UNIVERSITÀ DEGLI STUDI DI UDINE
hic sunt futura
www.uniud.it

Corso di dottorato di ricerca in:
Ingegneria Industriale e dell'Informazione

Ciclo 32°

Driver's Stress Detection using Skin Potential Response Signals

Dottorando
Alessandro Piras

Supervisore
Roberto Rinaldo
Co-supervisore
Pamela Zontone

Anno 2020

Abstract

Stress in car drivers represents a risk, especially for professional car drivers which are more likely to be exposed to it for prolonged periods. A persisting stress state leads to mental and physical pathologies and increases the probability of causing accidents. Thus, the monitoring of drivers' mental state could allow an immediate action before the problem degenerates.

In the present work, two main methods to detect a subject's sympathetic reaction to stress are developed.

In the first method (*Method 1*) we measure Skin Potential Response (SPR) and record the Steering Wheel angle excursion. Then, we process the measured signals with adaptive filters which remove the component related to Motion Artifact, exploiting the relation between hand movements to handle the Steering Wheel and MA. Next, we process the obtained Stress signal with a Smooth Nonlinear Energy Operator (SNEO) to locate stress events. Experiments which allow to define a ground-truth for stress events recognition, show that, by appropriately processing, it is possible to efficiently detect stress events, obtaining a mean Recall of 95 %.

The second method (*Method 2*) uses a double channel SPR sensor to measure the SPR from both the hands, and the Electrocardiogram is recorded with a triple channel ECG sensor. SPR measurements are processed through an algorithm which selects the smoother signal. In this way we obtain a Stress signal without the Motion Artifact component. Several experiments carried out in laboratory and with a real car professional simulators, reveal the efficacy of the proposed system, which outperformed Principal Component Analysis and Independent Component Analysis.

Next, machine learning techniques are employed to classify features obtained from the Stress signal and from Heart Rate Variability. In particular, we considered a Support Vector Machine (SVM) and a feed-forward Neural Network (NN). The system is tested in experiments carried out on a professional driving simulator. Classifying 15 seconds long time intervals, we obtained a balanced accuracy of 76.72% for SVM and 77.15% for NN. Applying a relabeling method based on the previous time intervals, the performances raised to 78.74% for SVM and 78.26% for NN. We then, tested the ability to identify stress intervals and obtained a balanced accuracy of 89.58% for SVM and 91.92% for NN.

Abstract

Lo stress rappresenta un rischio per chi guida, specialmente per chi guida per professione dato che sono soggetti a stress per periodi prolungati. Un persistente stato di stress porta a patologie sia fisiche che psicologiche e incrementa la probabilità di causare incidenti.

Nel presente lavoro sono stati sviluppati due sistemi per rilevare la risposta simpatetica di soggetti sotto stress.

Nel primo metodo (Metodo 1), la Skin Potential Response (SPR) viene misurata e l'escursione della posizione del volante (Steering Wheel, SW) viene registrata. Le misure prese vengono elaborate da filtri adattativi che rimuovono la componente dell'SPR legata al Motion Artifact (MA), sfruttando la relazione che c'è tra i movimenti della mano eseguiti per far girare il volante e il Motion Artifact. Il segnale Stress (S) ottenuto, viene poi elaborato da un operatore Smooth Nonlinear Energy Operator (SNEO) per localizzare gli eventi di stress. Gli esperimenti grazie ai quali viene definita una ground-truth per la rilevazione di stress, mostrano che, attraverso le opportune elaborazioni, è possibile rilevare eventi che causano stress, ottenendo una Recall del 95 %.

Il secondo metodo (Metodo 2) usa un sensore per la misura di SPR a due canali, che viene utilizzato per misurare il segnale SPR da entrambe le mani. In aggiunta viene misurato l' Elettrocardiogramma (ECG) con un sensore a tre canali. I segnali SPR misurati vengono poi processati attraverso un algoritmo che seleziona il segnale più smussato. In questo modo, dopo la cancellazione si ottiene un Segnale stress privo dal MA. Diversi esperimenti eseguiti sia in laboratorio che in simulazione realistiche, illustrano l'efficacia del sistema proposto, che ha fornito prestazioni migliori degli algoritmi Principal Component Analysis e Independent Component Analysis.

Successivamente, vengono utilizzate tecniche di Machine Learning per classificare intervalli temporali estraendo delle feature dal segnale Stress e dal segnale Hear Rate Variability. Nello specifico, vengono utilizzati i classificatori Support Vector Machine (SVM) and a feed-forward Neural Network (NN). Il sistema sviluppato viene poi testato mediante esperimenti eseguiti su un simulatore di guida professionale. Classificando intervalli di tempo lunghi 15 s, si è ottenuta una balanced accuracy del 89.58% con l' SVM e del 91.92% con la NN.

Thesis Structure

The objective of this thesis is to devise methods and procedures to automatically detect a driver's stress condition from sensed physiological signals. To describe our work, we structured the thesis as follows.

In Chapter 1 we discuss the motivations behind our research activity. An overview about the state of the art regarding techniques for emotion recognition, with special focus on car driving contexts, follows. The contribution of the present work is then outlined.

In Chapter 2 we briefly introduce a description of the skin structure and of the sweating process. Next, the characteristics of Electrodermal signals, such as Skin Conductance and Skin Potential, are explained with the help of several electrical models specifically designed for the purpose.

Chapter 3, begins with a description of the cardiovascular system and, successively, we explain how the Electrocardiogram (ECG) is related to the heart activity. The subsequent sections focus on techniques which are widely employed to process the ECG signal and obtain the Heart Rate (HR) parameters useful to detect mental stress.

Chapter 4 introduces the Motion Artifact (MA) issue and, also by means of preliminary experiments, we show how the MA disturbances affect the Electrodermal signals we employ in our work. Next, we present the methods we devised to remove Motion Artifact disturbances in order to extrapolate the SPR component related to the subjects' emotional state. To address the task, we followed two different procedures. The first method removes Motion Artifact by using adaptive filters, which exploit the relation between the hand movements and the Steering Wheel rotation. The second method involves the processing of two SPR signals measured on both the hands. The Motion Artifact removal is carried out by implementing a smoothing algorithm that outputs the stress component which is common to both the SPR signals. A stress signal (S) is obtained and the effectiveness of the methods is evaluated through the results of experiments carried out in different driving contexts.

Chapter 5 is about the activity that follows the Motion Artifact removal addressed in Chapter 4. It begins with the theory of the classifiers which have been employed to detect stress in the subjects under test, in particular Support Vector Machines (SVM) and feed-forward Neural Networks (NN). New experiments, in which the Skin Potential Response and Electrocardiogram measurements are taken, are presented. Next, we describe the extrapolation

of features from the SPR and HR signals and how they can be useful to evaluate the subjects' state. Implementing the SVM and NN classifiers, we process the collected features, showing the respective results and performance.

Finally, in Chapter 6 the conclusions and the plans for future work are presented. We suggest to employ more advanced feature extraction methods, alternative more complex classifiers from the Machine Learning and Deep Learning literature, and to enlarge the set of data and experiments, also in autonomous driving scenarios.

Publications

•Driver's stress detection using Skin Potential Response signals A Affanni, R Bernardini, A Piras, R Rinaldo, P Zontone Measurement 122, 264-274.

• Zontone, Pamela, et al. "Low-Complexity Classification Algorithm to Identify Drivers' Stress Using Electrodermal Activity (EDA) Measurements." The World Thematic Conference-Biomedical Engineering and Computational Intelligence. Springer, Cham, 2018.

• Affanni, Antonio, et al. "Dual channel Electrodermal activity sensor for motion artifact removal in car drivers' stress detection." 2019 IEEE Sensors Applications Symposium (SAS). IEEE, 2019.

• P. Zontone, A. Affanni, R. Bernardini, A. Piras, R. Rinaldo, F. Formaggia, D. Minen, M. Minen, and C. Savorgnan, "Car Driver's Sympathetic Reaction Detection through Electrodermal Activity (EDA) and Electrocardiogram (ECG) Measurements", IEEE Transactions on Biomedical Engineering (TBME), submitted for publication, 2019.

• P. Zontone, A. Affanni, R. Bernardini, A. Piras and R. Rinaldo, "Stress Detection Through Electrodermal Activity (EDA) and Electrocardiogram (ECG) Analisys in Car Drivers", Proceedings of the European Signal Processing Conference (EUSIPCO), La Coruña, Spain, 2019.

Contents

List of Figures

List of Tables

Chapter 1

State of the art and contribution

1.1 Motivation

The problem of assessing the psycho-physical well-being of drivers has recently attracted the interest of the research and industry communities. Pressure due to continuous driving in traffic during commuting or freight delivery, for professional drivers, can indeed cause mental as well as physical exhaustion, fatigue, shoulder and back pain and hypertension [9, 11], thus increasing the number of potential traffic violations and car accidents, with great costs for the community [119]. In a research among taxi drivers [52], positive correlation has been found between negative affecting state caused by conflicts in the workplace and accidents rate. During experiments carried out in [45] the performance of drivers decreases along with the raising of anxiety, leading to a higher probability to cause accidents.

The psychological state of a subject is reflected by sympathetic reactions, which identify mental workload and pressure or, as they are commonly referred to in literature, stress conditions. Recently, many systems which are able to detect the mental stress of drivers on the basis of their physiological response, have been designed. Detecting driver's stress conditions can allow to manage the phenomenon under the physiological and contextual aspect and eventually intercede for preventing the arise of health problems and/or accidents.

1.2 Related Work

Signal processing for emotion recognition involves two main branches [47]: physical/behavioral signals [36] and physiological signals.

Physical traits based method
Behavioral traits are exposed by facial expression which are detected through imaging techniques. The large employment of facial expression recognition induced researches to develop standard protocols [86, 33]. Furthermore, facial expression techniques advanced enough to allow development of methods able to recognize micro-expressions [111].
In addition to facial expression, imaging techniques can be employed to obtain a measurement of tissue oxygen saturation (StO2) as a feature for detecting human stress [27].
Behavior is well depicted by human gesture, which is analyzed in [83] using video techniques. Eye tracking systems are capable to detect eye movements as well as pupil dilatation and blink rate [75], [48] and through audio signal based techniques it is possible to detect tones and inflections that reveal inner stress [87], [89].

Physiological traits based methods
Among physiological signals, heart rate has been considered one of the most reliable for stress detection [47, 36, 86, 55, 63, 88]. Photoplethysmography optically measures blood volume changes and has many applications for stress and emotional detection [105, 74, 120]. In [96] it is shown how the Electrodermal Activity (EDA) characteristics differ depending on the emotional state, while in [102] it can be seen how the Electrodermal Activity can be used for detecting and discriminate mental stress and cognitive load.

To detect arousal, stress and emotional state of subjects, machine

learning techniques are widely employed for feature classification. Several methods are employed to train and classify physiological signals features. Such implementations can be found in [56] for linear discriminant analysis (LDA), Naïve Bayes, classification and regression tree (CART), self-organization map (SOMs), and support vector machine (SVM) or in [104] for Naive Bayes, Support Vector Machine (SVM) with linear and RBF kernel. In [76], a comparison of different classifiers concluded that KNN and SVM outperform Discriminant Analysis and Classification Trees. SVM is broadly used and its efficacy is confirmed in [70, 8].

The use of Neural Networks (NN) is taking root; not only because they are the most advanced type of classifiers, but also because their structure allows to create a lot of variants based on the users' specific needs and applications. We can find NN applied to emotional classification in [88, 25]; furthermore several NN variants are being applied, such as Back Propagation NN [88], Hierarchical NN [114], Convolutional NN [100], Unsupervised Deep Belief Network (UDBN) [51].

Several works have been focused on feature selection and extraction methods. Many features exist for detecting mental state, but not all of them represent a good indicator for a given application. Hence, choosing the optimal features can make the system computationally lighter and increase the performance at the same time. In [117], an analysis of Heart Rate features has been carried out to evaluate the ones whose response is mostly correlated to stress. An analogous work for EDA can be found in [103] where features belonging to EDA measurements are collected and automatically selected through Joint and Conditional Mutual Information and Double Input Symmetrical Relevance filters. Authors in [46] developed a system to evaluate features efficacy considering their impact on arousal and valence state. In [61] features are selected to recognize mental state considering their sensitivities to local changes.

EEG methods acquire data from a large number of channels [13],

and feature selection is needed, to discard the redundant ones [66, 67].

