible to identify that the highest MAE distances were due to some particular circumstances. Figure 6.3 will display samples that display the worst MAE score to further serve as examples when explaining the major concerns of this methodology. As already mentioned, there are some samples with disturbances happening during the first seconds of acquisition. In the example of subfigure 6.3a, the first identified point is farther from its respective reference point than the other instances. This actually represents the greatest distance error acquired in the entire *UCI* database, and it is due to the motion before the acquisition being very similar to the first *state*, which in this case was just «standing». This, of course, is more an error of annotation than an error of the algorithm's performance, but once again, it

reveals that it can still be sensitive to noise and that this concern must always be taken into account. In the second subfigure 6.3b, we can observe a different case. This is an interesting case because despite it having a maximum Accuracy score of 1, it has the lowest scores of MAE. Moreover, it also follows a very different structure than most other sample time series from this database, only having the equivalent to a block 1 displaying only the motions of sitting, standing, and lying. The explanation for this worst measurement is because the *states* are without one exception considerably significant in time, especially when considering the size of the time series, which is much more reduced. This leads to an enlargement of the windowing parameters, in this case, more of the $Wind_{len}$ and the $Kernel_{len}$ with the final objective of not identifying the lower-level motions which are not of interest. The risk of a greater extension of these parameters is the over smoothing and loss of temporal definition of the data, consequently results in a greater lag between the time instant where the information had initially been and where it is actually represented. In subfigure 6.3b it is visually evident that the $Conv_f$ and the SSM have a meager resolution due to parameters that were very highly defined. Resulting in a considerably great difference between the point detected B_{Ref} and their matching ground annotations B_{Est}.

Table 6.2: Statistical summary of the parameters, for the detection of type 1 events, over the *UCI* Database. The discriminated results for each sample are presented in Annex B.2. This table focus on presenting the distribution of the selected parameters through: the first and third quartiles (1^{st} Q and 3^{rd} Q, respectively), the mean, maximum (max), and minimum values (min), for all the time series samples of *UCI* Database.

	$Wind_{len}$ (s)	$Overlap_{frac}$	$Kernel_{len}$ (s)	$SubWind_{len}$ (s)	$SmoothWind_{len}$ (s)	$Thresh_{frac}$
Min	1	0.5	5	5	0	0
1^{st} Q	1	0,6	10	10	5	0
Mean	5.13	0.71	13.93	15.16	6.31	0.06
3^{rd} Q	9	0,8	20	20	10	0,1
Max	17	0.8	25	20	15	0.2

As a note regarding the parameterization of this function, it is self-evident that there is some considerable variability within the values considered, being hard to interpret any infallible rule of selection within these results. Moreover, as the parameter selection method was a simple grid search, there is still much space to make optimizations, and it means that these values might just be brute approximations of what are actually the best set of parameters for each time serie sample. Non-withstanding, there are some interesting patterns that surged when considering the entire set of the data results. As for starters it seems that the value of the $Wind_{len}$ tends to be close to the duration of the shortest *state* which vary between 2.8s and 7.1 s, per time serie sample. Meanwhile, the mean $Kernel_{len}$ value tend to be closer to the average value of*state* duration, which varies between 10.28s and 19.92s, per sample. These conclusions in retrospective make sense, as a $Wind_{len}$ closer to the smallest *state* duration is essential to make sure that all *states* are

somewhat differentiated before transporting that information to the SSM. Otherwise, this smaller *states* would, in theory, tend to be enveloped in the representation step by their neighborhood. Concerning the $Overlap_{frac}$, its value much interestingly tends towards 0.7, proving that it is unnecessary to increase this parameter too much, as it also has the disadvantage of considerably enhancing the time requirements of the algorithm. The parameters $SubWind_{len}$ and $SmoothWind_{len}$ vary a lot between the set of values provided in the grid search, which do not allow to retrieve any definite conclusion. The threshold is always considerably small, even within the small values which were provided, leading to the conclusion that the *states* are considerably well discriminated, with each one being differentiated enough to be easily detected. This is an exciting conclusion as similar *states* like «walking upstairs» and «walking downstairs» would, at first sight, seem to be very hard to distinguish when using an unsupervised algorithm without previous information related to the data. This analysis is important as the assessment of these relationships can facilitate parameter selection for future processes.

6.1.3 Working period segmentation

The following section will serve to present and discuss the results retrieved from the*Industrial* database. Most samples that were selected for this analysis were composed chiefly by instances of periodic AW motion, whose duration would vary between 187 s and 1169.11 s. These instances would be intercalated with 1 to 3 moments of rest, which tended to be much shorter, lasting on average only 128.53 s. Also, all of these samples had a calibration time at the start of the acquisition, which might vary considerably between 65s and 211.39s, with some cases also having another calibration phase at the end of the sample acquisition. During NAW periods, beyond the calibration posture and just being at rest, the worker may be waiting for the next piece of the assembly line, it may be giving or receive instructions or transit between different workstations. However, it is important to retain that these moments of NAW were not necessarily periods of inactivity. As such, the main objective of this section will be to segment, from the complex motion time series, these various *states* of AW and NAW, by detecting the events which represented the instances of transition.

Figure 6.4 schematizes this process, where a SSM constructed over the time series of sample "Opr2 Wkst1&2"is decomposed into 7 blocks along the main diagonal, with 4 of them representing NAW *states* and 3 blocks representing AW *states*. The NAW blocks translate instances when the worker is not performing the actual intended work. In this case, the first and last blocks of the time sample represent processes of device calibration at the start and the end of the acquisition. There is also a period in the middle of Wkst1 when there is a stop in the assembly line, and the subject remains still while waiting, and another third block where the worker moves into a new workstation. The other three AW blocks represent periods of work in two different workstations, with the symbols "A"and "B"distinguishing between the workstation type. The different structure of these blocks

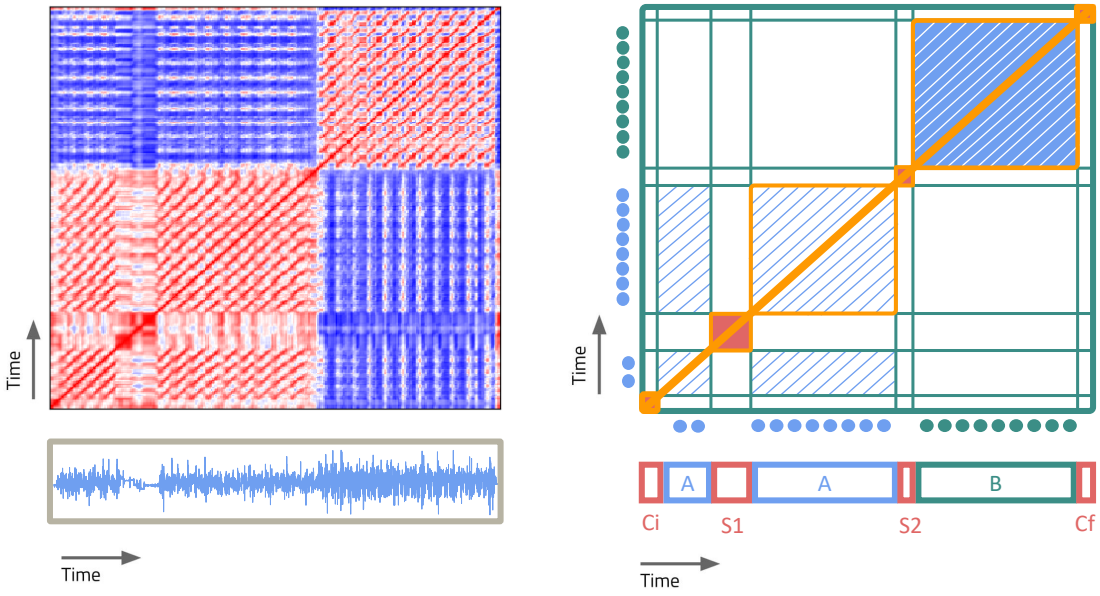


Figure 6.4: Schematic of type event 1 (work period transition) detection. The presented example comes from time serie sample of "Opr2 Wkst1&2".The left column has an SSM in the top corner and the raw data of the x coordinate of the Acc in the bottom corner. The left column presents a schematic division into blocks of the SSM, where the borders of these blocks present the events which are intended to be detected. The blocks are also coded in bottom corner with various symbols: Ci - NAW,Initial calibration; A - AW, work on Wkst1; B - AW; work onWkst2; S1 - NAW, first resting phase; S2 - NAW, second resting phase, Cf - NAW, Final calibration.

Table 6.3: Results of type event 1 (work period transition), discriminated per time serie samples. Measurements of P, R, F, A and MAE, of each according sample.

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

along the SSM main diagonal is visually evident, and the objective of this section will be to identify these point transitions between blocks, so that these events may summarize the entire time series into smaller nonoverlapping motion *states*.

Overall this process was very successful without the necessity of a very intense search, and by performing a simple manual optimization of the parameters, most events were able to be detected. The measurements of the algorithm's performance are summarized in the following table 6.3. 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 FP, which, as previously explained, are more affordable given the current context when compared to FN classifications. As such, it was given greater significance on the measurement of R against P. This means that the identified FP events just tended to divide the NAW *states* into various sub-*states*, which might be comprehensible as there might be a more complex motion description within these time instances. However, this is a manageable error, as long as there is still a clear segmentation between the AW time instance from the remaining time series, the segmentation of the NAW is irrelevant for the context of the problem. Examples of when such a thing can be seen are present at 102 s and 1062 s of "Opr1 Wkst1"(subfigure 6.5a) and in the 486 s of"Opr 2 Wkst1&2"(subfigure 6.5c). This results in a maximum recall score for all the time series samples, and a precision value high but yet imperfect for "Opr1 Wkst1", "Opr 2 Wkst1&2"and "Opr 3 Wkst1".

The time series which were provided in this database were not composed by many *states*, having only between 3 - 8 *state* motions. But like the previous section 6.1.2 there was a significant time variability within these samples transitioning between very large *states* of AW, to shorter *states* of NAW. As predictable, this fact tended to diminish the performance of the algorithm, as it was not possible to smooth areas of the data extensively because of the shorter NAW *states*. However, due to this fact, surged an interesting pattern that was not so noticeable in the previous section 6.1.2 where the parallel diagonals were much more prominently observable in the SSMs. This leads to a much noisier $Conv_f$, with a lot of high-intensity peaks constructed as a consequence of those parallel diagonals. This can be, to a degree, observed in the various figures of 6.5, where most of these peaks tend to appear in time instants of transition between parallel diagonals. As a response to this, there was a significant increase over the values of $Thresh_{frac}$ to filter out these more intense noise peaks. As those still tended to be smaller than the peak events that indicated a transition between AW and NAW blocks. However, if the NAW block which is intended to be detected is considerably short, those peak noises may acquire significantly high intensities. In the case of the sample "Opr 3 Wkst1,"this happens, resulting in the only FP point which was detected within an AW period.

Regarding the MAE value, these tended to be greater when compared with the previous section 6.1.2, being on average 11.25 s, and varying between 3.21 s and 27.77 s. This was expectable as the applied parameters were considerably larger, leading to a higher smoothing factor that distanced the events from their ground annotations. Notwithstanding, when considering that the AW period of interest tended to be about 1169.11 s and even the smaller *states* tended to be of 128.52, these MAE values can be seen as too short to be considered a significant hurdle for the analysis process.

Usually, the methodology of kernel convolution is pointed in the literature as a technique for change point detection, between *states*. In practice, what it does is to search in a SSM for two intensely consistent blocks, which are at the same time considerably different.

As already it was already seen, when fitting this database to the proposed methodology, the greatest difficulty of an optimized detection was that the SSMs constructed tended to have prominent diagonals. During a paper publication related to this thesis work, instead of this methodology, a technique more closely related to the one of the next section(whose methodology is described in section 5.4.2) was proposed [177]. Instead of constructing a convolution function $Conv_f$, it was calculated the similarity function $similarity\_funct$ and the time instances with a significantly lower similarity score would be considered as anomalies and annotated as time instances of NAW. This technique is more directed towards the analysis of motifs over SSM, and as described in the paper, it would also be able to be used under this specific context. However, it was not the methodology pursued in this work because it would not work for too similar motions. If the objective was to detect moments of transition between two similar workstations, this technique would not probably work as efficiently. Notwithstanding, it is essential to refer the potentiality of this type of processing, as it was able to detect most of the specified NAW periods with a high average accuracy of about 82.67%[177].

Overall, this has proven to be a valid proposal to segmentate the time series of interest. Despite the small amount of data, it was possible to segmentate the database with great success into the intended ground annotation, meanwhile still having a significantly low number of false classifications. Not only that, as most of these errors tended to be FP and within the NAW periods, which made them negligible inaccuracies. The greatest limitation yet to improve upon in respect to the analysis through SSM relates with the storage and time complexity of this method. The size of the SSM can cause memory errors for signals with a large number of samples, increasing the memory requirements by a quadratic function. Adding to that, the extraction of features step turns the process considerably slower. These are considerations that should be further optimized in future work.