Another issue that could undermine the reliability of features is given by inter-subject and contextual differences [104]. In [113], a clustering method for categorizing the users is developed to avoid the inaccuracies arising when trying to make the subject data homogeneous or fitting the system on the current subject every time.
Another approach is to use unsupervised Machine Learning techniques to automatically find significant features as authors have done in [51], merging the obtained features with the standard ones using data fusion methods.

Applications in the driving environment
Several stress detection works, involving the techniques discussed above, are set in the car driving scenarios. A crucial aspect to take into account, is the right signal choice depending on the purposes of the research activity. Authors in [68] compare a large number of physiological signals used to evaluate drowsiness and cognitive state during car driving, discussing related advantages and limitations about the employment of such data depending on the specific application. The surveys in [81, 69] compare machine learning approaches and put in evidence the issues which still have to be challenged in a deeper way, such as the assumptions to do in order to evaluate which situations cause a determined kind of reaction, such as stress or appraisal, the experimental setup that differs among the experiments, inter-subject differences, invasiveness of sensors and reliability of tests to resemble real driving scenario.

In [106], tests on a professional car driver during a car race, are carried out to analyze features belonging to different bio-signals. The results showed that HR and sweat rate are good indicators for mental stress detection while electromyogram (EMG) is suitable for recognizing physical stress but not mental stress. Differently from [106], authors in [60, 65] consider EMG a good indicator of mental state and use it, along with other physical

and physiological signals, to distinguish negative and positive mental conditions. In [64], mechanical parameters of the vehicle are used, in addition to SCR and ECG data, to detect stress and testing several features' datasets and it is realized that EDA measurements are extremely useful for detecting stress. On the other hand, in [30] several tests are done to find the optimal feature set for drowsiness detection. Differently from stress, it seems that EDA measurements are less efficient to recognize drowsiness. These results suggest that the experimental task has to be well defined to make the right choice about the employed signals. This consideration also concerns features belonging to the same signal. During experiments for analyzing the influence of GPS instructions on a driver in [115], an evaluation of HR features concluded that MeanRR, SDNN and HRVTri are the most effective ones to recognize sudden changes in drivers' mental stress during short time lapses.

There are several works that employ a large number of inputs related to physiological [18] and/or other kinds of data. Authors in [90] warn about the drawbacks of relying on multi-modal measurements, such as the computational cost, or the redundant features that impede target classification learning.

In [92], the available dataset includes physiological, video and environment data; then several test are carried out to rank the usefulness of features to recognize stress and fatigue. In [29], the optimal set of features, including PPG, GSR, temperature (TEM), acceleration (ACCEL), and the rate of rotation (GYRO), is selected through algorithms based on variance analysis (ANOVA) and sequential forward floating selection (SFFS). The approach used in [73] differs from the previous mentioned because the features are not only selected but also automatically extracted by the Time Series Feature Extraction based on Scalable Hypothesis tests (TSFRESH) system, which extracts and selects hundreds of features. Automatic feature extraction methods are especially common for electroencephalogram (EEG) [49], since EEG requires data from many channels.

In [32], besides stress detection, method for stress reduction are also reviewed.

The aspect of driver and passengers mental state is critical in autonomous driving as it is in manual driving. In [50], authors have provided to the subjects a chaffeur, who plays the role of an autonomous driving system, for monitoring their behavior about their car trips when they do not need to drive themselves. Usually people are not worried about the driver behavior, but this may not be the same for autonomous driving systems. Then, the analysis of the passenger mental state is needed to check how much people trust autonomous driving systems and the social impact of this new technology. Moreover, switching to manual driving could always occur, for example when a pedestrian suddenly crosses the street, especially in dense traffic [44]. In [116], it is shown how the vehicle sensors performances deteriorate during adverse weather conditions. Studies have shown how people drowsiness rises during autonomous driving, deteriorating their reactions when they have to take control of the vehicle [39]. Even if the drivers have to monitor the driving system, after some time their attention decreases and in some cases they fall asleep in about 60 minutes [71]. This brought researchers to search for tasks that do not make people get drowsy during trips and at the same time do not distract them in order to be able to act when it is necessary [38], [58]. Authors in [84], developed a system, based on eye-tracking techniques, that highlights pedestrians and other critical objects if the subject is distracted by the tasks which are assigned to them during the trip. In [112], authors presented an advanced eye-tracking system which is able to analyze multiple events at the same time and to classify them on the basis of their importance in an automated driving scenario. Moreover they created the Berkeley DeepDrive Attention (BDD-A) database with the experimental data. In [19] and [82], EEG is employed to check the drowsiness and the emotional state of self driven vehicle passengers.

1.3 Contribution

The contribution of this work is focused on the following points:

Employment of Skin Potential Response (SPR) signal
In the clinical sector we can find several studies on the behavior of endosomatic Electrodermal measurements, known as Sympathetic Skin Response (SSR) or Skin Potential Response (SPR), connected with sweating [31], muscle contraction [107], electrical and magnetic stimuli [99, 37]. In [108, 109, 110] the rate of waveforms that compose the SPR signal in response of stimuli is studied.
The technological sector however, hasn't take advantage of SPR enough for designing systems able to automatically detect psychological state. As we can deduce researching the state of the art related to psychological state detection and monitoring applications, Skin Conductance has always been chosen among EDA measurements.
Our purpose then, is to investigate how Skin Potential can be employed for psychological state analysis applications, as an alternative to Skin Conductance, and how we can take advantage from its properties.

Motion Artifact (MA) removal

Since EDA measurements are usually taken from the hands, the measured signal has a strong component related to hand movements. This component is defined as Motion Artifact (MA) and it is an affliction when taking EDA measurements during driving activity. If not appropriately processed, it can make the acquired data completely useless.

There are not related works dedicated to Motion artifact impact on Skin Potential. Regarding Skin Conductance instead, several methods to detect and correct MA interferences can be found. In [118], an unsupervised machine learning method can successfully recognize and label MA interferences. The system developed in [62], is able to correct minor MA interferences employing sparse recovery with the aid of appropriate atoms. In [28],

a system based on wavelet transform is able to clean SCR from MA identifying the waves that characterize the signal.

In our application, the methods above cannot work because drivers could need to move the hands suddenly and quickly to drive in a proper way and MA energy equals or even overtakes the energy of the emotional component we are interested in. This effect is mostly marked in the SC phasic part, SCR, whose energy can be smaller that that of the disturbances introduced by MA.

We developed two ways to remove the Motion Artifact component from SPR. In the first approach, we relate MA to hand movements with an adaptive filtering techniques. In the second approach, we measure SPR in both hands and extrapolate the emotional component through a smoothing selection process.

Feature extraction and classification
The final step is to correctly classify the subject state. For this purpose it is essential to choose the right features, that are the ones which represent the presence or the absence of stress. Useful features for detecting mental state extracted from SPR, differ from the kind of features commonly used when SCR is employed. This is due to the differences between SPR and SCR signals. Furthermore, SPR features that work for machine learning classification is a topic which has not been completely investigated, then an evaluation has to be made to evaluate which kind of features we can use for our purpose. Once the features have been collected, we employ Support Vector Machine and Neural Networks to reveal the mental stress of subjects under test.

Chapter 2

Skin Structure and Electrodermal Properties

In this chapter, we first describe the structure of the skin and the mechanism of the sweating process. The explanation will include the necessary knowledge to understand the signals employed in this work. The reader can refer to [21, 94], to deepen the topic. Secondly, we show the principal electrical models created to resemble the electrical behavior of the skin during the sweating process.

2.1 Skin Structure

The skin is composed by several layers which are listed in Table 2.1 and depicted in Figure 2.1

The epidermis, Figure 2.2, is mainly composed by keratinocytes cells which are produced in germinativum layer. The growing

TABLE 2.1: Skin structure

Cutis (skin)	Epidermis	Stratum corneum Stratum lucidum Stratum granulosum Stratum spinosum Stratum germinativum
Subcutis (hypodermis)	Dermis (cutis vera)	Stratum papillare Stratum reticulare

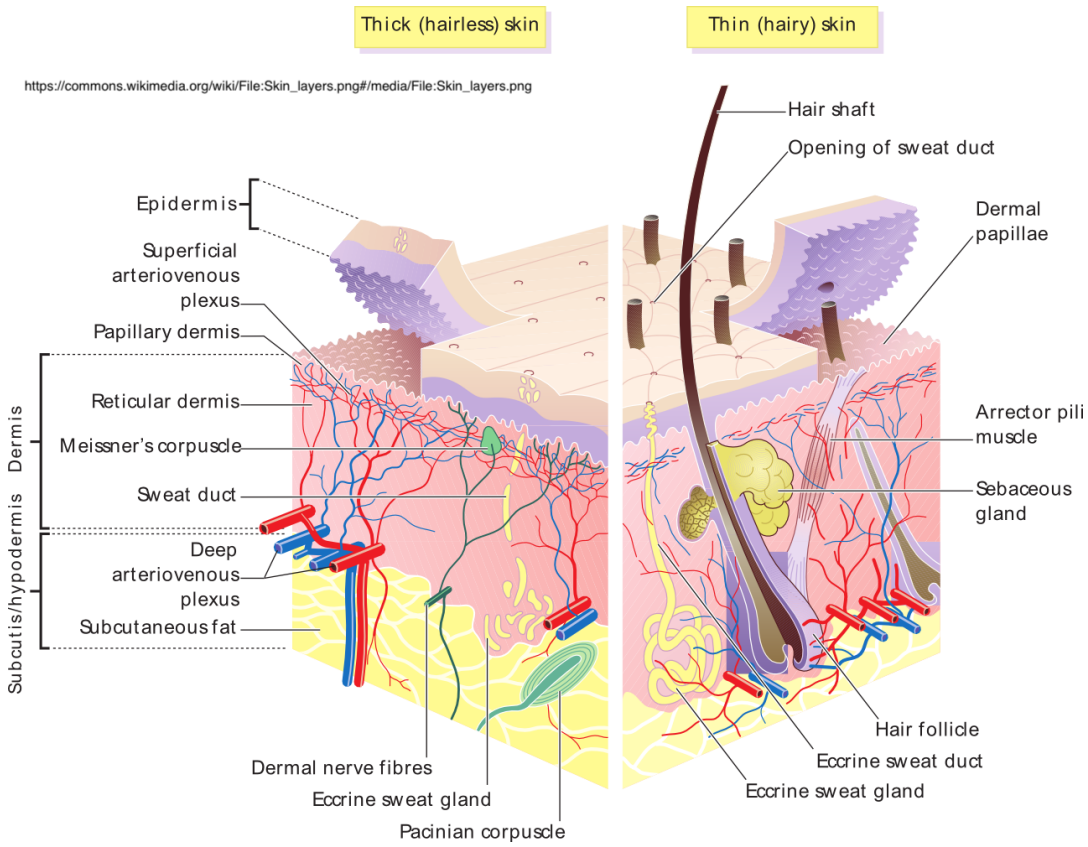


FIGURE 2.1: Skin structure.

pressure of the cells pushes keratinocytes through the upper layers. During this process, keratinocytes gradually absorb keratine and take a hornier shape until they reach the corneum layer becoming full exfoliated horny plates.

The dermis, also labeled as the corium, is much thicker than the epidermis. In the dermis, arterial and venous blood vessels and receptor organs are present. The terminal parts, which form the capillary net, are located in the papillar sub-layer along with melanocytes, which produce melatonine (skin pigment). The collagen is present in both the sub layers, but in the reticular sub-layer it is much stronger, giving high resistance to the skin.

The hypodermis is a connecting tissue that allows horizontal mobility to the skin on the muscles surface. In the hypodermis the fat is stored for thermal and mechanical insulation. It also contains nerves and vessels which supply the skin, the hair follicles and glands and the secretory part of the sweat glands.

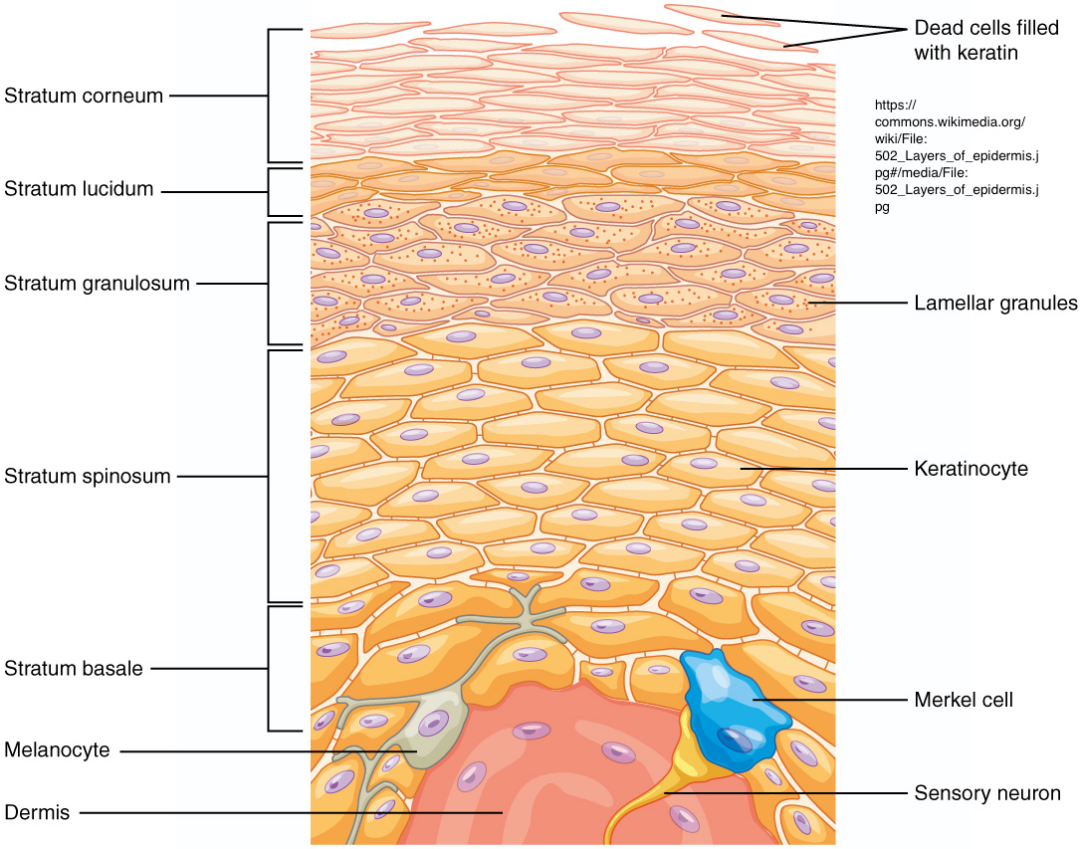


FIGURE 2.2: Epidermis.

Sweat glands are distributed across several parts of the body. They can be of eccrine types, which represent the majority of the total sweat glands, and apocrine types, which are exclusively located in hairy areas of the skin and loose cytoplasm in addition to the sweat.

The eccrine sweat gland can be subdivided into the secretory segment, which is located in both dermis and hypodermis, and is formed by layers of outer basal and inner luminal cells, and the duct, which is located in the epidermis and has no cells on its own.

In addition, skin contains several types of other glands, receptors, hair, muscles (i.e., the pilo-erecting muscle). For details, the reader can refer to [21].

2.2 Electrodermal Activity

Sympathetic fibers, which are innervated along the skin, activate the eccrine sweat glands secretory segment; the terminals of sympathetic fibers are called sudomotor fibers. The terminals reach as well the pilo-erecting muscles, whose task is to erect the hair, retract the skin and compress the sebaceous glands between them. Finally, the fibers reach the blood vessels, in order to regulate the blood flow, constituting the vasoconstrictory fiber system. Besides sudomotor and vasoconstrictory fibers, the skin is innervated by motor and sensory fibers as well by the central nerves from which the peripheral branches come from [21].

Sweat glands are distributed across several areas of the body. Particular attention has been dedicated to the sweat glands located on the palmar and plantar sites. Several experiments reveal that, differently from the sweat glands present on the rest of the body, palmar and plantar sweat glands don't take part to the body thermo-regulation process, unless the body reaches extremely high temperatures. On the other hand, palmar and plantar sweat glands are exceptionally active in the emotional sweating process.

The secretory segment of the sweat gland, which is surrounded by a layer of myoepithelial cells, consists of clear and dark secretory cells. The clear cells produce the liquid part of secretion, while the mucin molecules, produced by the dark cells, may have protecting functions within the lumen cavity of cells. Through the serous cells, water and ions pass from the plasma to the lumen. While the sweat passes through the duct, it looses most of the ions because of the ions reabsorption mechanism. The sweat moves from the duct to the skin surface in a pulsating manner, by means of rhythmic contractions of the myoepithelia that surrounds both the secretory and the ductile part. Sweating have several functions. One of them is thermoregulation, which regulates the body temperature by sweat vaporization. Another cause is due to the person emotional condition, that could stimulate the sweating glands. Also, sweat can arise while eating an

especially sour, salted or spicy food or when a lesion to the sympathetic nerve occurs. Reflex sweating occurs when a skin area is stimulated by radiation, heat, needle punctures, or electricity. Furthermore, sweat can also be caused by pharmacological induction.

2.3 EDA dynamics and electrical properties of the skin

The structure of the skin and the sweating process influences the electrical properties of the skin, on which electrodermal measurements, skin conductance, skin resistance and skin potential, are based. Electrical properties of the skin are usually classified in two major branches: passive and active. The dynamic of skin conductance is modeled using passive properties while the dynamic of skin potential is modeled by both active and passive properties.

The epidermal barrier, formed by the lower corneal zone, provides electrical resistance to the skin, since it is relatively impermeable to water and solutions. Electrical resistance depends on the amount of lipids and essential fatty acids that permeate the cell membrane, which prevents water diffusion, and skin temperature, whose increase goes along with an increase of water permeability of the skin. The other parts as the dermis and hypodermis don't contribute much to the skin resistance. The electrical resistance of the corneum depends on its hydratation. The factors that influence corneal hydratation are enivironment humidity, sweating, skin reabsorbtion, rate of sweating and age, because the skin becomes drier along with aging. Another property which influences skin electrical resistance, is the electrolytic content of the skin, which allows ions to flow through the skin. Furthermore, the skin presents capacitive properties due to cell membranes, which get polarized during the flow of ions.

Besides the passive properties, the skin has active membranes that regulate sweating by changing their potential when stimuli occour. These membranes are located in the secretory part

of the sweat glands, in the epidermal duct at the level of the stratum germinativum, in the inner corneal zone, in the dermal and in the epidermal part of the duct during the reabsorption of sodium successively to the sweating and in the myoepithelial cells.

2.4 EDA circuital models

Several circuital models have been developed to simulate the behavior of Electrodermal activity, especially focusing on the sweating process. In this section the main models, that differ on the focusing aspects and in the relative mechanisms, details will be described.

Figure 2.3 shows the Montagu and Coles model [77] , which is widley used to predict the dynamic of skin conductance. In the circuit, R_1 and R_2 represent the resistance of derms and stratum corneum respectively, $r_1.....r_n$ represent the sweat gland ducts resistances and are connected to the circuit by the shunts, whose dynamic depends on the sweating process. The sweat gland duct resistors can be modeled as a single variable resistance which varies on the basis of the global dynamic of the sweat ducts. It can be seen that the circuit is composed by passive components only, since the changing of the sweat duct resistance is represented by the shunts.

The voltage divider model, depicted in Figure 2.4, has been developed by Edelberg [94]. It is employed to characterize skin potential excursion. The skin potential measured at point P, depends on the voltage sources S and E, that represent the potential of the sweat glands and the corneum, the current ripartition between the resistance of sweat duct R_S and the resistance of extraductal epidermis R_e. If the resistance of the sweat gland decreases, for example during the filling of the ducts, the surface potential goes more negative. If the resistance of the corneum decreases, when the sweat is expelled, the surface potential goes more positive.

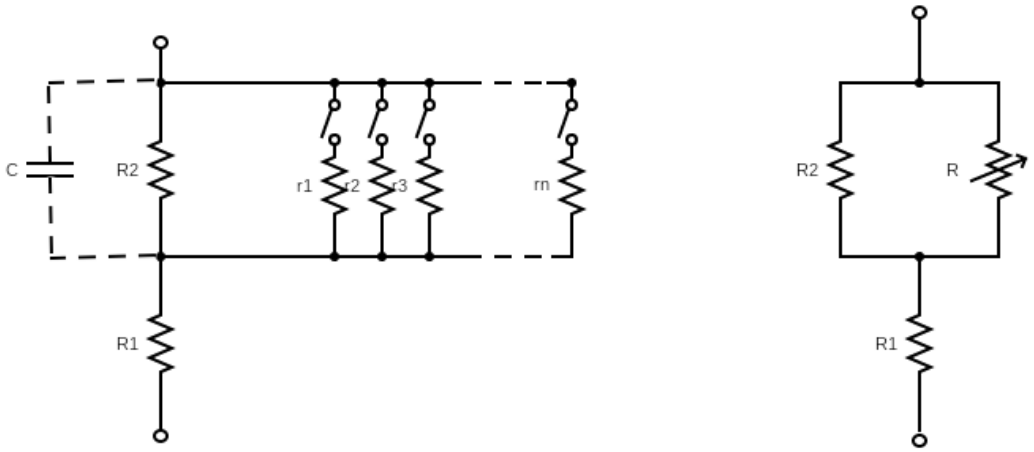


FIGURE 2.3: Montagu-Coles circuital model for EDA).

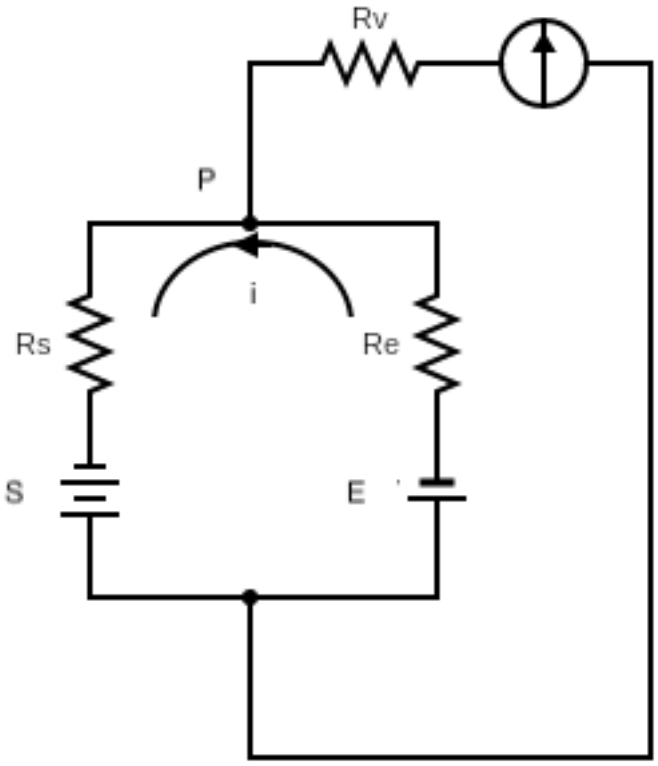


FIGURE 2.4: Voltage divider EDA model.

Starting from the Edelberg model, Fowles developed a more complex version showed in Figure 2.5. E_1 originates in the ductal wall in the dermis, and is mainly caused by the sodium ionic

concentration inside the cells. The dermal duct resistance is represented by R_4 and R_2, which depend on hydratation and duct filling respectively.

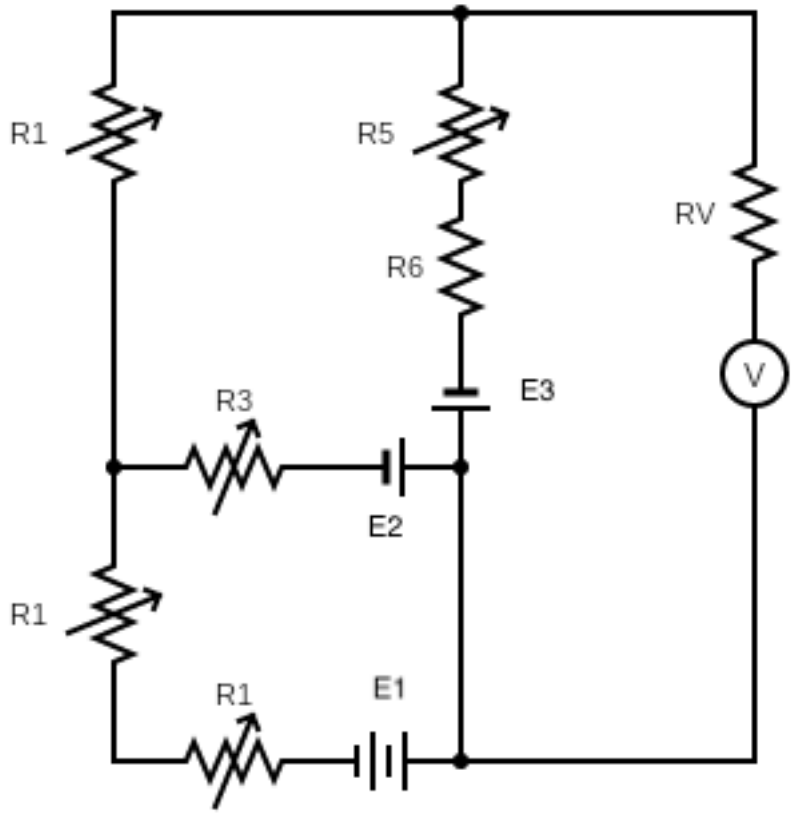


FIGURE 2.5: Fowles EDA model.