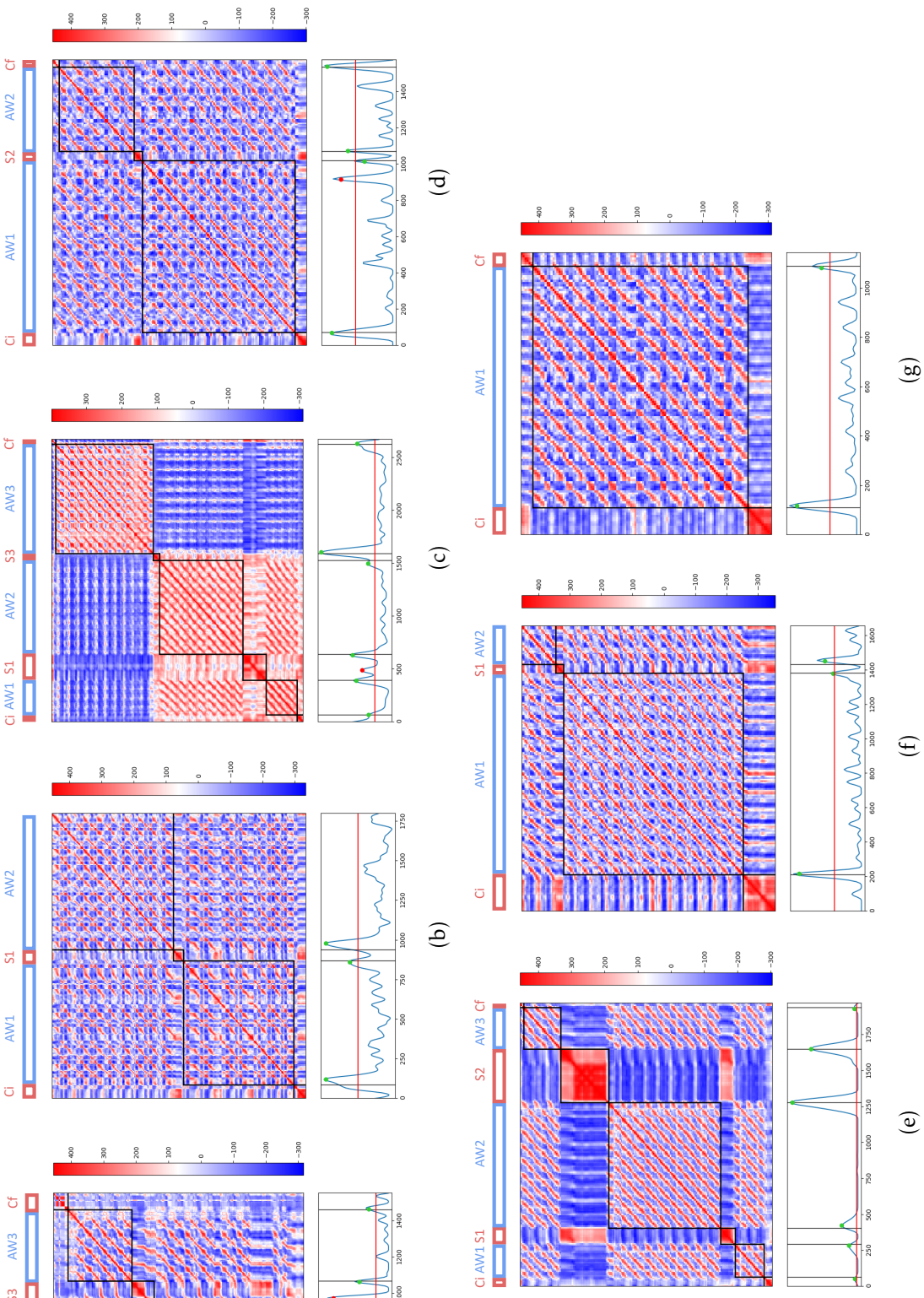


Figure 6.5: Representation Results of each time serie sample. Each subfigure represents a different time serie sample: (a) Opr1 Wkst1, (b)Opr1 Wkst2, (c)Opr2 Wkst1&2, (d)Opr3 Wkst1, (e)Opr4 Wkst1, (f)Opr5 Wkst1, (g)Opr5 Wkst2. In the center of each subfigure there is the SSM for that sample, with the top codifying the intended blocks of interest. The red blocks represent NAW, while the AW is represented by the blue blocks. The symbols "Ci"and "Cf"represent the initial and final calibration of the acquisition, respectively; the "AW"means that it is an active working period; and the "S"means that is an intermediate NAW period. The bottom depicts the respective filtered $Conv_f$ as a blue signal. The black lines present the time instances of the ground annotations B_{Ref}. The points over the peaks of the $Conv_f$ represent the estimated events B_{Ref}, with green colored representing TP instances, while the red represent FP. The red horizontal line informs about where $Thresh_{frac}$ parameter is applied.

6.2 Cycle Segmentation

The cyclic segmentation was described in section5.4.1, where there is taken advantage of how a signal described by the systematic repetition of a single motion will be represented in a SSM by several parallel diagonals across the matrix. When calculating the sum of the matrix over a single axis, what is constructed is a similarity function Sim_f where each point coordinate represents how similar that point coordinate is with the remaining time series' coordinates. This of course will be minimal in the little transition instance between the ending part of a cycle and the starting of a new one. By detecting the valleys within this Sim_f it's expected to detect all events of interest. This section will describe how well this methodology works when applied over the *HAR* database and then under the *Industrial* database. The first describes a set of data where there is a cyclic repetition of specific human motions. In the end, the analysis of the *Industrial* database will be resumed from the point it was left in section 6.1.3.

6.2.1 Evaluation metrics

Despite also performing a point event detection, unlike the previous section 6.1, the interest of this study will not necessarily insist on the detection of the instant transition, but instead on how consistent and close to the reality are the motifs that are recursively repeated along the time series. Motifs can be a subjective tool, considering the case represented in figure 6.6, it is possible to see discriminated into sub instances a work cycle from the sample "Opr2 Wkst1&2", of *Industrial* database. However, the motifs detected may not necessarily be the same as those represented in the image, and this does not mean that the detection was not well performed. In other words, instead of the motif being considered a sequence of the instances 1, 2, 3, 4, 5, 6, like it was in 6.6, the detection algorithm might segment the data in the transition from 2 to 3, returning a motif described by the sequence 3, 4, 5, 6, 1, 2. This is not 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.

The required metrics will, as such, need to be different to assess the performance of the event detection process properly. In the following section, there will be two measurements to make the analysis of this method:

1. **number of cycles detected**

2. **calculate the error between the duration of the ground truth segmentation cycles and the algorithms segmentation cycles**

The first point is somewhat evident. The focus was to assess the consistency of the found

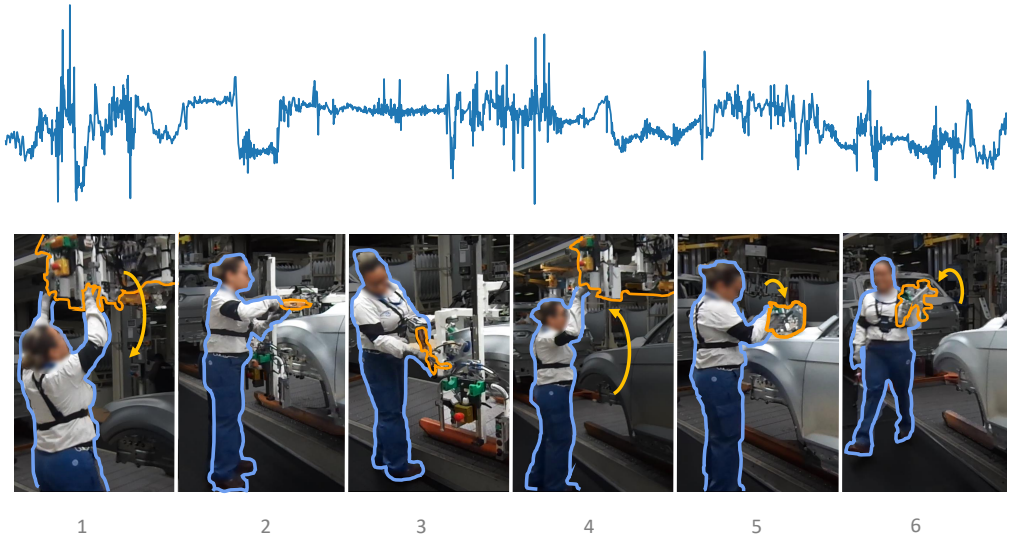


Figure 6.6: Representation of a work cycle from the time sample "Opr2 Wkst1&2", of *Industrial* database. The example represents a work cycle from the first work station Wkst1. The top region depicts the signal from the x coordinate of the Acc acquisition. The lower region has various intant representation from a video recording, displaying the various phases of work of the work cycle, during which a worker interacts with various tools and performs various minor motions.

moments. For this, there are counted the number of cycles correctly segmented. A cycle segment is considered correct if the moments at which the cycles are segmented correspond to a consistent position on the signal. Even if the segmentation of cycles occurs delayed from the ground truth-selection, what is evaluated in this category is the algorithm's consistency in defining the working cycles. Meanwhile, the second point intends to assess how different are the events found from the ground annotations. The ground-truth duration of cycles is compared with the duration of the detected cycles by calculating the absolute difference between these time intervals. This error is expressed in terms of seconds per a 100 s cycle or equivalently the duration percentage of the cycle. Overall, this last measurement intends to return how squished/extended the motifs detected compared to the ground truth.

To clarify the second metric D_e of error between the ground truth and the segmented algorithm, let us consider a B_{Est} and a B_{Ref} of equal lengths. Then the first step to retrieve the error is to associate each point of $b^{Ref} \in B_{Ref}$ to a point of$b^{Est} \in B_{Est}$. This relation is made in such a way that the error distance between all matched points D_{pp} is as minimized as possible.

$$D_{pp}(b^{Ref}, b^{Est}) = b^{Est} - b^{Ref}$$

, with b^{Ref}, b^{Est}, being the detected event and ground annotation which were matched. By expanding this definition to the organized vectors of B_{Est} and B_{Ref}, it's possible to

define the error between matched events and annotations also as a vector defined as

$$D_{pp}(B_{Ref}, B_{Est}) = \left[{b^{Est}}_0 - {b^{Ref}}_0 \quad \dots \quad {b^{Est}}_n - {b^{Ref}}_n \right]$$

, where n is the number of retrieved events. With this, the error between each cycle D_{cc} will be calculated by determining the distance by which each estimated cycle c^{Est} is dislocated from the ground truth annotation cycle c^{Ref}.

$$D_{cc}(c^{Est}, c^{Ref}) = c^{Est} - c^{Ref}$$

, where c^{Est} and c^{Ref} were cycle durations, with c^{Est} defined by the distance between two sequent b^{Est}; and c^{Ref} by two sequent b^{Ref}. As such this formula can be further simplified to

$$D_{cc}(c^{Est}, c^{Ref}) = ({b^{Est}}_{n+1} - {b^{Est}}_n) - ({b^{Ref}}_{n+1} - {b^{Ref}}_n)$$

$$D_{cc}(c^{Est}, c^{Ref}) = ({b^{Est}}_{n+1} - {b^{Ref}}_{n+1}) + ({b^{Ref}}_n - {b^{Est}}_n)$$

$$D_{cc}(c^{Est}, c^{Ref}) = D_{pp}({b^{Est}}_{n+1}, {b^{Ref}}_{n+1}) - D_{pp}({b^{Ref}}_n, {b^{Est}}_n)$$

Once again, this formula can be generalized to B_{Ref} and B_{Est} , obtaining the distance error between all considered cycles

$$D_{cc} = \left[D_{pp}({b^{Est}}_1, {b^{Ref}}_1) - D_{pp}({b^{Ref}}_0, {b^{Est}}_0) \quad \dots \quad D_{pp}({b^{Est}}_n, {b^{Ref}}_n) - D_{pp}({b^{Ref}}_{n-1}, {b^{Est}}_{n-1}) \right]$$

Having the formula of comparison of duration between the c^{Est} and c^{Ref}, it's still nonetheless an irrelevant value if it does not have any point of reference. As such the measurement of D_e will be expressed in relation to the duration of the reference cycles c^{Ref}, which are being analyzed. In the end, to unify the various error annotation we only consider the mean absolute value of this vector. In summary, the final measurement can be expressed as

$$D_e = mean\left(\left[\frac{\left| D_{pp}({b^{Est}}_1, {b^{Ref}}_1) - D_{pp}({b^{Ref}}_0, {b^{Est}}_0) \right|}{{b^{Est}}_1 - {b^{Est}}_0} \quad \dots \quad \frac{\left| D_{pp}({b^{Est}}_n, {b^{Ref}}_n) - D_{pp}({b^{Ref}}_{n-1}, {b^{Est}}_{n-1}) \right|}{{b^{Est}}_n - {b^{Est}}_{n-1}} \right] \right)$$

$$D_e = \frac{\frac{\left| D_{pp}({b^{Est}}_1, {b^{Ref}}_1) - D_{pp}({b^{Ref}}_0, {b^{Est}}_0) \right|}{{b^{Est}}_1 - {b^{Est}}_0} + \dots + \frac{\left| D_{pp}({b^{Est}}_n, {b^{Ref}}_n) - D_{pp}({b^{Ref}}_{n-1}, {b^{Est}}_{n-1}) \right|}{{b^{Est}}_n - {b^{Est}}_{n-1}}}{n} \times 100\%$$

These results will also be retrieved with a matrix profile (MP) algorithm for the purpose of comparing this method with some more established methodology. The choice of this algorithm in specific relates to the fact that the MP has already proven itself as a valid tool for detection of similar events in a workplace environment, as previously referred to in section 3.2. For the implementation of the matrix profile, it was used the stumpy [178] library.

6.2.2 Cyclic Motion Segmentation

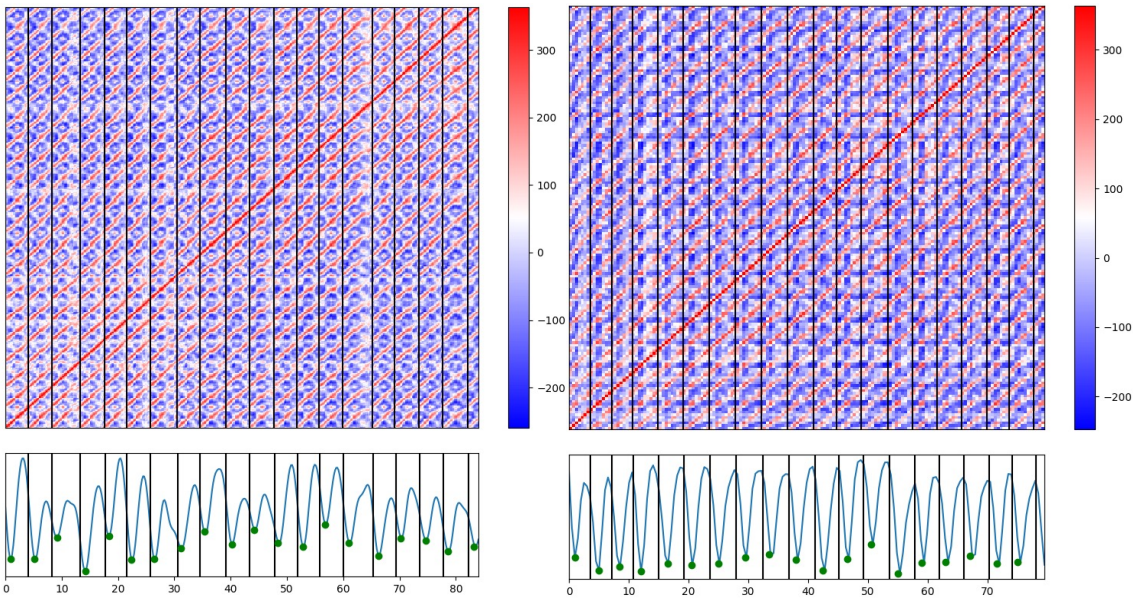
The following section will serve to present and discuss the results retrieved from the *HAR* database. The *HAR* database has about 15 different time series samples, each one representing the cyclic action of a specific human motion. This database is not as extensive as the previous *UCI* database. Still, unlike the other databases which only consider time series retrieved from Acc and Gyro sensors, each sample has a much greater dimension resultant of several biosensors retrieved during the acquisition phase (EMG sensors, electrogoniometer, airborne microphone, a piezoelectric microphone, force sensor beyond the already refereed 3-axial Acc and Gyro sensors), as it was already explained during section 4.1.