E_2 depends on the concentration of sodium and chloride ions at the stratum stratum germinativum level of the duct. E_2 is smaller than E_1 because the membrane in the stratum stratum germinativum is less selective on ions flow. In the stratum stratum germinativum, the resistances of the duct are modeled by R_3 for the duct wall, and by R_1 for the epidermal part of the duct, depending on the duct filling. E_3 represents the localized membrane potential, located in the lower zone of the stratum corneum and depends on potassium ionic concentration in the interstitial fluid. R_6 is the resistance of the compact keratinized layer zone, R_5 is the resistance of the upper layers of the corneum, varying with their hydration. R_6 varies very little due

to its low permeability, then its resistance can be assumed constant in this model.

2.5 Electrodermal Signals

An example of a Skin Conductance (SC) measured signal is plotted in Figure 2.6. The signal is composed by a tonic part, Skin Conductance Level (SCL), and a phasic part, Skin Conductance Response (SCR). SCR varies faster than SCL and is caused by the phasic electrodermal activity. SCR reveals reactions to sudden stimuli and sudden emotional state changes while SCL represents SC when no stimulus occurs. SCL and SCR are plotted respectively in the top and bottom graphs of Figure 2.6. SCL and SCR have been obtained from SC through non-negative deconvolution decomposition explained in [15], using the dictionary composed by sigmoid-exponential, Bateman and chi-square waveforms developed in [26].

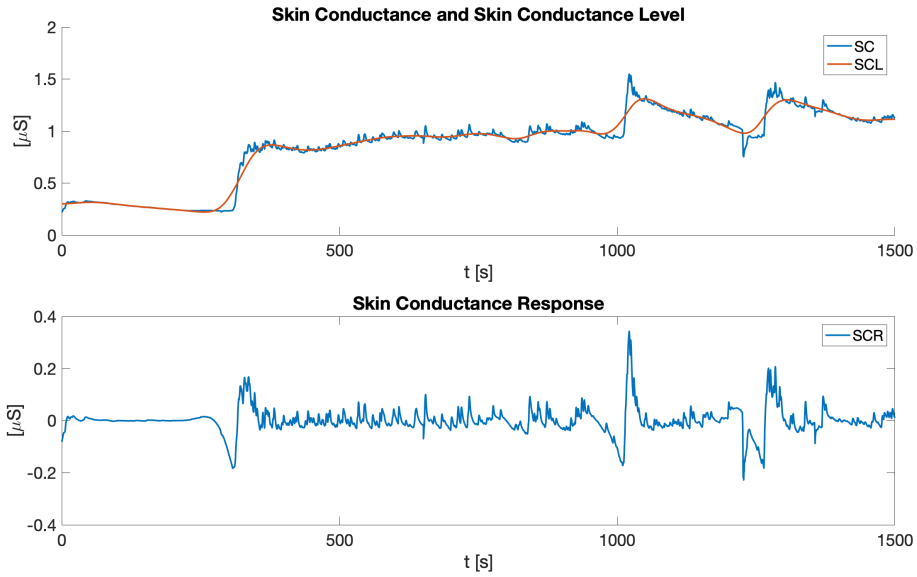


FIGURE 2.6: SC and SCL measurements (top). SCR measurements (bottom).

In Figure 2.7 some examples of waveforms belonging to single SCR impulses, modeled as Bateman functions [26] are plotted.

These impulses arise due to the sweating process. Skin conductance quickly arises while the sweat is being expelled, then the sweat evaporates and skin conductance slowly returns to its resting value represented by the skin conductance level for the SC signal or by zero for SCR. In addition to SCL, the analysis of the signals typically relies on the properties of SCR impulses, in particular on magnitude, energy, rise time, recovery time, impulse frequency, slope of the impulses [21].

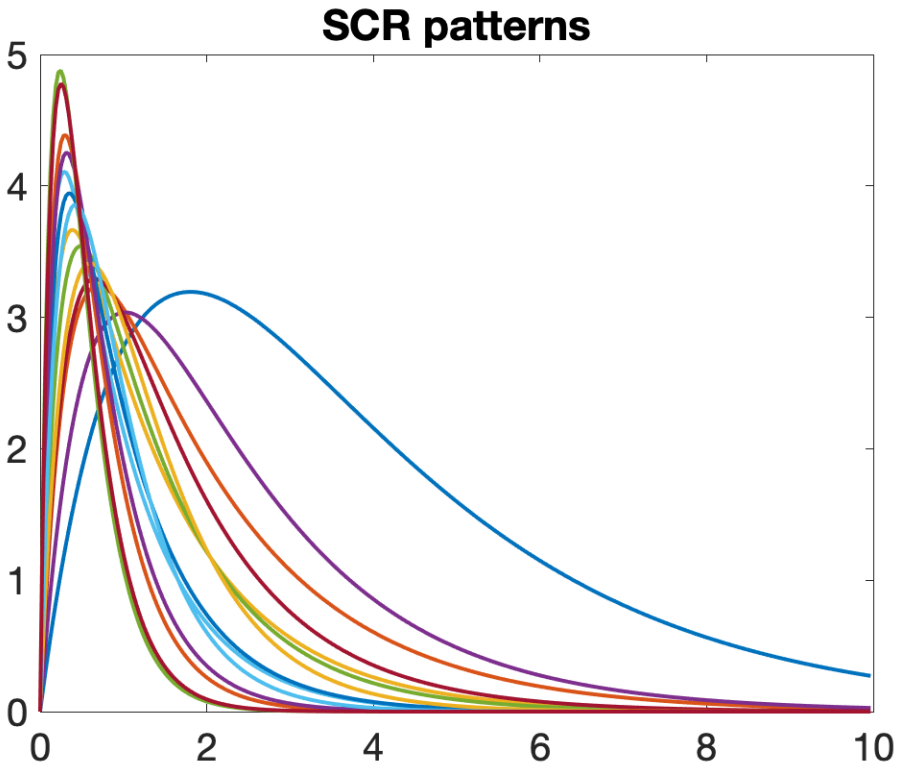


FIGURE 2.7: SCR wavelet patterns modeled as Bateman functions [26].

Skin Potential, on which our work is focused, can range between 10 and -70 mV. As Skin Conductance, Skin Potential has a slowly varying tonic component, Basal Skin Potential Level (BSPR) and a fast varying phasic component, Skin Potential Response (SPR). Differently from SCR, SPR waves can be composed by either positive and/or negative components. The positive component (P pattern in Figure 2.8 (a) is caused by hydration changing which, as in SCR, happens when the sweat is expelled on the skin surface from the duct, followed by a rapid

recovery. On the contrary, the negative component in Figure 2.8 (b) appears as a reaction to stimuli when the sweat ducts are still recovering after sweating. In this type of reaction, sweat is not expelled [42].

Depending on the electrodermal activity, different wave types could occur. The test undertaken in [107] involves repeated stimuli, and shows how the positive waves are more common during the first session. In the following sessions, the trend reverses and the subjects present more negative then positive waves due to habituation. The situation switches once more when a different type of stimuli occurs and the rate of positive waves, as reactions to the stimuli, raises again. SPR impulses don't always present a single wave type and often we could have waves composed by both positive and negative components (M pattern in Figure 2.8 (c)). As described in [108], M pattern waves could vary, in amplitude and latency, more than N and P waves, considering the same number of repeated stimuli.

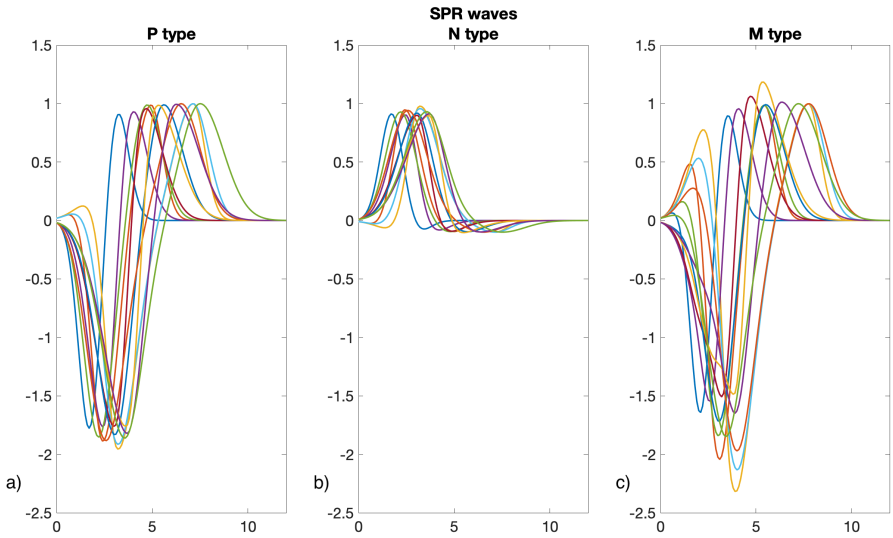


FIGURE 2.8: SPR wavelet patterns modeled as combination of Gaussian functions.

2.6 Conclusions

In this chapter Electrodermal Activity has been described under the physiological aspect to provide a background about the mechanisms behind the signals we measure and use in this work. The medical research which have been cited in this chapter, can give a wider comprehension on how the EDA signals behave. Although in our context the behavior of EDA will be more complicated, this is a fundamental base from which to start our research. The equivalent circuital models in the end, provide patterns that we can recognize to analyze the signals we measure. In Chapter 4, it will be explained how we apply EDA signals to carry out our purpose to automatically detect stress mental state in vehicle drivers.

Chapter 3

Electrocardiogram

In this chapter, we introduce the functioning of the cardiovascular system to give the basis to understand the information carried by a signal which depicts the heart activity (Section 3.1). Section 3.2 is about the characteristics of the Electrocardiogram (ECG) and the interpretations of the waves to evaluate if the subject is in a healthy state or could suffer some diseases. Section 3.3 is dedicated to the processing of ECG to detect the ECG waveform patterns and obtain a signal that represents the excursion of the Heart Rate (HR). Section 3.4 gives an overview of the Heart Rate Variability parameters and how they can provide information about a subject state, including mental stress.

3.1 The cardiovascular system

The functions of the cardiovascular system are: the transport of nutrients, oxygen, and hormones to the cells and the removal of metabolic waste; the body protection from foreign microbes and toxins by white blood cells and antibodies; the mechanism of coagulation to prevent blood loss after injuries; the regulation of the body temperature, fluid pH, and water content of cells.

The cardiovascular system is a closed double circuit (see Figure 3.1):

◦ the pulmonary circuit carries blood to and from the lungs for gaseous exchange; it transports the venal blood into the lungs through the right side of the heart, where the carbon dioxide leaves the system and the blood is loaded with oxygen, then the

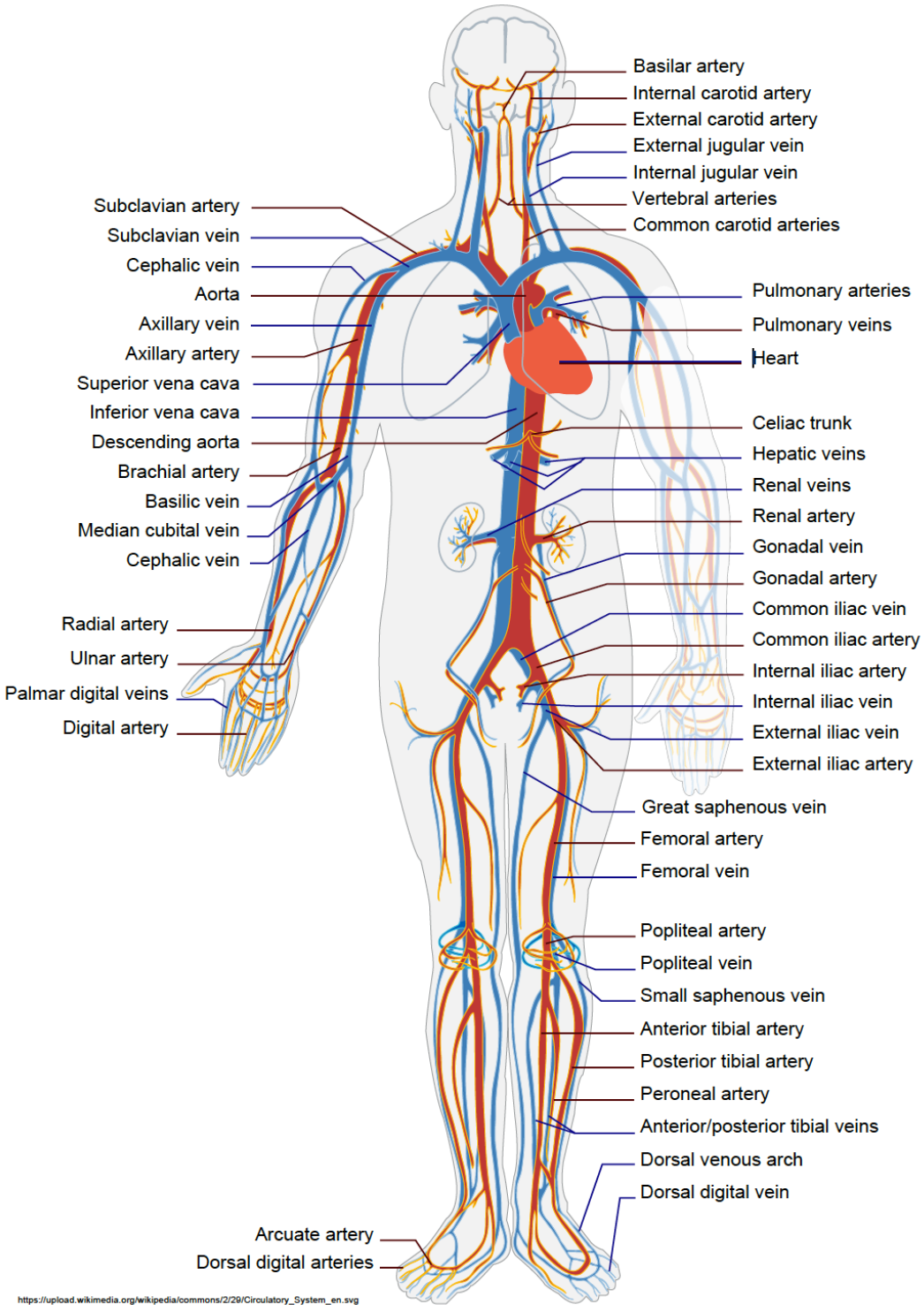


FIGURE 3.1: Human circulatory system. Red: arteries. Blue: veins.

blood is carried to the systemic circuit through the left side of the heart

- the systemic circuit transports the blood from the left side of

the heart, through the arteries, arterioles and capillaries, to the organs and tissues where oxygen and nutrients are delivered. The blood then, loaded with carbon dioxide, returns to the pulmonary circuit which directs the blood to the heart, where the circulation restarts.

The cardiovascular system is activated by the heart, that pumps the blood through the circuit. The heart has four chambers: two atria (left and right) and two ventricles (left and right). The atria are located in the upper part of the heart and pump blood to the respective left and right ventricles after receiving blood from the veins. The right atrium receives deoxygenated blood from the superior vena cava, then pumps the blood into the right ventricle. The right ventricle pumps the blood into the pulmonary artery, through which it reaches the lungs, where the blood oxygenation occurs. Oxygenated blood leaves the lung via the pulmonary vein and enters the left atrium. The left atrium pumps the blood into the left ventricle, which is much thicker than the right one. Finally, blood is pumped into the aorta, which carries oxygenated blood around the body.

The heart activity is activated by the sinoatrial node, which produces the action potentials that activate the cardiac muscle, the myocardium, which contracts the heart chambers, pumping the blood trough the system. Furthermore it acts as a natural pacemaker and regulates the heart beat rytmh.

For a more detailed description of the cardiovascular system see [43].

3.2 Electrocardiogram

The electrical activity of the cardiovascular system is well represented by Electrocardiogram (ECG). ECG depicts the excursion of the electrical charges that polarize and depolarize cardiac muscles.

ECG is composed by three waveforms, Figure 3.3

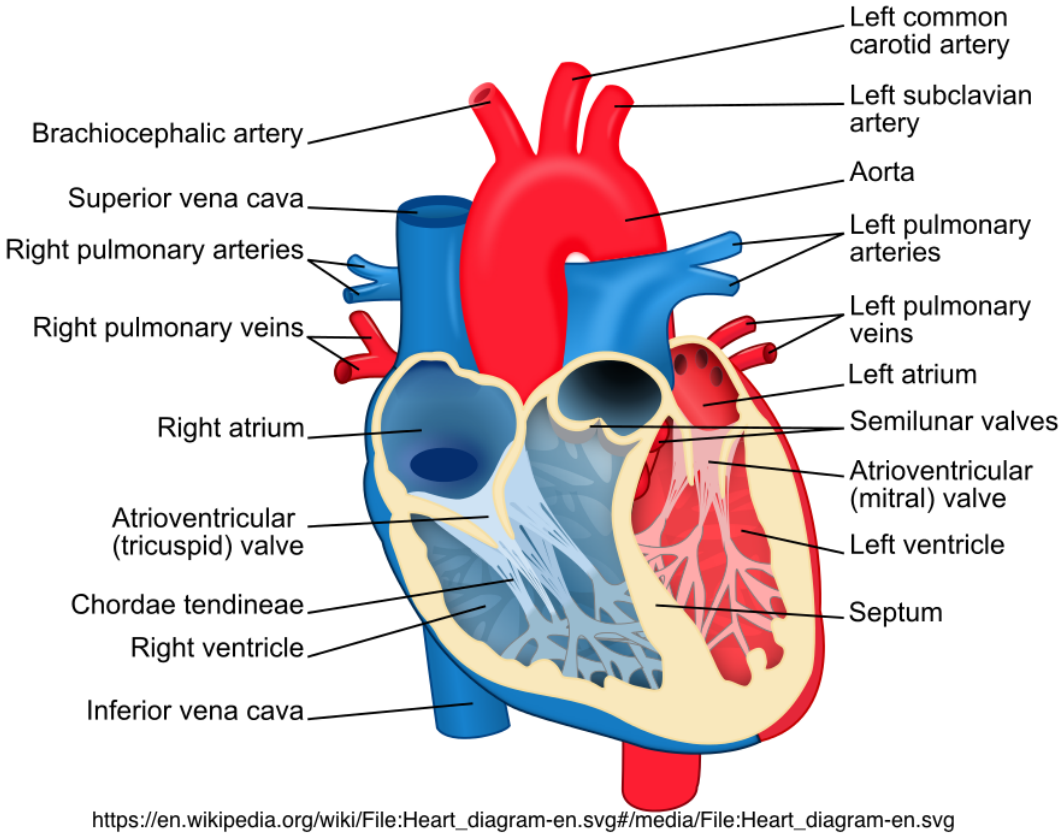


FIGURE 3.2: Human heart

- The P wave on the ECG represents atrial depolarization, which results in an atrial contraction

- The QRS complex represents the diffusion of the electrical stimulus through the myocardium

- The T wave represents the repolarization of the ventricles; in this phase we have a relaxation of the ventricular myocardium.

ECG provides information about the health state of the subject; abnormalities in ECG give indication of possible diseases. Abnormalities can regard magnitude and/or duration or even distortion or wave absence. Figure 3.4 and Figure 3.5 show some examples of ECG waveforms that indicate potential diseases.

ECG is related to a person emotional state too. Emotions influence the ECG cycle rate which can be used to estimate mental state; [57] and [101] present experiments during which several type of emotions and their intensity are estimated on the basis

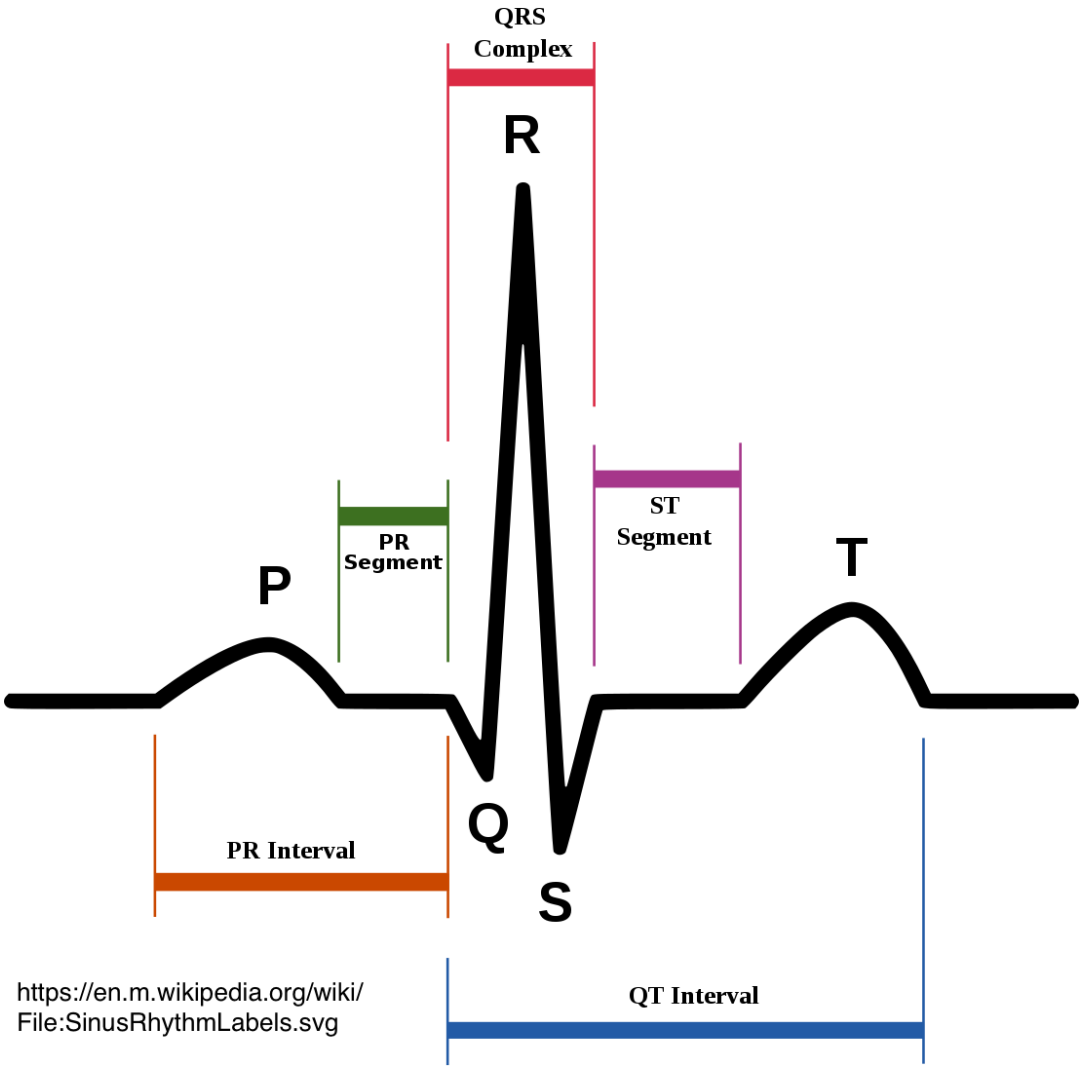


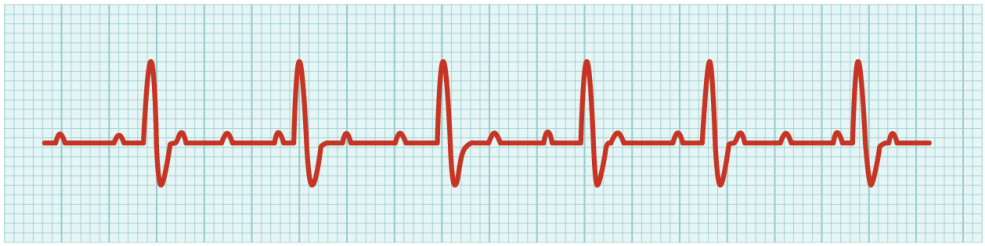
FIGURE 3.3: A group of patterns created to represent the ECG signal waves.

of ECG characteristics. For a comprehensive view of ECG interpretations the reader can refer to [41]

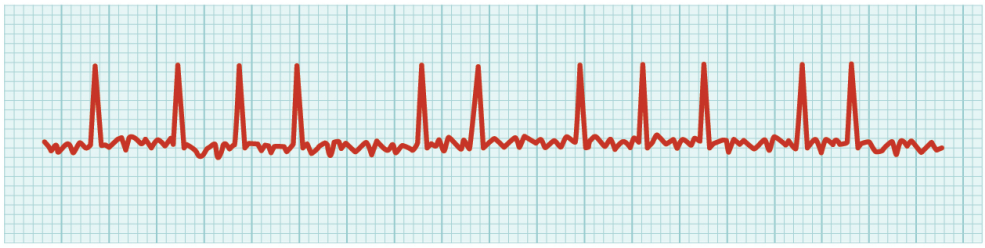
3.3 ECG processing

Emotional state recognition through cardiac activity, is often based on Heart Rate Variability (HRV) parameters which have to be extracted by Heart Rate (HR) and RR peaks signals. HR and RR signals are, in turn, extracted from ECG.