Table 6.4: Detected cycles and D_e results of the detection of type 2 events, over the *HAR* database.

TS sample	Detected Cycles	D_e(%)
1	19/19	13,25
2	19/19	8,81
3	19/19	8,82
4	19/19	5,82
5	19/19	9,73
6	19/19	6,69
7	19/19	6,18
8	19/19	5,26
9	19/19	6,66
10	19/19	6,97
11	31/31	7,22
12	19/19	8,19
13	19/19	8,96
14	19/19	12,30
15	19/19	6,81

Despite using only very simple methods of valley detection and of hyperparameters optimization, it was still a very successful technique as it identified all the intended cycles, with a considerably short D_e. This proves that the Sim_f is indeed a very useful tool, valid for the analysis of the recurrence behavior of the SSM. Regarding the values of D_e, these tended to be significantly small as well, proving that the found cycles were very similar in duration with the ground annotations. Despite this, most events were rarely intersected with the ground annotations. Instead, the most common was that all the events had some delay in relation to the ground annotations, which was perpetuated over the time series. The D_e were in average 8.56%. This means that on average, if the reference cycle c^{Ref} had 100s of duration, the cycles detected c^{Est} were either squeezed or stretched by about 8.56 s. Some examples of good detection are further displayed in figure 6.7.

These D_e values present a very small deviation and prove that the algorithm accurately detected the intended cycles. Most time-series samples had D_e values relatively close to



(a) (b)

Figure 6.7: The 2 best case scenario examples of a detection of type 2 events, in the *HAR* database. The larger squared image in the center represents the SSM of the time serie sample 4 for subfigure (a) and 8 for subfigure (b). The signal of blue color represents the respective Sim_f, the dark lines represent all the B_{Ref} instants and the green points represent TP instances of B_{Est}

this average of 8.56 %, as can be seen by a very small interquartile range of only 2 %, against the 9.9 % max-min distance. However, there were still some outliers with yet significantly higher D_e values (13,24 % in sample 1; 15,16 % in the sample 11; and 12,3 % in sample 14), which means that despite the overall success of this methodology there still needs to be taken care on the process of analysis, as there might exist some cycles detected which are less precise than others.

6.2.3 Working Cycle segmentation

Once again, the data provided from the *Industrial* database will be tested to understand if this methodology is valid for a labor context acquisition. However, unlike section 6.1.3 where the data was used in its integrity to detect the transition *state* events, this section will instead just focus on the intervals of interest. As such, by applying a mask over the time intervals of NAW, this analysis may focus only on the intended cycle segmentation. Additionally, the time series which were not considered in the analysis of section 6.1.3 will be used in this analysis. This totalizes into 11 time-series samples where six operators perform their work in 2 different workstations. The method of selection of the parameters $Wind_{len}$, $Overlap_{frac}$ and $SmoothWind_{len}$ will also be made manually.

Table 6.5 presents all the results of interest, with these also being further compared with an event detection made by the matrix profile algorithm. As previously mentioned in section 3.2, the matrix profile algorithm has proven itself to be a good technique for

Table 6.5: Detected cycles and D_e results of the detection of type 2 events, over the *Industrial* database. With the results being separated according with two different methodologies with the *SSM* being the technique described at section 5.4.2, while the other is the Matrix Profile algorithm described at section 3.2

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%

the segmentation of work cycles in a time series. As such, by attempting to compare the proposed algorithm with a more well-established methodology, it was implemented the Matrix Profile (MP) algorithm from *stumpy* python library, with the only parameter received being a time scale of the repeating pattern. For each time series sample, the average size of the working cycle was taken into consideration for the usage of the MP algorithm.

Overall, the measures considered for this analysis portray an algorithm that can indeed identify these work cycles with a high amount of certainty. Almost all the work cycles of interest were detected, as 154 cycles were detected within the 157 ground annotation cycles. The duration error was mostly good with an average value of 6.12% of the working cycle, however still significantly high in some outlier cases. These results did not change considerably between the MP, and the SSM proposed methodology. With the latter one having only one more detected cycle and with both having similar duration errors D_e, despite this last measurement being a bit better for the SSM algorithm as well. The slight difference in duration error might have occurred because of the smoothing factor used for the Matrix Profile, which was higher, and therefore increased the chance of errors in the duration of the working cycle. These considerations prove that the proposed algorithm can indeed segmentate the data retrieved from a labor context into the smaller constituent work cycles. The proposed methodology also shows that it can work for several scenarios, namely different types of workstations made by the same worker, as well as different workers performing under the same workstation, with the same performance as more well-established techniques.

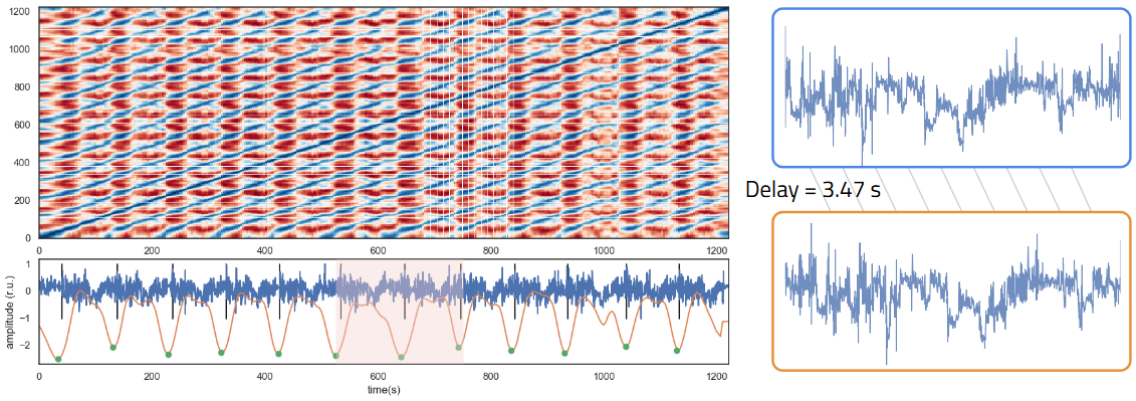


Figure 6.8: Schematic representation of a type 2 (work cycle transition) event detection.

An interesting characteristic of the final results is a delay between the detected cycles and their respective ground annotations. Through a visual assessment of the example of Figure 6.8, there can be observed that despite the identification of the cycles being very consistent, along with the time series sample, there is a constant delay between the point events identified and the actual ground annotations. Resulting in cycles with different margins, as illustrated in the right half of Figure 6.8, where the first and last 3.47 s of the signals are different from what was expected. This delay is comprehensible considering that the unsupervised algorithm does not have a reference of where the cycle has the ”real” start. The algorithms consider the beginning of the data as a reference, which would not always match with the instant the operator would start the cycle. The lag detected does not affect the detection of the entire cyclic information since both algorithms are consistent with their decision of where the ”start” is and are able to identify the cyclic pattern.

When considering all the time series samples, the average value of Duration error D_e is acceptably small, only being 6.12% of the working cycle. In other words, if we consider a working period with cycles of 100 seconds, there would be, on average, a difference of about 6 seconds between the detected cycles and the ground annotations. This value is considerably small to do not interfere significantly with the detection process. However, there are some time series with significant values, the highest one being 15.83%, for the Opr1 Wkst2. The major contributions to this error might have come from the loss of resolution when extracting features or when smoothing the similarity function to detect the minimums. Moreover, like in the previous section 6.1.3, the ground annotations probably are not perfect as they were annotated from video footage of the performed work. This might have created errors when transposing the annotations to the time series, also increased because as the cycles are considerably short, then any significant divergence would seriously affect the end result.

6.3 Search With Query

This final section intends to display how by making an analysis based on the SSM there is also another promising alternative analysis. To evaluate occupational exposure in a manufacturing scenario, it is relevant to analyze specific motion and human postures. Some human motions, due to their unergonomic posture, intensity, duration, or frequent repetition, may display a serious risk for musculoskeletal disorders. The methodology of search by example proposed in section 5.4.3 will further be tested in the *Industrial* database to assess its validity.

The actual detection entails constructing a distance function $Dist_f$, which describes how similar the example sub-sequence is to the rest of the working period. Then, the remaining process just requires a valley detection, where the points of more significant similarity are annotated. As this method is based on the match of equal-sized columns, it's expected that this event point will indicate the start of the motion of interest, with the duration being the same for every repetition of this motion.

In the example presented in Figure 6.9, it is demonstrated this process. The top image represents the SSM built from the time series sample "Opr2 Wkst1&2". In blue, there is illustrated one dimension of this multivariate time series. The signal in orange is the same as the previous section, representing the similarity function, and the minimums signaled with green dots describe the work cycle transitions. The grey signal represents the aforementioned distance function $Dist_f$, with the minimums indicating the events of interest. The green areas represent the time intervals where the motion of interest is repeated, always being present in the same position of the working cycle, as expected. The red window indicates the time interval, which served as an example for the query.

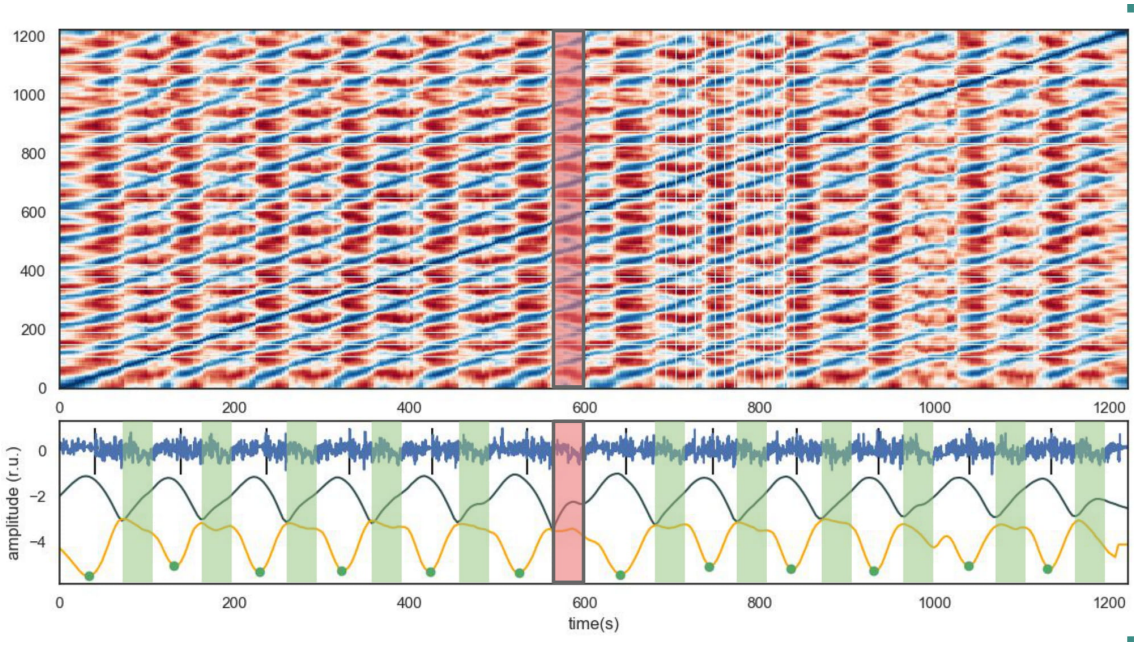


Figure 6.9: Schematic of type event 3 (sub-segment example search) detection

Chapter 7

Conclusions and Future Work

In this chapter will be presented the main conclusions that derived from the current work and some prospects for future works in this research line. As such, the structure of this chapter is divided into two sections: section 7.1 for the main conclusions, and section 7.2 for future work suggestions.

7.1 Conclusions

There has been an increased interest in the field of occupational medicine. This interest has surged in part due to the high prevalence of MSDs in society, along with the increased burden on the well-being that these conditions bring to subjects. Additionally, it is also due to a realization that monitorization and job analysis are essential tools to implement intervention strategies, which have proven to significantly reduce the prevalence of MSDs and their comorbidities. [179] This thesis aimed to tackle one of the issues that is still lacking for a complete implementation of direct measurement of risk assessment in the work environment. To promote the use of direct sensing techniques, this work proposed a methodology of event detection that structures the data acquired by inertial sensors into a more comprehensive representation to be used by professionals in the field of ergonomy. Through a pre-segmentation of the periods of interest, it is expected to facilitate the hard work of analysis over an entire database and instead to make the use of inertial sensors in a work setting as a more approachable option. The methods used were inspired from the sound analysis domain into human motion analysis field, an approach which until recently had not been extensively explored. Through processing methodologies applied over the SSM, it was possible to detect the three event types of interest with success. With these having been tested in a database retrieved from a real-world manufacturing scenario and by two external sources of data with simple representations

of human motion.

The objective predefined for this thesis was the design of an unsupervised algorithm of event detection to be applied under a labor context. In the previous section 1.3, there were described in more detail a list of requirements necessary for this algorithm to have. With the completion of this thesis, it is possible to conclude how well those requirements were accomplished:

Agnostic to time series type: The most tested type of data was inertial sensing data, as the two of the three databases used (*UCI* and *Industrial* databases) only Acc and Gyro time series. However, the *HAR* database used a set of different types of data which included EMG. Not only this type of data seemed not to affect the analysis process in any negative way, but it might have been a positive addition as the *HAR* database had some of the best performing results. As SSM based analysis measures how each point of time is similar to the remaining signal, it is actually expected that by describing each time instant with multiple measurements that then the comparison between time instances will be closer to reality, and that the algorithm will gain with that.