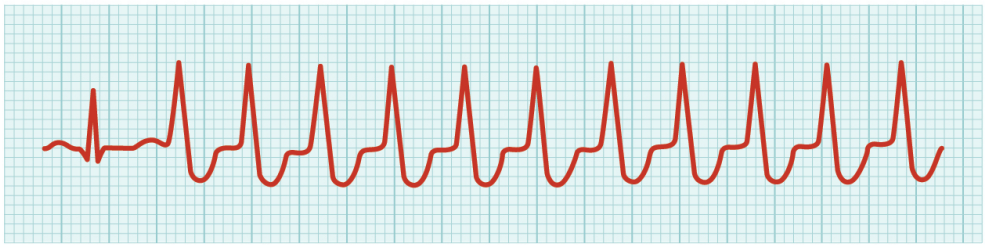
In this section, the ECG processing to obtain the parameters we need is depicted, describing operations and algorithms which



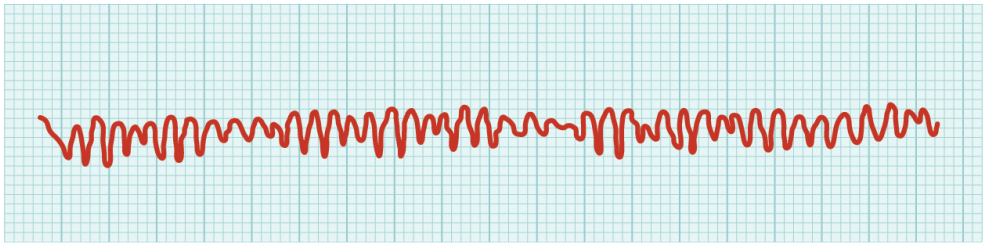
(a) Second-degree (partial) block



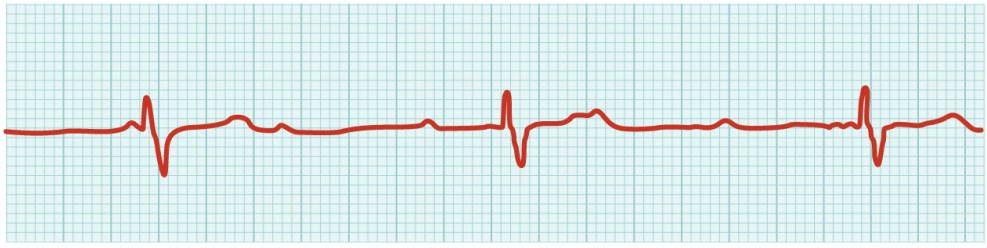
(b) Atrial fibrillation



(c) Ventricular tachycardia



(d) Ventricular fibrillation



(e) Third-degree block https://upload.wikimedia.org/wikipedia/commons/b/ba/2024_Cardiac_Arrhythmias.jpg

FIGURE 3.4: Abnormal ECG waveforms related to different types of diseases

have been developed for this purpose.

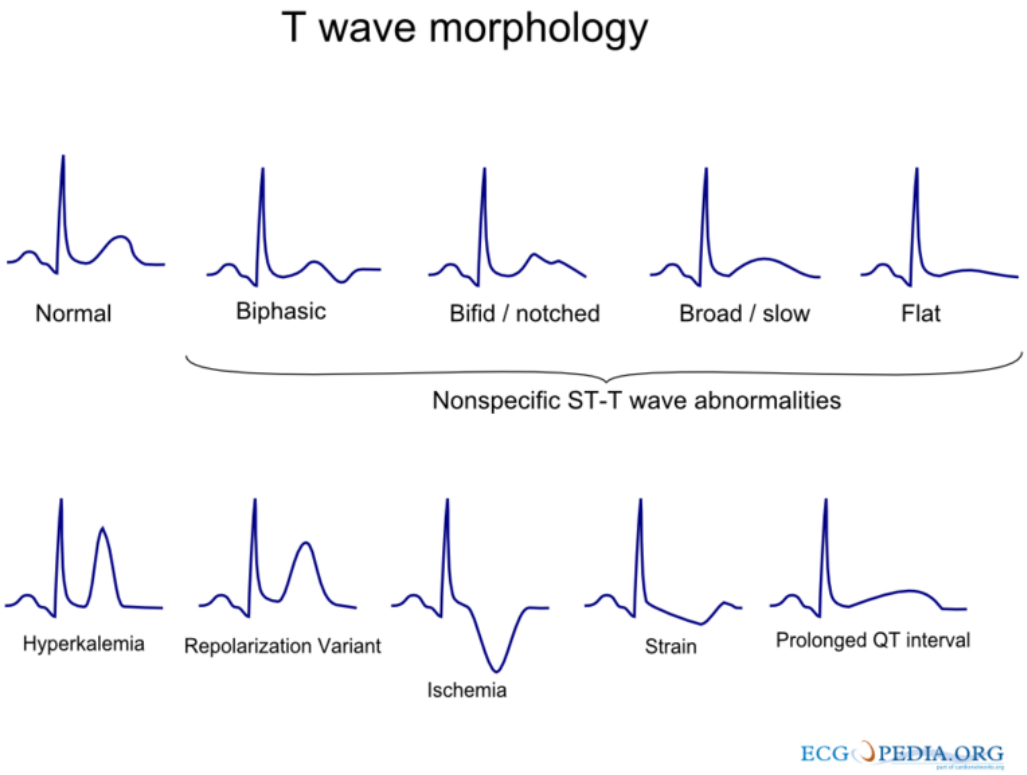


FIGURE 3.5: Abnormal T waves

3.3.1 QRS detection

The fist step is to detect the pattern of QRS complexes in ECG waveforms. In our work, we use the Pan-Tompkins algorithm, which has been developed to recognize QRS complex in noisy ECG signals [80]. In the following, the Pan-Tompkins algorithm steps (Figure 3.6) will be described. The ECG signal is band-pass filtered. The band pass filter is designed as a cascade of a second order low-pass filter followed by a first order high-pass filter to have more design flexibility for the passband.

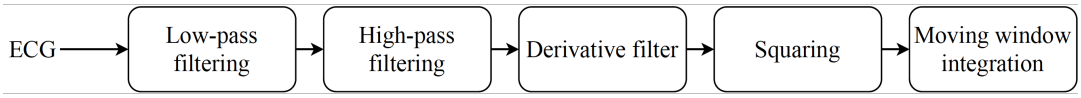


FIGURE 3.6: Pan-Tompkins algorithm block scheme.

After the filtering, the signal is differentiated to provide the QRS complex slope information (Figure 3.8 (top)) and then squared to emphasize peaks (Figure 3.8 (middle)). An integration is then carried out (Figure 3.8 (bottom)) to obtain the waveform feature

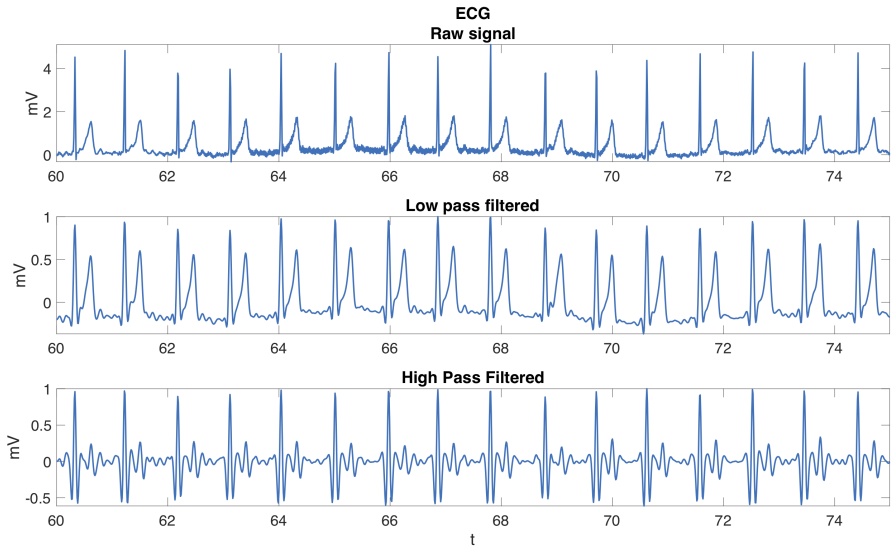


FIGURE 3.7: ECG signal: Raw (bottom), processed with low-pass filter (middle), processed with highpass filter (bottom)

information in addition to the slope of the R wave. The integration window length has to be chosen considering the widest QRS complex. If the window is too wide, the integrated waveforms will merge the QRS and T complexes together. If it is too narrow, some QRS complexes will produce several peaks in the integration waveform. These can cause difficulty in subsequent QRS detection processes. The QRS complex corresponds to the rising edge of the integration waveform. The time duration of the rising edge is equal to the width of the QRS complex. A fiducial mark for the temporal location of the QRS complex can be determined from this rising edge according to the desired waveform feature to be marked such as the maximal slope or the peak of the R wave.

To reduce the possibility of wrongly selecting a noise peak as a QRS, each peak amplitude is compared to a threshold that takes into account the available information about already detected QRS and the noise level.

Two sets of thresholds are automatically adjusted to float over the noise. Low thresholds are possible because of the improvement of the signal-to-noise ratio by the bandpass filter.

The higher of the two thresholds in each of the two sets is used for the first analysis of the signal. The lower threshold is used

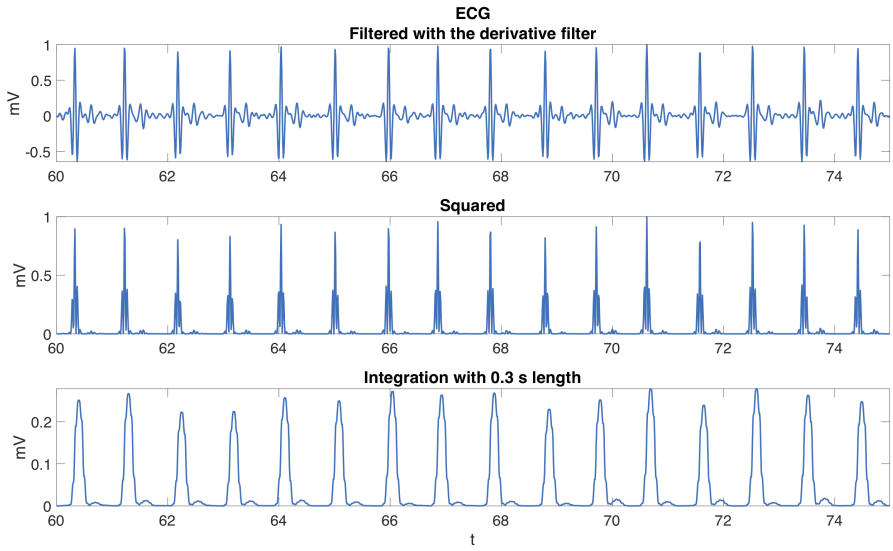


FIGURE 3.8: ECG: processed with derivative filter (top), squared (middle), integrated (bottom)

if no QRS is detected in a certain time interval so that a search-back technique is necessary to look back in time for the missed QRS complexes.

The thresholds are based upon running estimates of $PK_{Isignal}$ and PK_{Inoise}, which are the signal and noise peaks respectively [80]. New values of these variables are computed in part from their prior values. When a new peak is detected, it must first be classified as a noise peak or a signal peak. To be a signal peak, the peak must exceed a threshold Thr_{I1} as the signal is first analyzed or a threshold Thr_{I2} if searchback is required to find the QRS. When the QRS complex is found using the second threshold, the noise level is updated. In Figure 3.8 (bottom), the estimated noise and signal level and the thresholds related to the ECG integrated signal are plotted.

For irregular heart rates, the first threshold of each set is reduced by half so as to increase the detection sensitivity and to avoid missing beats.

To be identified as a QRS complex, a peak must be recognized as such in both the integrated and bandpass-filtered waveforms.

Two RR-interval averages are maintained. One is the average calculated on the basis of the eight most-recent beats. The other

one is the average of the eight most-recent beats having RR intervals that fall within certain limits. The reason for maintaining these two separate averages is to be able to adapt to quickly changing or irregular heart rates. The first average is the mean of the eight most-recent sequential RR intervals regardless of their values.

$$AVG_{RR1n} = mean([RR_{n-7} + ... + RR_n]) \tag{3.1}$$

where RR_n is the most-recent RR interval. The second average is based on selected beats:

$$AVG_{RR2n} = mean([RR'_{n-7} + ... + RR'_n]) \tag{3.2}$$

where RR' is the most recent RR interval that fell between the acceptable low and high RR-interval limits. This is the case for normal sinus rhythm.

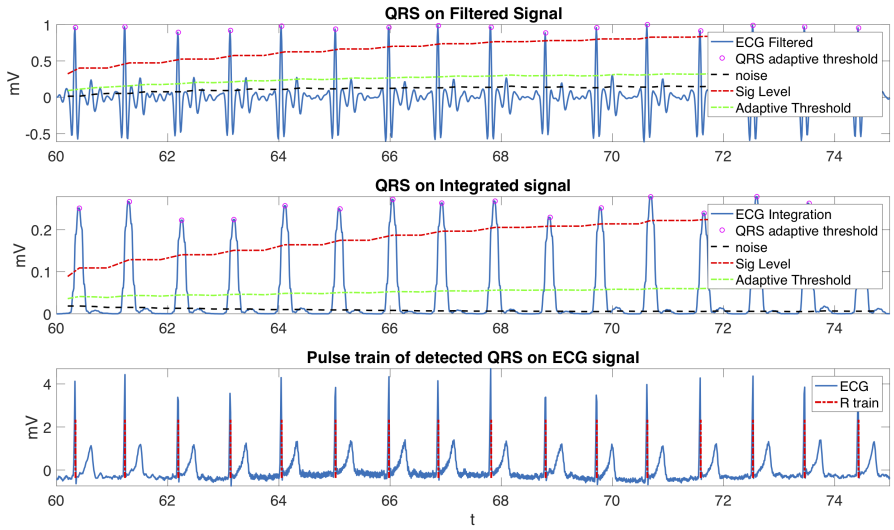


FIGURE 3.9: QRS detection on filtered ECG (top), QRS detection on integrated ECG (middle), ECG and pulse train of detected R waves (bottom)

Figure 3.9 shows QRS detection with the estimated signal and noise level and thresholds for Filtered (top) and Integrated (middle) ECG. The bottom graph depicts the ECG and the resulting QRS detection train.

When an RR interval is less than 360 ms (it must be greater than the 200 ms latency), a judgment is made to determine whether

the current QRS complex has been correctly identified or whether it is really a T wave. If the maximal slope that occurs during this waveform is less than half that of the QRS waveform that preceded it, it is identified to be a T wave; otherwise, it is called a QRS complex.

3.3.2 Artifact correction

Once QRS complexes have been detected, RR and HR signals can be computed, Figure 3.10:

$$RR = diff(t_R) HR = 60/RR \tag{3.3}$$

where t_R are the time instants of detected QRS complexes.

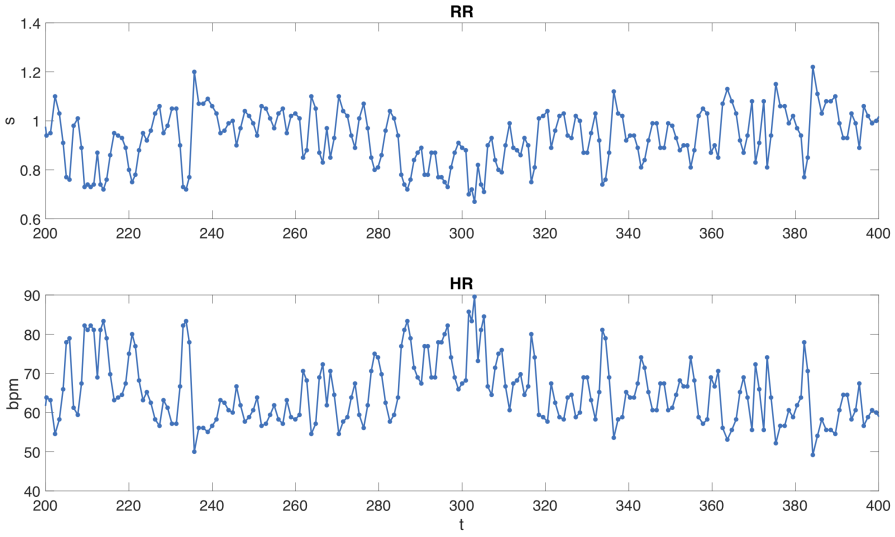


FIGURE 3.10: RR (top) and HR(bottom)

However, some errors could still be present due to artifacts. An error correction procedure then, is recommended. First, a smoothed version of RR, $RR_{smoooth}$, is obtained.

For this purpose, RR is processed by a Savitzky–Golay filter [97]. The Savitzky–Golay filter is preferable to a moving window averaging filter which could introduce bias in the presence

of local maxima [85]. Instead of averaging the signal, the Savitzky–Golay filter replaces a signal block, within a moving window, with a polynomial function of predefined order. The polynomial coefficients are selected according to the values reported in [97], based on the window length and the polynomial order. The Savitzky–Golay filter uses the polynomial coefficients that optimally fit the data, in a computationally much cheaper way than the least-squares fitting method.

Then, RR intervals whose difference with $RR_{smoooth}$ exceed a threshold are discarded and new values are obtained through interpolation.

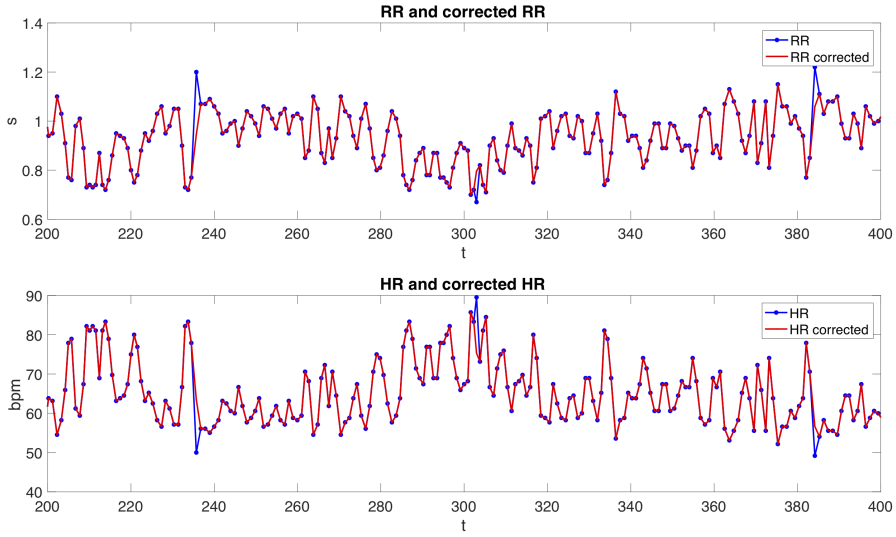


FIGURE 3.11: RR: before and after error correction (top), HR: before and after error correction (bottom)

In Figure 3.11, RR and HR before and after the artifact correction are plotted.

3.4 Heart Rate Variability

Heart Rate Variability (HRV) represents the amount of heart rate fluctuations around the mean heart rate and depicts the functioning of cardiovascular control system and Autonomous Nervous System (ANS) [91, 1].

The Autonomous Nervous System has sympathetic and parasympathetic components. Sympathetic stimulation, occurring in response to stress, exercise and heart disease, causes an increase in HR by increasing the firing rate of pacemaker cells in the heart's sino-atrial node. Parasympathetic activity, primarily resulting from the function of internal organs, trauma, allergic reactions and the inhalation of irritants, decreases the firing rate of pacemaker cells and the HR, providing a regulatory balance in physiological autonomic function.

Heart Rate Variability then, is an indicator for the health state of a person. HRV clinical applications cover a large number of fields: it can reveal cardiovascular and neurological diseases, monitor diabetes, renal and hypertension status. Also, HRV can highlight the effects of drugs, smoking, alcohol and it gives information about the sleeping quality and, as we can see from the references cited in 1, HRV is extremely useful to recognize stress.

In the following tables HRV parameters included in the standard developed by the Task Force of The European Society of Cardiology and The North American Society of Pacing and Electrophysiology [24], are listed. HRV parameters are sorted in statistical time domain features (Table 3.1), geometric time domain features (Table 3.2), frequency domain short-term features (usually taken in a window of some seconds or some minutes) (Table 3.3) and frequency domain long-term features (usually taken in a window of 24 hours) (Table 3.4).

In our work we will use HRV features (jointly with SPR features) to detect stress in driving subjects. We will not use all the feature listed in this chapter, but will do evaluations for finding the HRV features that are more suitable for our purposes.

TABLE 3.1: HRV time domain statistical features

Variable	Units	Description function
SDNN	ms	Standard deviation of all NN intervals.
SDANN	ms	Standard deviation of the averages of NN intervals in all 5 min segments of the entire recording.
RMSSD	ms	The square root of the mean of the sum of the squares of differences between adjacent NN intervals.
SDNN index	ms	Mean of the standard deviations of all NN intervals for all 5 min segments of the entire recording.
SDSD	ms	Standard deviation of differences between adjacent NN intervals.
NN50 count		Number of pairs of adjacent NN intervals differing by more than 50 ms in the entire recording. Three variants are possible counting all such NN intervals pairs or only pairs in which the first or the second interval is longer.
pNN50	%	NN50 count divided by the total number of all NN intervals.

TABLE 3.2: HRV time domain geometric features

Variable	Units	Description function
HRV triangular index		Total number of all NN intervals divided by the height of the histogram of all NN intervals measured on a discrete scale with bins of $7 \cdot 8125$ ms (1/128 s). (Details in Fig. 2)
TINN	ms	Baseline width of the minimum square difference triangular interpolation of the highest peak of the histogram of all NN intervals (Details in Fig. 2.)
Differential index	ms	Difference between the widths of the histogram of differences between adjacent NN intervals measured at selected heights (e.g. at the levels of 1000 and 10 000 samples).
Logarithmic index		Coefficient ϕ of the negative exponential curve $k \cdot e^{-\phi t}$ which is the best approximation of the histogram of absolute differences between adjacent NN intervals.

TABLE 3.3: HRV frequency domain short-term features

Variable	Units	Description function	
5 min total power	$[ms^2]$	The variance of NN intervals over the temporal segment	approximately $\leq 0.4Hz$
VLF	$[ms^2]$	Power in very low frequency range	$\leq 0.4Hz$
LF	$[ms^2]$	Power in low frequency range	$0.04 - 0.15$ Hz
LF norm	n.u.	LF power in normalized units $LF/(TotalPower - VLF)x100$	
HF	$[ms^2]$	Power in high frequency range	$0.15 - 0.4Hz$
HF norm	n.u.	HF power in normalised units $HF/(TotalPower - VLF)x100$	
LF/HF		Ratio $LF[ms^2]/HF[ms^2]$	

TABLE 3.4: HRV frequency domain long-term features

Variable	Units	Description function	
Total power	$[ms^2]$	Variance of all NN intervals	approximately $\leq 0.4Hz$
ULF	$[ms^2]$	Power in the ultra low frequency range	$\leq 0.003Hz$
VLF	$[ms^2]$	Power in the very low frequency range	$0.003 - 0.04Hz$
LF	$[ms^2]$	Power in the low frequency range	$0.04 - 0.15Hz$
HF	$[ms^2]$	Power in the high frequency range	$0.4 - Hz$
α		Slope of the linear interpolation of the spectrum in a log-log scale	$\leq 0.04Hz$

Chapter 4

Motion Artifact Removal

The main topic of the chapter, is the development of the Motion Artifact removal system for a Skin Potential Response signal.