Multidimensionality: A caution which was taken throughout the construction of the algorithm was how the various dimensions of the data could be analyzed at the same time. Each algorithm step had various processing steps where it actively tried to vectorize the more time-demanding operations along the various dimensions of the data. When designing the feature retrieval, the subsegments of each window iteration were segmented from the original data and overlayed before actual retrieving the features from all the windows simultaneously; when constructing the SSM, it was necessary to compare various segments and to facilitate this procedure it was made the algebraic dot product between vectors. In the actual process of detection, independently of the dimension of the SSM, the detection is simplified to the analysis of a single univariate signal, $Conv_f$, Sim_f or $Dist_f$ depending on what is the event of interest. This was also possible due to the *numpy* and *scipy*[180] python libraries, which allow for the construction of dense matrices and to perform algebraic functions over them in an optimized manner, and also with the usage of the *libfmp* library, which provided methods to compute the SSM and extract relevant information from it[181].

Time-Efficiency: The procedure is still slow and needs to be otimized. This was a great limitation observed during the results as an increase in the dimension of the data was accompanied by a significant increase in time demands as well. Despite not being a highly limiting time demand for the current context, it still is a detail that will require improvement in future work. This issue is also associated with the exponential increase in memory for the computation of the SSM.

Interpretability: This criterium was fulfilled during the definition of three event types (work period transition, work cycle transition, and sub-segment example search). The three events were selected with the final objective of structuring the work period into smaller sections that could be more easily understood while also providing a tool that segments the periods of interest, with the criterium of interest still remaining adjustable

by the user.

Additionally, the results retrieved by applying the algorithm in different database settings were overall very good. This proves that the algorithm is not only valid to be applied in this context but that it is also malleable enough to possibly work in different work settings. Due to the current pandemic of Covid-19, it was impossible to extend the acquisition database already provided by the work of [164]. As such, the testing of the algorithm in a real-work scenario was only made for a relatively small database of 11 time series samples. Despite the retrieved results being significantly high, the testing of this methodology should still be extended to more and different work settings in future studies.

In conclusion, this work has fulfilled most of the required expectations which were initially proposed. Successfully identifying the events of 1) work period transition; 2) work cycle transition; and 3) sub-segment example search. The application of this method in the ergonomic context can be of great interest since it can improve the current approaches of ergonomic evaluations in these scenarios. This strategy turns the process more flexible, allowing to identify in the working cycle sources of risk factors. Moreover, this allows not only to compare the occupational exposure of different workstations for the same worker but also to compare the occupational exposure throughout the working period. Ultimately, this method can decrease the workload associated with the manual identifying of working cycles and anomalies while improving the accuracy of the evaluation.

Overall this algorithm constructs a good base of work to further built on, if intended to construct a summarizing algorithm focused towards real-work acquisitions. The most significant utility provided by this algorithm is the automation of processes related with the identification of periods of relevance, segmentation of the signal, and further identification of periodic regions. Moreover, the detection of events for the description of patterns in a labor acquisition is something yet scarce in the literature, and the proposed analysis by means of SSM seems to be an innovative approach with the potential to be expanded upon.

7.2 Future Work

This work has been a promising step into the objective of creating analysis methodologies that could facilitate the analysis of ergonomic data. However, it is still work that could be expanded upon. As referred in the previous section, the greatest limitation with the proposed methodology relates with the time and storage demand and, as such, the development of optimization strategies could be considered so to surpass these constraints. A simple methodology would be a reduction of the number of parameters and of features selected for the algorithm.

Another consideration which could be made would be to further expand the technique to other types of processing to test if there is something which could be improved upon. An example could be a change in the similarity function, because as seen in section 3.3.3

there could also be applied instead the Euclidean, Cosine between angles or Kullback Leibler distances. Another example could be the feature selection. Multiple features were used in the context of this work, as it is the normal procedure when dealing with shallow learning of human motion data. However, these features could be expanded upon and through reduction techniques like PCA, and understand which of them are more useful for the type of signals acquired in a working labour context.

Introduced in section 2.2.1, it was explained the multiple dimensions of search that an algorithm may focus on. As this proposal focused exclusively on a uni scale analysis, an interesting improvement would be expand to a multi scale analysis, where there could be analysed the recurrence of events adapted to the window scale. This would allow for a more complete analysis of the events, where more events of interest could identified.

Another interesting path would be to further classify the segmentated motion states, as well as the found events.

Bibliography

[1] The Editors of Encyclopaedia Britannica. *Occupational medicine*. 2018. URL: https://www.britannica.com/science/occupational-medicine (visited on 09/07/2020).

[2] *Occupational Medicine | American Medical Association*. URL: https://www.ama-assn.org/specialty/occupational-medicine (visited on 09/07/2020).

[3] American Board of Medical Specialties. *Guide To Medical Specialties 2020*. Tech. rep. 2020. URL: https://www.abms.org/wp-content/uploads/2020/11/ABMS-Guide-to-Medical-Specialties-2020.pdf.

[4] J. Burton, W. H. Organization, and Others. *WHO Healthy workplace framework and model: Background and supporting literature and practices*. World Health Organization, 2010.

[5] World Health Organization. "WHO Global Plan of Action on Workers Health (2008-2017)." In: (2013). URL: http://bit.ly/146pn27.

[6] M. R. Cullen, M. G. Cherniack, and L. Rosenstock. *Occupational medicine: First of Two Parts*. 1990. DOI: 10.1056/NEJM199003013220905. URL: http://www.nejm.org/doi/abs/10.1056/NEJM199003013220905.

[7] M. R. Cullen, M. G. Cherniack, and L. Rosenstock. *Occupational Medicine: second of Two Parts*. 1990. DOI: 10.1056/NEJM199003083221007. URL: http://www.nejm.org/doi/abs/10.1056/NEJM199003083221007.

[8] World Health Organization. *Musculoskeletal conditions*. URL: https://www.who.int/news-room/fact-sheets/detail/musculoskeletal-conditions (visited on 08/25/2020).

[9] 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. URL: /pmc/articles/PMC2757646/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC2757646/.

[10] 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. URL: https://dx.plos.org/10.1371/journal.pone.0090633.

[11] 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. URL: https://pubmed.ncbi.nlm.nih.gov/22071318/.

[12] 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. URL: https://link.springer.com/articles/10.1186/1471-2474-10-143https://link.springer.com/article/10.1186/1471-2474-10-143.

[13] 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. URL: http://ard.bmj.com/.

[14] United States Bone and Joint Initiative (USBJI). *The Burden of Musculoskeletal Diseases in the United States*. 4th. 2016. URL: https://www.boneandjointburden.org/print/book/export/html/964.

[15] *A problem worth solving: the rising cost of musculoskeletal conditions in Australia*. 2013. DOI: https://doi.org/APO-35957. URL: https://apo.org.au/node/35957.

[16] D. Bliuc, N. D. Nguyen, V. E. Milch, T. V. Nguyen, J. A. Eisman, and J. R. Center. "Mortality risk associated with low-trauma osteoporotic fracture and subsequent fracture in men and women." In: *JAMA - Journal of the American Medical Association* 301.5 (2009), pp. 513–521. ISSN: 00987484. DOI: 10.1001/jama.2009.50. URL: https://pubmed.ncbi.nlm.nih.gov/19190316/.

[17] S. Dadoun, N. Zeboulon-Ktorza, C. Combescure, M. Elhai, S. Rozenberg, L. Gossec, and B. Fautrel. "Mortality in rheumatoid arthritis over the last fifty years: Systematic review and meta-analysis." In: *Joint Bone Spine* 80.1 (2013), pp. 29–33. ISSN: 1297319X. DOI: 10.1016/j.jbspin.2012.02.005. URL: https://pubmed.ncbi.nlm.nih.gov/22459416/.

[18] E. Nüesch, P. Dieppe, S. Reichenbach, S. Williams, S. Iff, and P. Jüni. "All cause and disease specific mortality in patients with knee or hip osteoarthritis: Population based cohort study." In: *BMJ* 342.7798 (2011), p. 638. ISSN: 17561833. DOI: 10.1136/bmj.d1165. URL: http://www.bmj.com/.

[19] J. Widdifield, S. Bernatsky, J. M. Paterson, G. Tomlinson, K. Tu, B. Kuriya, J. C. Thorne, J. E. Pope, S. Hollands, and C. Bombardier. "Trends in excess mortality among patients with rheumatoid arthritis in Ontario, Canada." In: *Arthritis Care and Research* 67.8 (2015), pp. 1047–1053. ISSN: 21514658. DOI: 10.1002/acr.22553. URL: https://pubmed.ncbi.nlm.nih.gov/25623141/.

[20] R. Cooper, D. Kuh, and R. Hardy. *Objectively measured physical capability levels and mortality: Systematic review and meta-analysis*. 2010. DOI: 10.1136/bmj.c4467. URL: https://pubmed.ncbi.nlm.nih.gov/20829298/.

[21] *GBD Results Tool | GHDx*. URL: http://ghdx.healthdata.org/gbd-results-tool (visited on 09/29/2020).

[22] T. Vos et al. "Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015." In: *The Lancet* (2016). ISSN: 1474547X. DOI: 10.1016/S0140-6736(16)31678-6.

[23] A. Luttmann, M. Jager, B. Griefahn, G. Caffier, F. Liebers, and W. H. Organization. "Preventing musculoskeletal disorders in the workplace." In: (2003). ISSN: 924159053X.

[24] D. Hoy, J. A. Geere, F. Davatchi, B. Meggitt, and L. H. Barrero. *A time for action: Opportunities for preventing the growing burden and disability from musculoskeletal conditions in low- and middleincome countries*. 2014. DOI: 10.1016/j.berh.2014.07.006. URL: https://pubmed.ncbi.nlm.nih.gov/25481422/.

[25] Eurostat. *Self-reported work-related health problems and risk factors - key statistics - Statistics Explained*. 2017. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php/Self-reported{_}work-related{_}health{_}problems{_}and{_}risk{_}factors{_}-{_}key{_}statistics (visited on 08/24/2020).

[26] J. Lundkvist, F. Kastäng, and G. Kobelt. "The burden of rheumatoid arthritis and access to treatment: Health burden and costs." In: *European Journal of Health Economics* 8.SUPPL. 8 (2008). ISSN: 16187598. DOI: 10.1007/s10198-007-0088-8. URL: https://pubmed.ncbi.nlm.nih.gov/18157732/.

[27] Z. Gledović. "Epidemiology, Prevention." In: *Encyclopedia of Public Health*. Ed. by W. Kirch. Dordrecht: Springer Netherlands, 2008, pp. 354–356. ISBN: 978-1-4020-5614-7. DOI: 10.1007/978-1-4020-5614-7_1008. URL: https://doi.org/10.1007/978-1-4020-5614-7{_}1008.

[28] 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. URL: https://reliefweb.int/report/world/preventing-disease-through-healthy-environments-global-assessment.

[29] D. H. Seidel, D. M. Ditchen, U. M. Hoehne-Hückstädt, M. A. Rieger, and B. Steinhilber. *Quantitative measures of physical risk factors associated with work-related musculoskeletal disorders of the elbow: A systematic review*. 2019. DOI: 10.3390/ijerph16010130. URL: /pmc/articles/PMC6339038/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC6339038/.

[30] A. Leclerc, M. F. Landre, J. F. Chastang, I. Niedhammer, and Y. Roquelaure. "Upper-limb disorders in repetitive work." In: *Scandinavian Journal of Work, Environment and Health* 27.4 (2001), pp. 268–278. ISSN: 03553140. DOI: 10.5271/sjweh.614. URL: https://pubmed.ncbi.nlm.nih.gov/11560341/.

[31] A. M. Briggs, M. J. Cross, D. G. Hoy, L. Sànchez-Riera, F. M. Blyth, A. D. Woolf, and L. March. "Musculoskeletal Health Conditions Represent a Global Threat to Healthy Aging: A Report for the 2015 World Health Organization World Report on Ageing and Health." In: *The Gerontologist* 56.Suppl_2 (2016), S243–S255. ISSN: 0016-9013. DOI: 10.1093/geront/gnw002. URL: https://doi.org/10.1093/geront/gnw002.

[32] S. Roberts, P. Colombier, A. Sowman, C. Mennan, J. H. Rölfing, J. Guicheux, and J. R. Edwards. "Ageing in the musculoskeletal system: Cellular function and dysfunction throughout life." In: *Acta Orthopaedica* (2016). ISSN: 17453682. DOI: 10.1080/17453674.2016.1244750.

[33] J. M. Wolf, L. Cannada, A. E. Van Heest, M. I. O'Connor, and A. L. Ladd. "Male and Female Differences in Musculoskeletal Disease." In: *Journal of the American Academy of Orthopaedic Surgeons* 23.6 (2015), pp. 339–347. ISSN: 1067-151X. DOI: 10.5435/JAAOS-D-14-00020. URL: http://journals.lww.com/00124635-201506000-00003.

[34] A. M. Al-Bashaireh, L. G. Haddad, M. Weaver, D. L. Kelly, X. Chengguo, and S. Yoon. *The Effect of Tobacco Smoking on Musculoskeletal Health: A Systematic Review*. 2018. DOI: 10.1155/2018/4184190. URL: /pmc/articles/PMC6077562/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC6077562/.

[35] K. T. Palmer, H. Syddall, C. Cooper, and D. Coggon. "Smoking and musculoskeletal disorders: Findings from a British national survey." In: *Annals of the Rheumatic Diseases* 62.1 (2003), pp. 33–36. ISSN: 00034967. DOI: 10.1136/ard.62.1.33. URL: http://ard.bmj.com/.

[36] M. Abate, D. Vanni, A. Pantalone, and V. Salini. *Cigarette smoking and musculoskeletal disorders*. 2013. DOI: 10.11138/mltj/2013.3.2.063. URL: /pmc/articles/PMC3711704/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC3711704/.

[37] G. W. Gaddini, R. T. Turner, K. A. Grant, and U. T. Iwaniec. *Alcohol: A Simple Nutrient with Complex Actions on Bone in the Adult Skeleton*. 2016. DOI: 10.1111/acer.13000. URL: https://pubmed.ncbi.nlm.nih.gov/26971854/.

[38] J. A. Kanis, H. Johansson, O. Johnell, A. Oden, C. De Laet, J. A. Eisman, H. Pols, and A. Tenenhouse. "Alcohol intake as a risk factor for fracture." In: *Osteoporosis International* 16.7 (2005), pp. 737–742. ISSN: 14332965. DOI: 10.1007/s00198-004-1734-y. URL: https://pubmed.ncbi.nlm.nih.gov/15455194/.

[39] A. D. Woolf, F. C. Breedveld, and T. K. Kvien. *Controlling the obesity epidemic is important for maintaining musculoskeletal health*. 2006. DOI: 10.1136/ard.2006.058172. URL: /pmc/articles/PMC1798351/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC1798351/.

[40] A. Anandacoomarasamy, I. Caterson, P. Sambrook, M. Fransen, and L. March. *The impact of obesity on the musculoskeletal system*. 2008. DOI: 10.1038/sj.ijo.0803715. URL: https://pubmed.ncbi.nlm.nih.gov/17848940/.

[41] R. Shiri, J. Karppinen, P. Leino-Arjas, S. Solovieva, and E. Viikari-Juntura. "The association between obesity and low back pain: A meta-analysis." In: *American Journal of Epidemiology* 171.2 (2010), pp. 135–154. ISSN: 00029262. DOI: 10.1093/aje/kwp356. URL: https://pubmed.ncbi.nlm.nih.gov/20007994/.

[42] M. Collins and S. M. Raleigh. "Genetic risk factors for musculoskeletal soft tissue injuries." In: *Medicine and Sport Science* 54 (2009), pp. 136–149. ISSN: 02545020. DOI: 10.1159/000235701. URL: https://pubmed.ncbi.nlm.nih.gov/19696512/.

[43] National Research Council (US) and Institute of Medicine (US) Panel on Musculoskeletal Disorders and the Workplace. *Musculoskeletal Disorders and the Workplace*. National Academies Press, 2001. DOI: 10.17226/10032. URL: https://www.ncbi.nlm.nih.gov/books/NBK222440/.

[44] N. R. Council. *Work-Related Musculoskeletal Disorders*. National Academies Press, 1999. DOI: 10.17226/6431.

[45] S. Borjanović. "ErgonomicsErgonomics." In: *Encyclopedia of Public Health*. Ed. by W. Kirch. Dordrecht: Springer Netherlands, 2008, pp. 357–360. ISBN: 978-1-4020-5614-7. DOI: 10.1007/978-1-4020-5614-7_1019. URL: https://doi.org/10.1007/978-1-4020-5614-7{_}1019.

[46] J. O. Crawford. "The Nordic Musculoskeletal Questionnaire." In: *Occupational Medicine* 57.4 (2007), pp. 300–301. ISSN: 0962-7480. DOI: 10.1093/occmed/kqm036. URL: https://academic.oup.com/occmed/article/57/4/300/2751338.

[47] I. Kuorinka, B. Jonsson, A. Kilbom, H. Vinterberg, F. Biering-Sørensen, G. Andersson, and K. Jørgensen. "Standardised Nordic questionnaires for the analysis of musculoskeletal symptoms." In: *Applied Ergonomics* 18.3 (1987), pp. 233–237. ISSN: 00036870. DOI: 10.1016/0003-6870(87)90010-X. URL: https://pubmed.ncbi.nlm.nih.gov/15676628/.

[48] G. C. David. *Ergonomic methods for assessing exposure to risk factors for work-related musculoskeletal disorders*. 2005. DOI: 10.1093/occmed/kqi082. URL: https://pubmed.ncbi.nlm.nih.gov/15857898/.

[49] D. Wang, F. Dai, and X. Ning. "Risk Assessment of Work-Related Musculoskeletal Disorders in Construction: State-of-the-Art Review." In: *Journal of Construction Engineering and Management* 141.6 (2015), p. 04015008. ISSN: 0733-9364. DOI: 10.1061/(asce)co.1943-7862.0000979. URL: https://ascelibrary.org/doi/abs/10.1061/{\%}28ASCE{\%}29CO.1943-7862.0000979.

[50] E. Viikari-Juntura, S. Rauas, R. Martikainen, E. Kuosma, H. Riihimäki, E. P. Takala, and K. Saarenmaa. "Validity of self-reported physical work load in epidemiologic studies on musculoskeletal disorders." In: *Scandinavian Journal of Work, Environment and Health* 22.4 (1996), pp. 251–259. ISSN: 03553140. DOI: 10.5271/sjweh.139. URL: https://pubmed.ncbi.nlm.nih.gov/8881013/.

[51] G. Å Hansson, I. Balogh, J. Unge Byström, K. Ohlsson, C. Nordander, P. Asterland, S. Sjölander, L. Rylander, J. Winkel, and S. Skerfving. "Questionnaire versus direct technical measurements in assessing postures and movements of the head, upper back, arms and hands." In: *Scandinavian Journal of Work, Environment and Health* 27.1 (2001), pp. 30–40. ISSN: 03553140. DOI: 10.5271/sjweh.584. URL: https://pubmed.ncbi.nlm.nih.gov/11266144/.

[52] O. Karhu, P. Kansi, and I. Kuorinka. "Correcting working postures in industry: A practical method for analysis." In: *Applied Ergonomics* 8.4 (1977), pp. 199–201. ISSN: 00036870. DOI: 10.1016/0003-6870(77)90164-8. URL: https://pubmed.ncbi.nlm.nih.gov/15677243/.

[53] L. McAtamney and E. Nigel Corlett. "RULA: a survey method for the investigation of work-related upper limb disorders." In: *Applied Ergonomics* 24.2 (1993), pp. 91–99. ISSN: 00036870. DOI: 10.1016/0003-6870(93)90080-S. URL: https://pubmed.ncbi.nlm.nih.gov/15676903/.

[54] S. Hignett and L. McAtamney. "Rapid entire body assessment (REBA)." In: *Applied ergonomics* 31.2 (2000), pp. 201–205. ISSN: 0003-6870.

[55] K. Schaub, G. Caragnano, B. Britzke, and R. Bruder. "The European Assembly Worksheet." In: *Theoretical Issues in Ergonomics Science* 14.6 (2013), pp. 616–639. ISSN: 1464536X. DOI: 10.1080/1463922X.2012.678283. URL: https://www.tandfonline.com/doi/abs/10.1080/1463922X.2012.678283.

[56] J. S. Moore and A. Garg. "The strain index: A proposed method to analyze jobs for risk of distal upper extremity disorders." In: *American Industrial Hygiene Association Journal* 56.5 (1995), pp. 443–458. ISSN: 00028894. DOI: 10.1080/15428119591016863. URL: https://pubmed.ncbi.nlm.nih.gov/7754975/.

[57] J. Winkel and S. E. Mathiassen. "Assessment of physical work load in epidemiologic studies: Concepts, issues and operational considerations." In: *Ergonomics* 37.6 (1994), pp. 979–988. ISSN: 13665847. DOI: 10.1080/00140139408963711. URL: https://pubmed.ncbi.nlm.nih.gov/8026455/.

[58] U.S. Bureau of Labor Statistics. *AMERICAN TIME USE SURVEY-2019 RESULTS*. Tech. rep. U. S. Departament of Labor, 2020. URL: www.bls.gov/tus/data.htm..

[59] Eurostat. *Duration of working life - statistics - Statistics Explained*. 2020. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php/Duration{_}of{_}working{_}life{_}-{_}statistics (visited on 10/17/2020).

[60] Eurostat. *Hours of work - annual statistics - Statistics Explained*. 2020. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Hours{_}of{_}work{_}-{_}annual{_}statistics{\#}Working{_}hours{_}for{_}employees (visited on 10/17/2020).

[61] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth. *From Data Mining to Knowledge Discovery in Databases*. Tech. rep. 3. 1996, pp. 37–37. DOI: 10.1609/AIMAG.V17I3.1230. URL: www.ffly.com/.

[62] M. Hong, A. Demers, J. Gehrke, and M. Riedewald. "Event and Pattern Detection over Streams." In: *Encyclopedia of Database Systems*. Ed. by L. LIU and M. T. ÖZSU. Boston, MA: Springer US, 2009, pp. 1029–1033. ISBN: 978-0-387-39940-9. DOI: 10.1007/978-0-387-39940-9_155. URL: https://doi.org/10.1007/978-0-387-39940-9{_}155.

[63] V. Cerqueira, L. Torgo, and C. Soares. "Early Anomaly Detection in Time Series: A Hierarchical Approach for Predicting Critical Health Episodes." In: (2020). arXiv: 2010.11595. URL: http://arxiv.org/abs/2010.11595.

[64] B. Liu, J. Li, C. Chen, W. Tan, Q. Chen, and M. Zhou. "Efficient motif discovery for large-scale time series in healthcare." In: *IEEE Transactions on Industrial Informatics* 11.3 (2015), pp. 583–590. ISSN: 15513203. DOI: 10.1109/TII.2015.2411226.

[65] A. Balasubramanian, J. Wang, and B. Prabhakaran. "Discovering Multidimensional Motifs in Physiological Signals for Personalized Healthcare." In: *IEEE Journal on Selected Topics in Signal Processing* 10.5 (2016), pp. 832–841. ISSN: 19324553. DOI: 10.1109/JSTSP.2016.2543679.

[66] S. Yazdanirad, A. H. Khoshakhlagh, E. Habibi, A. Zare, M. Zeinodini, and F. Dehghani. "Comparing the Effectiveness of Three Ergonomic Risk Assessment Methods-RULA, LUBA, and NERPA-to Predict the Upper Extremity Musculoskeletal Disorders." eng. In: *Indian journal of occupational and environmental medicine* 22.1 (2018), pp. 17–21. ISSN: 0973-2284 (Print). DOI: 10.4103/ijoem.IJOEM_23_18.

[67] E. J. Ciaccio, S. M. Dunn, and M. Akay. "Biosignal Pattern Recognition and Interpretation Systems." In: *IEEE Engineering in Medicine and Biology Magazine* 12.3 (1993), pp. 89–95. ISSN: 07395175. DOI: 10.1109/51.232348.

[68] L. S. Smith, S Shahid, A Vernier, N Mtetwa, and U. D. Franche-comté. "Finding events in noisy signals." In: *Proc. of Irish Signals and Systems Conference*. 2007, pp. 31–35.

[69] A. Ericsson, M. Berndtsson, and J. Mellin. "Event in Active Databases." In: *Encyclopedia of Database Systems*. Springer US, 2009, pp. 1044–1045. DOI: 10.1007/978-0-387-39940-9_512. URL: https://link.springer.com/referenceworkentry/10.1007/978-0-387-39940-9{_}512.

[70] B. M. Michelson. "Event-driven architecture overview." In: *Patricia Seybold Group* 2.12 (2006), pp. 10–1571.

[71] D. Margineantu, W.-K. Wong, and D. Dash. "Machine learning algorithms for event detection." In: *Machine Learning* 79.3 (2010), p. 257. ISSN: 0885-6125.

[72] A. Gensler and B. Sick. "Performing event detection in time series with SwiftEvent: an algorithm with supervised learning of detection criteria." In: *Pattern Analysis and Applications* 21.2 (2018), pp. 543–562. ISSN: 14337541. DOI: 10.1007/s10044-017-0657-0.

[73] F Xie, A Song, and V Ciesielski. "Event detection in time series by genetic programming." In: *2012 IEEE Congress on Evolutionary Computation*. 2012, pp. 1–8. DOI: 10.1109/CEC.2012.6256589.

[74] S. Aminikhanghahi and D. J. Cook. "A survey of methods for time series change point detection." In: *Knowledge and Information Systems* 51.2 (2017), pp. 339–367. ISSN: 02193116. DOI: 10.1007/s10115-016-0987-z.

[75] V. Guralnik and J. Srivastava. "Event detection from time series data." In: *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*. 1999, pp. 33–42.

[76] R. Sheikhrabori, M. Aminnayeri, and M. Ayoubi. "Maximum likelihood estimation of change point from stationary to nonstationary in autoregressive models using dynamic linear model." In: *Quality and Reliability Engineering International* 34.1 (2018), pp. 27–36. ISSN: 0748-8017.

[77] S Li, Y Cao, C Leamon, Y Xie, L Shi, and W Song. "Online seismic event picking via sequential change-point detection." In: *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. 2016, pp. 774–779. ISBN: VO -. DOI: 10.1109/ALLERTON.2016.7852311.

[78] Q. Wang, J. Tang, J. Zeng, S. Leng, and W. Shui. "Regional detection of multiple change points and workable application for precipitation by maximum likelihood approach." In: *Arabian Journal of Geosciences* 12.23 (2019), p. 745. ISSN: 1866-7538. DOI: 10.1007/s12517-019-4790-5. URL: https://doi.org/10.1007/s12517-019-4790-5.

[79] A. Ihler, J. Hutchins, and P. Smyth. "Adaptive Event Detection with Time-Varying Poisson Processes." In: *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '06. New York, NY, USA: Association for Computing Machinery, 2006, pp. 207–216. ISBN: 1595933395. DOI: 10.1145/1150402.1150428. URL: https://doi.org/10.1145/1150402.1150428.

[80] F. F. do Nascimento and W. V. Moura e Silva. "A Bayesian model for multiple change point to extremes, with application to environmental and financial data." In: *Journal of Applied Statistics* 44.13 (2017), pp. 2410–2426. ISSN: 0266-4763. DOI: 10.1080/02664763.2016.1254733. URL: https://doi.org/10.1080/02664763.2016.1254733.

[81] D. Preston, P. Protopapas, and C. Brodley. "Event discovery in time series." In: *Society for Industrial and Applied Mathematics - 9th SIAM International Conference on Data Mining 2009, Proceedings in Applied Mathematics*. Vol. 1. 2009, pp. 60–71. ISBN: 9781615671090. DOI: 10.1137/1.9781611972795.6. arXiv: 0901.3329.

[82] I. Cribben. "Change points in heavy-tailed multivariate time series: Methods using precision matrices." In: *Applied Stochastic Models in Business and Industry*. Vol. 35. 2. John Wiley and Sons Ltd, 2019, pp. 299–320. DOI: 10.1002/asmb.2373. URL: https://onlinelibrary.wiley.com/doi/full/10.1002/asmb.2373https://onlinelibrary.wiley.com/doi/abs/10.1002/asmb.2373https://onlinelibrary.wiley.com/doi/10.1002/asmb.2373.

[83] K. D. Feuz, D. J. Cook, C. Rosasco, K. Robertson, and M. Schmitter-Edgecombe. "Automated Detection of Activity Transitions for Prompting." eng. In: *IEEE transactions on human-machine systems* 45.5 (2015), pp. 575–585. ISSN: 2168-2291. DOI: 10.1109/THMS.2014.2362529. URL: https://pubmed.ncbi.nlm.nih.gov/27019791https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4805372/.

[84] F Desobry, M Davy, and C Doncarli. "An online kernel change detection algorithm." In: *IEEE Transactions on Signal Processing* 53.8 (2005), pp. 2961–2974. DOI: 10.1109/TSP.2005.851098.

[85] F. Li, G. C. Runger, and E. Tuv. "Supervised learning for change-point detection." In: *International Journal of Production Research* 44.14 (2006), pp. 2853–2868. ISSN: 0020-7543. DOI: 10.1080/00207540600669846. URL: https://doi.org/10.1080/00207540600669846.

[86] A. Mueen. "Time series motif discovery: dimensions and applications." In: *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 4.2 (2014), pp. 152–159. ISSN: 19424787. DOI: 10.1002/widm.1119. URL: http://doi.wiley.com/10.1002/widm.1119.

[87] R. Agrawal, C. Faloutsos, and A. Swami. "Efficient similarity search in sequence databases." In: *International conference on foundations of data organization and algorithms*. Springer, 1993, pp. 69–84.

[88] K.-P. Chan and A. W.-C. Fu. "Efficient time series matching by wavelets." In: *Proceedings 15th International Conference on Data Engineering (Cat. No. 99CB36337)*. IEEE, 1999, pp. 126–133. ISBN: 0769500714.

[89] F. Korn, H. V. Jagadish, and C. Faloutsos. "Efficiently supporting ad hoc queries in large datasets of time sequences." In: *Acm Sigmod Record* 26.2 (1997), pp. 289–300. ISSN: 0163-5808.

[90] M. Kontaki, A. N. Papadopoulos, and Y. Manolopoulos. "Similarity search in time series databases." In: *Encyclopedia of Database Technologies and Applications*. IGI Global, 2005, pp. 646–651.

[91] M. M. M. Fuad and P.-F. Marteau. "The extended edit distance metric." In: *2008 International Workshop on Content-Based Multimedia Indexing*. IEEE, 2008, pp. 242–248. ISBN: 1424420431.

[92] M. K. Das and H.-K. Dai. "A survey of DNA motif finding algorithms." In: *BMC Bioinformatics* 8.7 (2007), S21. ISSN: 1471-2105. DOI: 10.1186/1471-2105-8-S7-S21. URL: https://doi.org/10.1186/1471-2105-8-S7-S21.

[93] P. Patel, E. Keogh, J. Lin, and S. Lonardi. "Mining motifs in massive time series databases." In: *2002 IEEE International Conference on Data Mining, 2002. Proceedings.* IEEE, 2002, pp. 370–377. ISBN: 0769517544.

[94] P. P. Lin Jessica, Keogh Eammon, Leonardi Stefano. "Finding motifs in time series." In: *Proc. of the 2nd Workshop on Temporal Data Mining*. 2002, pp. 53–68.

[95] B. Chiu, E. Keogh, and S. Lonardi. "Probabilistic discovery of time series motifs." In: *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*. 2003, pp. 493–498.

[96] A. Mueen, E. Keogh, and N. Bigdely-Shamlo. "Finding time series motifs in disk-resident data." In: *2009 Ninth IEEE International Conference on Data Mining*. IEEE, 2009, pp. 367–376. ISBN: 1424452422.

[97] P. Esling and C. Agon. "Time-series data mining." In: *ACM Computing Surveys* 45.1 (2012). ISSN: 03600300. DOI: 10.1145/2379776.2379788.

[98] D. Rafiei and A. Mendelzon. "Efficient retrieval of similar time sequences using DFT." In: *arXiv preprint cs/9809033* (1998).

[99] P. W. Johnson, M. Hagberg, E. Wigaeus Hjelm, and D. Rempel. "Measuring and characterizing force exposures during computer mouse use." In: *Scandinavian Journal of Work, Environment and Health* 26.5 (2000), pp. 398–405. ISSN: 03553140. DOI: 10.5271/sjweh.560. URL: https://pubmed.ncbi.nlm.nih.gov/11103838/.

[100] Y. Levanon, A. Gefen, Y. Lerman, S. Portnoy, and N. Z. Ratzon. "Key Strike Forces and Their Relation to High Level of Musculoskeletal Symptoms." In: *Safety and Health at Work* 7.4 (2016), pp. 347–353. ISSN: 20937997. DOI: 10.1016/j.shaw.2016.04.008. URL: /pmc/articles/PMC5128007/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC5128007/.

[101] E. Bernmark and C. Wiktorin. "A triaxial accelerometer for measuring arm movements." In: *Applied Ergonomics* 33.6 (2002), pp. 541–547. ISSN: 00036870. DOI: 10.1016/S0003-6870(02)00072-8.

[102] G. Hansson, P. Asterland, N. G. Holmer, and S. Skerfving. "Validity and reliability of triaxial accelerometers for inclinometry in posture analysis." In: *Medical and Biological Engineering and Computing* 39.4 (2001), pp. 405–413. ISSN: 01400118. DOI: 10.1007/BF02345361. URL: https://link.springer.com/article/10.1007/BF02345361.

[103] W. S. Marras, F. A. Fathallah, R. J. Miller, S. W. Davis, and G. A. Mirka. "Accuracy of a three-dimensional lumbar motion monitor for recording dynamic trunk motion characteristics." In: *International Journal of Industrial Ergonomics* 9.1 (1992), pp. 75–87. ISSN: 0169-8141.

[104] W. S. Marras and K. P. Granata. "The development of an EMG-assisted model to assess spine loading during whole-body free-dynamic lifting." In: *Journal of Electromyography and Kinesiology*. 1997. DOI: 10.1016/S1050-6411(97)00006-0.

[105] U. C. Gatti, G. C. Migliaccio, S. Schneider, and R. Fierro. "Assessing physical strain in construction workforce: A first step for improving safety and productivity management." In: *Proceedings, 27th international symposium on automation and robotics in construction (ISARC), international association for automation and robotics in construction (IAARC)*. 2010, pp. 255–264.

[106] X. Ning. "Development of a new work-rest scheduling model based on inventory control theory." Doctoral dissertation. 2011. ISBN: 9781124871783.

[107] A. D. Nimbarte, M. Zreiqat, and X. Ning. "Impact of shoulder position and fatigue on the flexion-relaxation response in cervical spine." In: *Clinical Biomechanics* (2014). ISSN: 18791271. DOI: 10.1016/j.clinbiomech.2013.12.003.

[108] X. Ning and G. A. Mirka. "The effect of sinusoidal rolling ground motion on lifting biomechanics." In: *Applied Ergonomics* 42.1 (2010), pp. 131–137. ISSN: 00036870. DOI: 10.1016/j.apergo.2010.06.001. URL: https://pubmed.ncbi.nlm.nih.gov/20591409/.

[109] X. Ning, J. Zhou, B. Dai, and M. Jaridi. "The assessment of material handling strategies in dealing with sudden loading: The effects of load handling position on trunk biomechanics." In: *Applied Ergonomics* 45.6 (2014), pp. 1399–1405. ISSN: 18729126. DOI: 10.1016/j.apergo.2014.03.008. URL: https://pubmed.ncbi.nlm.nih.gov/24766903/.

[110] W. S. Marras and K. P. Granata. "The development of an EMG-assisted model to assess spine loading during whole-body free-dynamic lifting." In: *Journal of Electromyography and Kinesiology*. Vol. 7. 4. J Electromyogr Kinesiol, 1997, pp. 259–268. DOI: 10.1016/S1050-6411(97)00006-0. URL: https://pubmed.ncbi.nlm.nih.gov/11369269/.

[111] K. Schüldt, J. Ekholm, K. Harms-Ringdahl, U. P. Arborelius, and G. Németh. "Influence of sitting postures on neck and shoulder e.m.g. during arm-hand work movements." In: *Clinical Biomechanics* (1987). ISSN: 02680033. DOI: 10.1016/0268-0033(87)90003-9.

[112] J. G. Richards. "The measurement of human motion: A comparison of commercially available systems." In: *Human movement science* 18.5 (1999), pp. 589–602. ISSN: 0167-9457.

[113] M. Topley and J. Richards. "A Comparison of Currently Available Optoelectronic Motion Capture Systems." In: *Journal of Biomechanics* 106 (2020), p. 109820. DOI: 10.1016/j.jbiomech.2020.109820.

[114] M Friedrich. "Measuring lumbar sagittal posture in sewage workers using an ultrasonic device." In: *Journal of Musculoskeletal research* 6.03n04 (2002), pp. 135–145. ISSN: 0218-9577.

[115] H. HSIAO and W MONROE KEYSERLING. "A three-dimensional ultrasonic system for posture measurement." In: *Ergonomics* 33.9 (1990), pp. 1089–1114. ISSN: 0014-0139.

[116] J. M. Sabatier and A. E. Ekimov. *Ultrasonic methods for human motion detection*. Tech. rep. 2006.

[117] J. M. de Toledo, J. F. Loss, T. W. Janssen, J. W. van der Scheer, T. D. Alta, W. J. Willems, and D. H. E. J. Veeger. "Kinematic evaluation of patients with total and reverse shoulder arthroplasty during rehabilitation exercises with different loads." In: *Clinical Biomechanics* 27.8 (2012), pp. 793–800. ISSN: 0268-0033.

[118] G. R. Johnson and J. M. Anderson. "Measurement of three-dimensional shoulder movement by an electromagnetic sensor." In: *Clinical Biomechanics* 5.3 (1990), pp. 131–136. ISSN: 0268-0033.

[119] G. Li and P. Buckle. "Current techniques for assessing physical exposure to work-related musculoskeletal risks, with emphasis on posture-based methods." In: *Ergonomics* 42.5 (1999), pp. 674–695. ISSN: 0014-0139.

[120] N. Barbour and G. Schmidt. "Inertial sensor technology trends." In: *IEEE Sensors Journal* 1.4 (2001), pp. 332–339. ISSN: 1530437X. DOI: 10.1109/7361.983473.

[121] S. Lim and C. D'Souza. *A narrative review on contemporary and emerging uses of inertial sensing in occupational ergonomics*. 2020. DOI: 10.1016/j.ergon.2020.102937.

[122] G. S. Faber, I Kingma, C. C. Chang, J. T. Dennerlein, and J. H. van Dieën. "Validation of a wearable system for 3D ambulatory L5/S1 moment assessment during manual lifting using instrumented shoes and an inertial sensor suit." In: *Journal of biomechanics* 102 (2020), p. 109671. ISSN: 0021-9290.

[123] M. Brandt, P. Madeleine, A. Samani, M. D. Jakobsen, S. Skals, J. Vinstrup, and L. L. Andersen. "Accuracy of identification of low or high risk lifting during standardised lifting situations." In: *Ergonomics* 61.5 (2018), pp. 710–719.

[124] H. J. Busser, W. G. De Korte, E. B. C. Glerum, and R. C. Van Lummel. "Method for objective assessment of physical work load at the workplace." In: *Ergonomics* 41.10 (1998), pp. 1519–1526.

[125] S. Kim and M. A. Nussbaum. "An evaluation of classification algorithms for manual material handling tasks based on data obtained using wearable technologies." In: *Ergonomics* 57.7 (2014), pp. 1040–1051.

[126] R. C. Ailneni, K. R. Syamala, I.-S. Kim, and J. Hwang. "Influence of the wearable posture correction sensor on head and neck posture: Sitting and standing workstations." In: *Work* 62.1 (2019), pp. 27–35.

[127] D. Battini, A. Persona, and F. Sgarbossa. "Innovative real-time system to integrate ergonomic evaluations into warehouse design and management." In: *Computers & Industrial Engineering* 77 (2014), pp. 1–10.

[128] R. Bootsman, P. Markopoulos, Q. Qi, Q. Wang, and A. A. A. Timmermans. "Wearable technology for posture monitoring at the workplace." In: *International Journal of Human-Computer Studies* 132 (2019), pp. 99–111.

[129] L Peppoloni, A Filippeschi, E Ruffaldi, and C. A. Avizzano. "A novel wearable system for the online assessment of risk for biomechanical load in repetitive efforts." In: *International Journal of Industrial Ergonomics* 52 (2016), pp. 1–11.

[130] E. Valero, A. Sivanathan, F. Bosché, and M. Abdel-Wahab. "Musculoskeletal disorders in construction: A review and a novel system for activity tracking with body area network." In: *Applied ergonomics* 54 (2016), pp. 120–130.

[131] J Village, C Trask, N Luong, Y Chow, P Johnson, M Koehoorn, and K Teschke. "Development and evaluation of an observational Back-Exposure Sampling Tool (Back-EST) for work-related back injury risk factors." In: *Applied ergonomics* 40.3 (2009), pp. 538–544.

[132] S. Lambrecht, S. L. Nogueira, M. Bortole, A. A. G. Siqueira, M. H. Terra, E. Rocon, and J. L. Pons. "Inertial Sensor Error Reduction through Calibration and Sensor Fusion." eng. In: *Sensors (Basel, Switzerland)* 16.2 (2016), p. 235. ISSN: 1424-8220 (Electronic). DOI: 10.3390/s16020235.

[133] C. M. Lu and N. J. Ferrier. "Repetitive Motion Analysis: Segmentation and Event Classification." In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26.2 (2004), pp. 258–263. ISSN: 01628828. DOI: 10.1109/TPAMI.2004.1262196.

[134] C. M. Lu and N. J. Ferrier. "Automated Analysis of Repetitive Joint Motion." In: *IEEE Transactions on Information Technology in Biomedicine* 7.4 (2003), pp. 263–273. ISSN: 10897771. DOI: 10.1109/TITB.2003.821309.

[135] D. Wang, F. Dai, and X. Ning. "Risk assessment of work-related musculoskeletal disorders in construction: State-of-the-art review." In: *Journal of Construction Engineering and Management* 141.6 (2015). ISSN: 07339364. DOI: 10.1061/(ASCE)CO.1943-7862.0000979.

[136] Z. Zhang. "Microsoft kinect sensor and its effect." In: *IEEE multimedia* 19.2 (2012), pp. 4–10. ISSN: 1070-986X.

[137] C. Lu, H. Liu, and N. J. Ferrier. "Multidimensional motion segmentation and identification." In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Vol. 2. IEEE, 2000, pp. 629–636. DOI: 10.1109/cvpr.2000.854931.

[138] A. Vögele, B. Krüger, and R. Klein. "Efficient unsupervised temporal segmentation of human motion." In: *SCA 2014 - Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation*. 2014.

[139] K. C. Liu and C. T. Chan. "Significant change spotting for periodic human motion segmentation of cleaning tasks using wearable sensors." In: *Sensors (Switzerland)* 17.1 (2017). ISSN: 14248220. DOI: 10.3390/s17010187. URL: /pmc/articles/PMC5298760/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC5298760/.

[140] J. P. Eckmann, S. O. Kamphorst, and D Ruelle. "Recurrence plots of dynamical systems." In: *World Scientific Series on Nonlinear Science Series A* 16 (1995), pp. 441–446.

[141] N Marwan. "A historical review of recurrence plots." In: *The European Physical Journal Special Topics* 164.1 (2008), pp. 3–12. ISSN: 1951-6401. DOI: 10.1140/epjst/e2008-00829-1. URL: https://doi.org/10.1140/epjst/e2008-00829-1.

[142] J.-J. Aucouturier and M. Sandler. "Finding Repeating Patterns in Acoustic Musical Signals : Applications for Audio Thumbnailing." In: *journal of the audio engineering society* (2002).

[143] J. Foote. "Automatic audio segmentation using a measure of audio novelty." In: *IEEE International Conference on Multi-Media and Expo*. I/MONDAY. 2000, pp. 452–455. DOI: 10.1109/icme.2000.869637.

[144] J. Foote. "Visualizing music and audio using self-similarity." In: *Proceedings of the ACM International Multimedia Conference & Exhibition*. ACM, 1999, pp. 77–80. DOI: 10.1145/319463.319472.

[145] M. Müller. *Fundamentals of Music Processing*. Springer International Publishing, 2015. DOI: 10.1007/978-3-319-21945-5.

[146] M. Cooper and J. Foote. "Summarizing popular music via structural similarity analysis." In: *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. Vol. 2003-Janua. Institute of Electrical and Electronics Engineers Inc., 2003, pp. 127–130. ISBN: 0780378504. DOI: 10.1109/ASPAA.2003.1285836.

[147] M. J. Bruderer, M. F. McKinney, and A. Kohlrausch. "Structural boundary perception in popular music." In: *ISMIR*. 2006, pp. 198–201.

[148] R. B. Dannenberg and N. Hu. "Pattern discovery techniques for music audio." In: *International Journal of Phytoremediation* 21.1 (2003), pp. 153–163. ISSN: 15497879. DOI: 10.1076/jnmr.32.2.153.16738.

[149] G. Peeters. "Sequence Representation of Music Structure Using Higher-Order Similarity Matrix and Maximum-Likelihood Approach." In: *ISMIR*. 2007, pp. 35–40.

[150] G. Peeters, A. La Burthe, and X. Rodet. "Toward automatic music audio summary generation from signal analysis." In: 2002.

[151] G. Peeters. "Deriving musical structures from signal analysis for music audio summary generation: "Sequence"and "state"approach." In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 2771 (2004), pp. 143–166. ISSN: 16113349. DOI: 10.1007/978-3-540-39900-1_14. URL: http://www.ircam.fr.

[152] B. Meudic and E. St-James. "Automatic extraction of approximate repetitions in polyphonic midi files based on perceptive criteria." In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 2771 (2004), pp. 124–142. ISSN: 16113349. DOI: 10.1007/978-3-540-39900-1_13. URL: https://link.springer.com/chapter/10.1007/978-3-540-39900-1{_}13.

[153] F. Kaiser and T. Sikora. "Music Structure Discovery in Popular Music using Non-negative Matrix Factorization." In: *ISMIR*. 2010, pp. 429–434.

[154] S. Jun and E. Hwang. "Music segmentation and summarization based on self-similarity matrix." In: *Proceedings of the 7th International Conference on Ubiquitous Information Management and Communication, ICUIMC 2013*. New York, New York, USA: ACM Press, 2013, pp. 1–4. ISBN: 9781450319584. DOI: 10.1145/2448556.2448638. URL: http://dl.acm.org/citation.cfm?doid=2448556.2448638.

[155] S. Jun, S. Rho, and E. Hwang. "Music structure analysis using self-similarity matrix and two-stage categorization." In: *Multimedia Tools and Applications* 74.1 (2013), pp. 287–302. ISSN: 15737721. DOI: 10.1007/s11042-013-1761-9. URL: https://link.springer.com/article/10.1007/s11042-013-1761-9.

[156] M. A. Bartsch and G. H. Wakefield. "To catch a chorus: Using chroma-based representations for audio thumbnailing." In: *Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No. 01TH8575)*. IEEE, 2001, pp. 15–18. ISBN: 0780371267.

[157] M. Cooper and J. Foote. "Automatic Music Summarization via Similarity Analysis." In: *ISMIR*. 2002.

[158] A. L. Jacobson. "Auto-threshold peak detection in physiological signals." In: *2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Vol. 3. 2001, 2194–2195 vol.3. DOI: 10.1109/IEMBS.2001.1017206.

[159] *ARTHROKINEMAT | Hui Liu | Research Project*. URL: https://www.researchgate.net/project/Arthrokinemat (visited on 04/25/2021).

[160] H. Liu and T. Schultz. *A Wearable Real-time Human Activity Recognition System using Biosensors Integrated into a Knee Bandage*. 2019, pp. 47–55. DOI: 10.5220/0007398800470055.

[161] D. Dua and C. Graff. *{UCI} Machine Learning Repository*. 2017. URL: http://archive.ics.uci.edu/ml.

[162] 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.

[163] *UCI Machine Learning Repository: Smartphone-Based Recognition of Human Activities and Postural Transitions Data Set*. URL: https://archive.ics.uci.edu/ml/datasets/Smartphone-Based+Recognition+of+Human+Activities+and+Postural+Transitions (visited on 04/23/2021).

[164] S. M. L. V. P. dos Santos. "Explaining the Ergonomic Assessment of Human Movement in Industrial Contexts." Doctoral dissertation. Nova School f Science and Technology, 2019. URL: https://github.com/joaomlourenco/novathesis.

[165] Fraunhofer AICOS. *IOTIP | INTERNET OF THINGS IN A PACKAGE: WAFER LEVEL MODULAR ARCHITECTURE FOR INTERNET OF THINGS*. Tech. rep. URL: https://www.fraunhofer.pt/en/fraunhofer{_}portugal/news/news{_}archive/a-{_}made-in-portugal-technology-for-iot-growth.html.

[166] J. Zhao. "A Review of Wearable IMU (Inertial-Measurement-Unit)-based Pose Estimation and Drift Reduction Technologies." In: *Journal of Physics: Conference Series* 1087 (2018), p. 42003. ISSN: 1742-6588. DOI: 10.1088/1742-6596/1087/4/042003. URL: http://dx.doi.org/10.1088/1742-6596/1087/4/042003.

[167] W. Sousa Lima, E. Souto, K. El-Khatib, R. Jalali, and J. Gama. "Human activity recognition using inertial sensors in a smartphone: An overview." In: *Sensors* 19.14 (2019), p. 3213.

[168] S. S. Bangaru, C. Wang, and F. Aghazadeh. "Data Quality and Reliability Assessment of Wearable EMG and IMU Sensor for Construction Activity Recognition." In: *Sensors* 20.18 (2020). ISSN: 1424-8220. DOI: 10.3390/s20185264. URL: https://www.mdpi.com/1424-8220/20/18/5264.

[169] R. Khusainov, D. Azzi, I. E. Achumba, and S. D. Bersch. "Real-Time Human Ambulation, Activity, and Physiological Monitoring: Taxonomy of Issues, Techniques, Applications, Challenges and Limitations." In: *Sensors* 13.10 (2013), pp. 12852–12902. ISSN: 1424-8220. DOI: 10.3390/s131012852. URL: https://www.mdpi.com/1424-8220/13/10/12852.

[170] V Behravan, N. E. Glover, R Farry, P. Y. Chiang, and M Shoaib. "Rate-adaptive compressed-sensing and sparsity variance of biomedical signals." In: *2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. 2015, pp. 1–6. ISBN: 2376-8894 VO -. DOI: 10.1109/BSN.2015.7299419.

[171] A. L. Goldberger, L. A. Amaral, L Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley. "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals." eng. In: *Circulation* 101.23 (2000), E215–20. ISSN: 1524-4539 (Electronic). DOI: 10.1161/01.cir.101.23.e215.

[172] M. Barandas, D. Folgado, L. Fernandes, S. Santos, M. Abreu, P. Bota, H. Liu, T. Schultz, and H. Gamboa. "Tsfel: Time series feature extraction library." In: *SoftwareX* 11 (2020), p. 100456. ISSN: 2352-7110.

[173] 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.

[174] *C4S4_NoveltySegmentation*. URL: https://www.audiolabs-erlangen.de/resources/MIR/FMP/C4/C4S4{_}NoveltySegmentation.html (visited on 04/22/2021).

[175] M. Feurer and F. Hutter. "Hyperparameter optimization." In: *Automated machine learning*. Springer, Cham, 2019, pp. 3–33.

[176] A. Gensler and B. Sick. "Novel Criteria to Measure Performance of Time Series Segmentation Techniques." In: *LWA*. Citeseer, 2014, pp. 193–204.

[177] A. Santos, J. Rodrigues, D. Folgado, S. Santos, C. Fujao, and H. Gamboa. "Self-Similarity Matrix of Morphological Features for Motion Data Analysis in Manufacturing Scenarios." In: *BIOSIGNALS*. 2021, pp. 80–90.

[178] S. M. Law. "STUMPY: A powerful and scalable Python library for time series data mining." In: *Journal of Open Source Software* 4.39 (2019), p. 1504. ISSN: 2475-9066.

[179] N. R. Council. "Musculoskeletal disorders and the workplace: low back and upper extremities." In: (2001). ISSN: 0309072840.

[180] F Pedregosa, G Varoquaux, A Gramfort, V Michel, B Thirion, O Grisel, M Blondel, P Prettenhofer, R Weiss, V Dubourg, J Vanderplas, A Passos, D Cournapeau, M Brucher, M Perrot, and E Duchesnay. "Scikit-learn: Machine Learning in {P}ython." In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[181] M. Müller and F. Zalkow. "libfmp: A Python Package for Fundamentals of Music Processing." In: *Journal of Open Source Software* 6.63 (2021), p. 3326. DOI: 10.21105/joss.03326. URL: https://doi.org/10.21105/joss.03326.

Appendix A

Industrial Database Information

Table A.1: Individual characteristics of the testing subjects [164].

Subject's characteristics	
Number	12
Gender (m:f)	9:3
Age ($\overline{yr} \pm \sigma$)	36±9
Height ($\overline{h} \pm \sigma$)	172±7
Gender (r)	Right hand (11:12)

Table A.2: Time differences of the calibration t pose position. For each time subject, the time differences of the hand, forearm, arm and torso IMU signals being respectively represented. The 0 points in red were the first positions where this position was registered, and served as reference points for the others

Time axis	δ_{hand}	$\delta_{forearm}$	δ_{arm}	δ_{torso}	**UT len**
Opr1 Wkst1	9.45	4.46	0.0	18.9	1547.89
Opr1 Wkst2	10.3	10.1	9.7	0.0	1788.61
Opr2 Wkst1&2	0.0	2.14	0.24	2.05	1.35
Opr3 Wkst1	12.7	0.0	12.7	0.1	1562.81
Opr3 Wkst2	0.01	0.0	0.01	0.01	1577.31
Opr4 Wkst1	0.0	25.86	1.36	11.36	1952.04
Opr4 Wkst2	0.01	0.0	0.01	0.01	1228.75
Opr5 Wkst1	0.0	0.3	1.3	1.4	1640.3
Opr5 Wkst2	0.01	0.0	0.02	0.03	1641.58
Opr6 Wkst1	0.01	0.0	0.02	0.03	2010.76
Opr6 Wkst2	0.0	10.34	13.52	17.37	1533.18

Table A.3: Temporal information regarding the *Industrial* Database.

Time axis	i time	time len	sample len	Average Fs	min Fs	max Fs
Opr1 Wkst1						
Hand						
Acc	1994.15	1593.8	158082	99.19	0.44	102.4
Gyr	1994.15	1593.8	158136	99.22	0.45	102.4
Forearm						
Acc	1999.23	1552.36	153304	98.76	1.34	100.73
Gyr	1999.23	1552.36	153444	98.85	1.18	100.73
Arm						
Acc	2004.04	1583.43	68152	43.04	3.25	101.35
Gyr	2004.21	1583.31	81170	51.27	5.3	100.68
Torso						
Acc	0.0	1602.0	318148	198.59	197.21	200.1
Gyr	0.35	1601.65	318079	198.59	197.21	200.1
Opr1 Wkst2						
Hand						
Acc	1981.23	1800.02	181464	100.81	0.08	101.58
Gyr	1981.23	1800.02	181464	100.81	0.08	101.58
Forearm						
Acc	1981.42	1799.27	179164	99.58	0.08	100.83
Gyr	1981.42	1799.27	179154	99.57	0.08	100.83
Arm						
Acc	1981.63	1799.09	78600	43.69	0.08	101.68
Gyr	1981.63	1799.09	92356	51.33	0.08	101.68
Torso						
Acc	0.0	1789.12	355310	198.59	197.33	199.78
Gyr	0.36	1788.77	355239	198.59	197.33	199.78
Opr2 Wkst1&2						
Hand						
Acc	2706.85	2656.5	127834	48.12	1.85	102.53
Gyr	2706.85	2656.54	149080	56.12	1.99	102.53
Forearm						
Acc	2694.12	2669.44	268130	100.44	2.58	101.65
Gyr	2694.12	2669.44	268118	100.44	1.97	101.65
Arm						
Acc	2696.25	2667.24	262034	98.24	0.08	100.68
Gyr	2696.25	2667.22	262464	98.4	0.08	100.68

Table A.3: Temporal information regarding the *Industrial* Database.

Time axis	i time	time len	sample len	Average Fs	min Fs	max Fs
Opr3 Wkst1						
Hand						
Acc	59.7	1575.5	158744	100.76	0.1	101.58
Gyr	59.7	1575.5	158756	100.77	0.11	101.58
Forearm						
Acc	72.17	1562.81	156930	100.42	4.03	100.83
Gyr	72.17	1562.81	156946	100.43	5.3	100.83
Arm						
Acc	59.7	1575.5	158744	100.76	0.1	101.58
Gyr	59.7	1575.5	158756	100.77	0.11	101.58
Torso						
Acc	0.22	1569.41	77920	49.65	49.56	49.72
Gyr	0.08	1569.57	311709	198.59	197.24	199.72
Opr3 Wkst2						
Hand						
Acc	67.74	1577.32	160128	101.52	100.52	101.57
Gyr	67.74	1577.32	160128	101.52	100.52	101.57
Forearm						
Acc	67.56	1577.44	158728	100.62	99.6	100.83
Gyr	67.56	1577.44	158728	100.62	99.6	100.83
Arm						
Acc	67.39	1577.49	156148	98.99	4.38	100.68
Gyr	67.39	1577.47	155534	98.6	6.71	100.68
Torso						
Acc	0.0	1577.99	628685	398.41	44.31	415.74
Gyr	0.14	1577.86	628641	398.41	199.44	415.74
Opr4 Wkst1						
Hand						
Acc	5204.41	1953.42	198292	101.51	100.46	102.53
Gyr	5204.41	1953.42	198292	101.51	100.46	102.53
Forearm						
Acc	5183.07	1974.67	194288	98.39	0.07	101.65
Gyr	5183.07	1974.67	194556	98.53	0.07	101.65
Arm						
Acc	5203.93	1953.61	191314	97.93	3.99	101.35
Gyr	5204.14	1953.41	190304	97.42	6.71	101.35

Table A.3: Temporal information regarding the *Industrial* Database.

Time axis	i time	time len	sample len	Average Fs	min Fs	max Fs
Torso						
Acc	0.0	1968.75	390982	198.59	197.17	199.96
Gyr	0.36	1968.38	390910	198.59	197.17	199.96
Opr4 Wkst2						
Hand						
Acc	1637.45	1228.98	124748	101.51	100.49	102.41
Gyr	1637.45	1228.98	124748	101.51	100.49	102.41
Forearm						
Acc	1637.6	1228.75	123628	100.61	99.6	101.65
Gyr	1637.6	1228.75	123628	100.61	99.6	101.65
Arm						
Acc	1637.25	1229.13	123696	100.64	99.6	101.68
Gyr	1637.25	1229.13	123696	100.64	99.6	101.68
Torso						
Acc	0.0	1229.41	122970	100.02	83.35	689.05
Gyr	0.03	1229.39	614763	500.05	361.25	507.69
Opr5 Wkst1						
Hand						
Acc	643.75	1641.87	164874	100.42	5.34	101.58
Gyr	643.75	1641.87	164654	100.28	6.77	101.58
Forearm						
Acc	643.94	1641.83	165200	100.62	99.6	101.65
Gyr	643.94	1641.83	165200	100.62	99.6	101.65
Arm						
Acc	644.1	1641.61	165224	100.65	99.3	101.68
Gyr	644.1	1641.61	165224	100.65	99.3	101.68
Torso						
Acc	0.0	1642.43	654257	398.35	49.81	415.43
Gyr	0.15	1642.28	654209	398.35	199.42	415.43
Opr5 Wkst2						
Hand						
Acc	643.75	1641.87	164874	100.42	5.34	101.58
Gyr	643.75	1641.87	164654	100.28	6.77	101.58
Forearm						
Acc	643.94	1641.83	165200	100.62	99.6	101.65
Gyr	643.94	1641.83	165200	100.62	99.6	101.65

Table A.3: Temporal information regarding the *Industrial* Database.

Time axis	i time	time len	sample len	Average Fs	min Fs	max Fs
Arm						
Acc	644.1	1641.61	165224	100.65	99.3	101.68
Gyr	644.1	1641.61	165224	100.65	99.3	101.68
Torso						
Acc	0.0	1642.43	654257	398.35	49.81	415.43
Gyr	0.15	1642.28	654209	398.35	199.42	415.43
Opr6 Wkst1						
Hand						
Acc	38.98	2010.78	204028	101.47	6.77	102.53
Gyr	38.98	2010.78	204038	101.47	14.5	102.53
Forearm						
Acc	38.81	2010.82	202008	100.46	2.45	101.68
Gyr	38.81	2010.82	201948	100.43	3.73	101.68
Arm						
Acc	38.68	2011.05	202400	100.64	99.6	101.68
Gyr	38.68	2011.05	202400	100.64	99.6	101.68
Torso						
Acc	0.0	2011.41	801249	398.35	132.8	416.03
Gyr	0.13	2011.28	801201	398.35	199.44	416.03
Opr6 Wkst2						
Hand						
Acc	96.41	1533.19	155636	101.51	100.49	101.58
Gyr	96.41	1533.19	155636	101.51	100.49	101.58
Forearm						
Acc	96.23	1533.33	154224	100.58	20.12	101.65
Gyr	96.23	1533.33	154196	100.56	20.13	101.65
Arm						
Acc	96.06	1533.43	154328	100.64	99.6	101.35
Gyr	96.06	1533.43	154328	100.64	99.6	101.35
Torso						
Acc	0.0	1533.88	611046	398.37	99.87	412.37
Gyr	0.11	1533.77	611006	398.37	199.43	412.37

Appendix B

UCI Database Information

Table B.1: Range of values applyed for the hyperparameter otimization of the *UCI* Database

Parameters	Range of values for Grid Search
$Wind_{len}$ (s)	1, 5, 9, 13, 17
$Overlap_{frac}$	0.5, 0.6, 0.8, 0.9
$Kernel_{len}$ (s)	5, 10, 15, 20, 25
$SubWind_{len}$ (s)	0, 5, 10, 15, 20
$SmoothWind_{len}$ (s)	0, 5, 10, 15, 20
$Thresh_{frac}$	0, 0.05, 0.1, 0.15, 0.2, 0.3, 0.5

Table B.2: Selected parameter values for each time serie sample of the *UCI* Database

TS samples	$Wind_{len}$ (s)	$Overlap_{frac}$	$Kernel_{len}$ (s)	$SubWind_{len}$ (s)	$SmoothWind_{len}$ (s)	$Thresh_{frac}$
1	5,00	0,80	10,00	10,00	5,00	0,00
2	9,00	0,80	15,00	10,00	0,00	0,15
3	1,00	0,80	10,00	15,00	5,00	0,10
4	1,00	0,80	15,00	5,00	5,00	0,05
5	1,00	0,80	5,00	20,00	5,00	0,15
6	1,00	0,60	5,00	5,00	5,00	0,20
7	1,00	0,80	20,00	5,00	10,00	0,05
8	5,00	0,50	25,00	20,00	10,00	0,05
9	1,00	0,80	10,00	20,00	5,00	0,05
10	9,00	0,80	25,00	10,00	0,00	0,05
11	9,00	0,80	25,00	10,00	5,00	0,00
12	9,00	0,80	20,00	5,00	10,00	0,00
13	17,00	0,80	10,02	20,00	0,00	0,00
14	1,00	0,80	5,00	20,00	5,00	0,10
15	13,00	0,80	5,00	20,00	0,00	0,20
16	1,00	0,60	10,00	10,00	5,00	0,05
17	5,00	0,60	25,00	20,00	5,00	0,00
18	1,00	0,80	25,00	15,00	5,00	0,00
19	1,00	0,60	10,00	20,00	5,00	0,00
20	13,00	0,60	15,00	15,00	5,00	0,15
21	5,00	0,80	20,00	10,00	0,00	0,00
22	5,00	0,60	10,00	10,00	5,00	0,00

Table B.2: Selected parameter values for each time serie sample of the *UCI* Database

TS samples	$Wind_{len}$ (s)	$Overlap_{frac}$	$Kernel_{len}$ (s)	$SubWind_{len}$ (s)	$SmoothWind_{len}$ (s)	$Thresh_{frac}$
23	5,00	0,80	20,00	20,00	10,00	0,00
24	13,00	0,60	20,00	15,00	10,00	0,15
25	9,00	0,60	20,00	15,00	0,00	0,00
26	1,00	0,60	25,00	5,00	10,00	0,05
27	1,00	0,60	20,00	15,00	5,00	0,00
28	1,00	0,60	10,00	20,00	5,00	0,05
29	1,00	0,80	10,00	20,00	10,00	0,05
30	5,00	0,60	5,00	20,00	0,00	0,15
31	1,00	0,60	5,00	20,00	5,00	0,05
32	5,00	0,80	25,00	15,00	5,00	0,00
33	1,00	0,80	5,00	15,00	10,00	0,10
34	1,00	0,80	10,00	15,00	10,00	0,20
35	1,00	0,80	15,00	10,00	10,00	0,15
36	1,00	0,80	10,00	20,00	15,00	0,05
37	1,00	0,80	10,00	20,00	10,00	0,00
38	5,00	0,80	10,00	20,00	5,00	0,10
39	1,00	0,60	15,00	20,00	5,00	0,00
40	9,00	0,80	10,00	20,00	10,00	0,05
41	1,00	0,60	25,00	20,00	5,00	0,00
42	1,00	0,60	5,00	20,00	5,00	0,10
43	1,00	0,60	25,00	15,00	5,00	0,00
44	1,00	0,60	10,00	15,00	5,00	0,10

Table B.2: Selected parameter values for each time serie sample of the *UCI* Database

TS samples	$Wind_{len}$ (s)	$Overlap_{frac}$	$Kernel_{len}$ (s)	$SubWind_{len}$ (s)	$SmoothWind_{len}$ (s)	$Thresh_{frac}$
45	1,00	0,60	10,00	20,00	5,00	0,00
46	9,00	0,50	5,00	20,00	10,00	0,10
47	9,00	0,50	15,00	20,00	10,00	0,10
48	13,00	0,80	25,00	20,00	10,00	0,00
49	5,00	0,60	15,00	10,00	5,00	0,00
50	9,00	0,80	25,00	15,00	15,00	0,05
51	1,00	0,60	20,00	20,00	5,00	0,05
52	1,00	0,80	20,00	15,00	10,00	0,05
53	13,00	0,80	5,00	10,00	10,00	0,15
54	17,00	0,80	10,00	20,00	10,00	0,00
55	13,00	0,80	5,00	20,00	0,00	0,15
56	9,00	0,80	20,00	15,00	5,00	0,05
57	5,00	0,80	10,00	10,00	5,00	0,15
58	13,00	0,80	5,00	15,00	10,00	0,05
59	1,00	0,60	20,00	5,00	5,00	0,05
60	13,00	0,80	5,00	15,00	10,00	0,05
61	1,00	0,80	5,00	5,00	5,00	0,05

Table B.3: Results of A, R, P, F and MAE results, for the detection of type 1 events, over the *UCI* Database

TS Samples	A	R	P	F	MAE (s)
1	0,83	0,93	0,89	0,91	0,76
2	0,96	1,00	0,96	0,98	0,60
3	1,00	1,00	1,00	1,00	0,87
4	1,00	1,00	1,00	1,00	0,47
5	0,83	1,00	0,83	0,91	0,31
6	0,92	1,00	0,92	0,96	0,52
7	0,80	0,80	1,00	0,89	1,07
8	0,80	0,80	1,00	0,89	1,38
9	0,88	0,95	0,91	0,93	0,60
10	0,86	0,90	0,95	0,93	1,12
11	0,88	0,91	0,95	0,93	1,29
12	0,83	0,91	0,91	0,91	1,17
13	0,85	1,00	0,85	0,92	1,57
14	0,88	0,91	0,95	0,93	0,41
15	0,92	1,00	0,92	0,96	0,69
16	0,89	1,00	0,89	0,94	0,36
17	0,87	0,91	0,95	0,93	1,05
18	0,83	0,90	0,90	0,90	0,58
19	0,88	0,96	0,92	0,94	0,65
20	1,00	1,00	1,00	1,00	1,84
21	0,81	0,93	0,87	0,90	0,91
22	0,85	0,96	0,88	0,92	1,02
23	0,90	0,95	0,95	0,95	0,95
24	0,94	0,94	1,00	0,97	1,40
25	0,90	0,90	1,00	0,95	1,35
26	0,90	0,90	1,00	0,95	0,83
27	0,83	0,87	0,95	0,91	0,75
28	0,85	1,00	0,85	0,92	0,71
29	0,83	0,90	0,90	0,90	0,68
30	0,88	0,95	0,91	0,93	0,78
31	0,91	1,00	0,91	0,95	0,50
32	0,92	0,96	0,96	0,96	0,68
33	1,00	1,00	1,00	1,00	0,67
34	0,96	1,00	0,96	0,98	0,69
35	1,00	1,00	1,00	1,00	0,47
36	0,96	1,00	0,96	0,98	0,86

Table B.3: Results of *A*, *R*, *P*, *F* and MAE results, for the detection of type 1 events, over the *UCI* Database

TS Samples	A	R	P	F	MAE (s)
37	0,85	1,00	0,85	0,92	0,56
38	0,76	0,76	1,00	0,86	0,75
39	0,86	0,90	0,95	0,92	0,75
40	0,96	0,96	1,00	0,98	0,96
41	0,82	0,86	0,95	0,90	0,86
42	0,88	1,00	0,88	0,94	0,62
43	0,91	0,91	1,00	0,95	0,67
44	0,88	0,91	0,95	0,93	0,51
45	0,86	1,00	0,86	0,92	0,59
46	0,76	0,86	0,86	0,86	1,25
47	0,80	0,89	0,89	0,89	1,62
48	0,87	0,95	0,91	0,93	1,60
49	0,85	1,00	0,85	0,92	0,93
50	0,86	0,86	1,00	0,93	1,26
51	0,90	0,90	1,00	0,95	1,17
52	0,96	0,96	1,00	0,98	0,86
53	0,92	1,00	0,92	0,96	0,92
54	1,00	1,00	1,00	1,00	1,45
55	0,92	1,00	0,92	0,96	1,20
56	0,76	0,83	0,90	0,86	1,19
57	0,73	0,84	0,84	0,84	0,84
58	0,92	1,00	0,92	0,96	0,89
59	0,96	1,00	0,96	0,98	0,37
60	0,73	0,96	0,76	0,85	1,20
61	0,79	1,00	0,79	0,88	0,48

Appendix C

HAR Database Information

Table C.1: Range of values applyed for the hyperparameter otimization of the *HAR* Database.

Parameters	Range of values for Grid Search
$Wind_{len}$ (s)	2, 4, 6, 8, 10
$Overlap_{frac}$	0.5, 0.6, 0.8, 0.9
$SmoothWind_{len}$ (s)	0, 2, 4, 6, 8, 10 25