In Section 4.1 we begin to analyze Skin Potential measurements describing the SPR behavior with some simple tests during which, stress is induced by stimulation and Motion Artifact disturbances are introduced. A comparison between SPR and SCR signals follows, along with an explanation about how, in our application, SPR employment can bring better results respect to SCR.

In Section 4.2 the main topic of this chapter, MA removal is challenged. We present a system which removes Motion Artifact, exploiting its relation with the hand movements related to the Steering Wheel rotation. The system is based on adaptive filters and is tested through experiments set up in laboratory.

In Section 4.3 another approach for Motion Artifact removal is presented. The system processes two SPR measured on both the hands to remove MA. Several experiments are carried out in laboratory and in more realistic car driving contexts, to test the method performances.

Section 4.4 discusses the possible real time implementation of the proposed systems.

The conclusions are discussed in 4.5

4.1 Preliminary test and the Motion Artifact issue

This section is dedicated, in the first part, to the preliminary measurements and analysis of Skin Potential Response (SPR).
We begin with a description of the laboratory instruments and the experimental setup, then we show the behavior of Skin Potential Response signal to stress stimuli and how Motion Artifact prevents to correctly recognize the Stress component of the signal.
In the last part we compare SPR and SCR measurements showing their differences and explaining the reasons of our choice to employ SPR.

4.1.1 SPR Sensor

Here the sensor employed for measuring Skin Potential Response is described 4.1 (a). For a deeper explanation see [4].

SPR signal is acquired with two electrodes posed on the palm and the back of the hand respectively; a third reference electrode is added on the wrist to force the reference voltage V_{REF} . Referring to the scheme shown in 4.1 (a), the differential voltage V_{IND} between the palm and the back of the hand is conditioned by an analog front end, acquired by a Digital Signal Processor (DSP) and sent to a Bluetooth module. The device is supplied with a single cell Lithium Polymer battery, whose voltage is converted into + 3.3 V through a buck DC-DC converter.

A linear voltage reference provides $V_{REF} = 1.65V$ to the reference electrode and also to the analog front end. The differential voltage V_{IND} is filtered by passive high pass filters, in order to remove the DC voltage that may be present on the skin; the input impedance of the instrument is 100 MΩ in the pass-band, by far higher than skin output impedance (in the order of 1 MΩ). With this specification, the load error is lower than 1% and, moreover, the changes in skin impedance (e.g., due to long term acquisition or temperature changes) are desensitized. Measuring

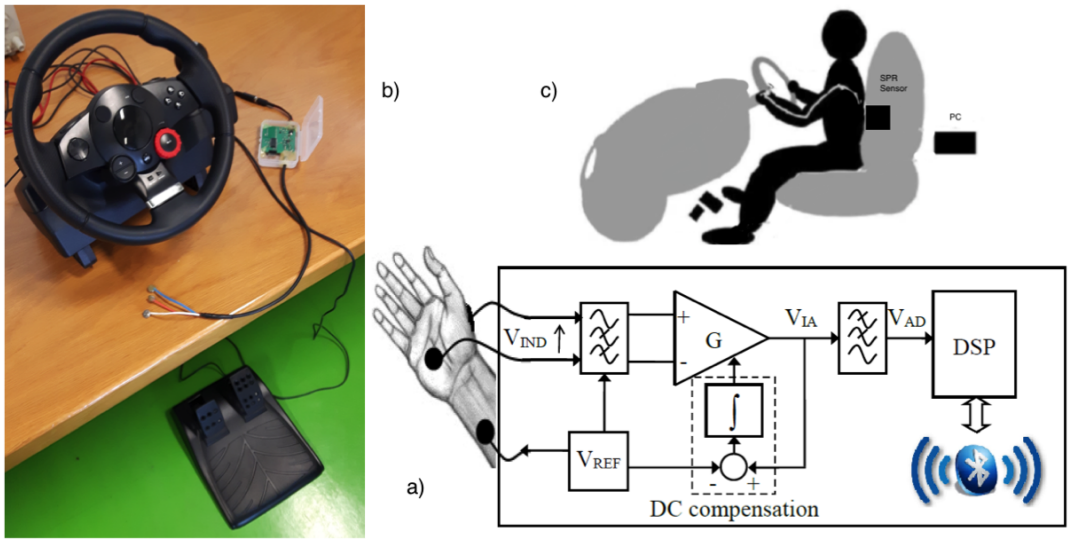


FIGURE 4.1: Laboratory experimental setup. EDA Sensor (a). Steering Wheel (b). An illustration resembling the experiment

SPR with a high input impedance instrument, in fact, makes the changes in skin electrical properties negligible (this is one of the advantages of SPR with respect to SCR).
The instrumentation amplifier then, amplifies V_{IND} with a gain set to 80, since the input range for SPR pulses is in the order of $\pm$ 20 mV and must be amplified into 3.3 V_{pp}.
The DC Compensation block in Figure 4.1 (a) has been introduced to remove the non-idealities of the instrumentation amplifier, like offset voltage (V_{OS}), bias current (I_B) and offset current (I_{OS}). The DC compensation is the inverting integration of $V_{IA} - V_{REF}$ thus shifting the DC component of the amplifier output to V_{REF}. This DC compensation, together with the input filters, behaves as a second order high pass filter with cutoff frequency at 0.08 Hz. At the output of the instrumentation amplifier, a third order Sallen-Key lowpass filter acts as anti-alias filter with cutoff frequency at 40 Hz.
After the analog front end, V_{AD} is acquired by the analog to digital converter on board on the DSP; the sample rate is 200 Sa/s and data are converted with 12-bit resolution. On the DSP a digital notch filter is implemented in order to reduce the EMI power line noise which affects the low-level signal. The implemented notch is a second order Butterworth filter centered at 50 Hz and

rejects power line disturbs by 38 dB. The processed data are finally sent to the Bluetooth module with data transfer Baud Rate of 19.2 kbps. The frequency response of the circuit allows detecting the SPR pulses, but the stress measurement is partially affected by skin stretch and body movements, which produce burst potentials. For this reason, it is mandatory to remove the motion artifact, as described in next section.

4.1.2 Steering Wheel

The subjects under test in laboratory use a Logitech G Driving Force GT Steering Wheel from which we record the Steering Wheel angle (Figure 4.1 (b)). The recorded values can range from -1, which means that the Steering Wheel angle is turned 90° in anticlockwise, to 1, which means that the Steering Wheel angle is turned 90° in clockwise. A value of 0 means that the steering wheel is at the reference position 0°.

4.1.3 PC connection and interface

The sensor and the steering wheel are connected to a PC (Figure 4.1 (c)) through a software developed in Matlab. The sensor transmits data via Bluetooth and the steering wheel angle is recorded in synchronous with the receiving of data sets from the sensor. In addition, the software manages the playing of audio and video files which are used to represent the test scenario and cause stress to the subject under test.

4.1.4 Endosomatic EDA and exosomatic EDA comparison

In Section 2.5 we showed patterns which represent endosomatic and exosomatic waveforms; we then described the properties of Skin Potential and Skin Conductance waveforms.

In this section we take real measurements of endosomatic EDA, (Skin Potential) and exosomatic EDA (Skin Conductance); then,

we compare SP and SC and make our considerations on the basis of real data. Our choice to use Skin Potential to detect the stress of a car driver then, is explained.

An example of different characteristics belonging to SCR and SPR signals in our application is shown in Figure 4.2, which plots a 70 seconds long recording of both SPR and SCR responses, in blue and red color, respectively. The two signals have been synchronized on the basis of the acquiring laptop timestamp, which has a resolution of 1 ms.
The test encompassed a free drive on a simple track, using the steering wheel with the right hand, which had no sensors applied, while the left hand was kept resting on the knee. It is from this left resting hand that the SPR signal was recorded, in addition to the SCR recorded with a commercial device produced by Shimmer [22]. At random times a stress inducing sound is played to the driver (evidenced by the black square markers in the figure).

From Figure 4.2 it can be seen that, as we have explained in Section 2.5, the SPR shows a faster reaction to stress stimuli compared to SCR. SPR rises and decays faster, thus making event recognition possibly simpler in our scenario. Once the experimental measurements confirmed the bibliographical hypothesis, we considered other aspects that led us to adopt Skin Potential for our work. SPR can be measured in a simpler way, without applying current to the body, and because it is less sensitive to the impedance of electrodes and to slow variations of skin impedance [3]. A detailed comparison of different EDA signals, including SCR and SPR, is also reported in [12].

4.1.5 Motion Artifact

The main problem concerning electrodermal activity measured while a person is driving, is represented by the Motion Artifact (MA). Since electrodermal measurements have to be taken from the hand, the signal will be strongly corrupted by components related to physical movements instead of subject mental state.

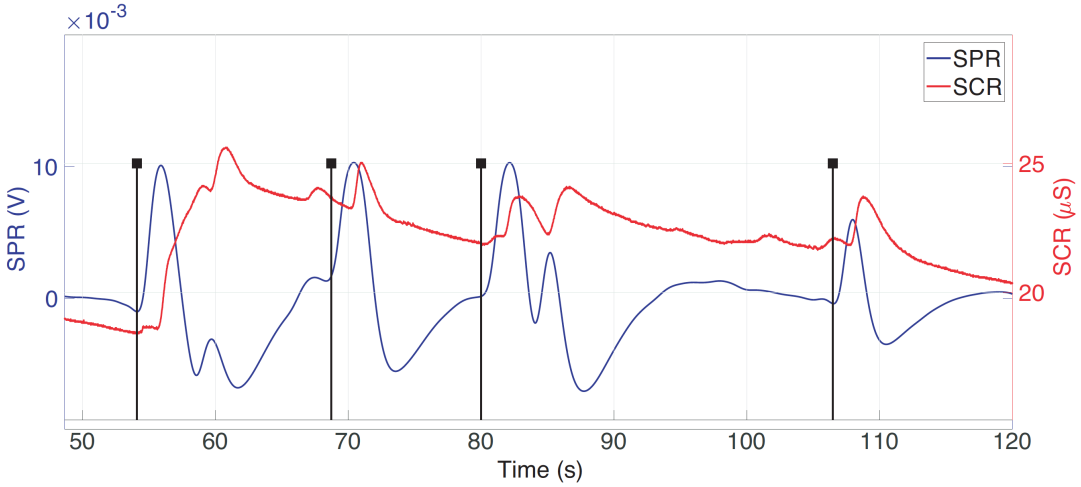


FIGURE 4.2: SPR and SCR responses comparison. Blue line: SPR signal. Red line: SCR signal.

Moreover, the power of motion artifact disturbances is comparable with the power of the waves that resemble stress; since a driver, almost always needs to move the steering wheel, it can lead to many false stress detections.

As underlined in [54], Motion Artifact related problems limit the situations in which EDA measurements can be used for stress detection in car driving scenarios and the applications are limited as well. Standard filters are not able to provide a solution due to the overlap between the frequency bands of stress and motion artifact components. Therefore, motion artifact cannot be processed and removed in the same manner as it can be done for noise or minor interferences that only partially overlap the frequency band of the signal of interest.

4.1.6 Measurements analysis

The initial test has the purpose of highlighting how the Motion Artifact obstructs the detection of stress response. SPR is measured on the subject's hand in three different conditions. In the first test, the subject does not move his hands. Audio record of clapping hands is played in random time instants to induce stress. In the second test, the subject moves his hand freely as he was driving. No stress is induced. In the third experiment,

stress is induced while the subject moves his hands as he was driving.

The obtained measurements are plotted in Figure 4.3 for test 1 (top), test 2 (middle) and test 3 (bottom). SPR is plotted in blue line, the Steering Wheel excursion in black dotted line and the instants in which the stress inputs are administered are marked with black stems. In test 1, we can clearly identify the subject's reaction to the stress stimuli. It is noticeable that SPR in the proximity of the stems resembles the patterns in Figure 2.8. In test 2, only the Motion Artifact component is present since no stress stimuli occur. We can notice that the signal power is comparable to the one of stress impulses in test 1. In test 3, it is evident the difficulty to discriminate the Motion Artifact and the stress response to stimuli. In these conditions, false stress detections will compromise the reliability of the analysis.

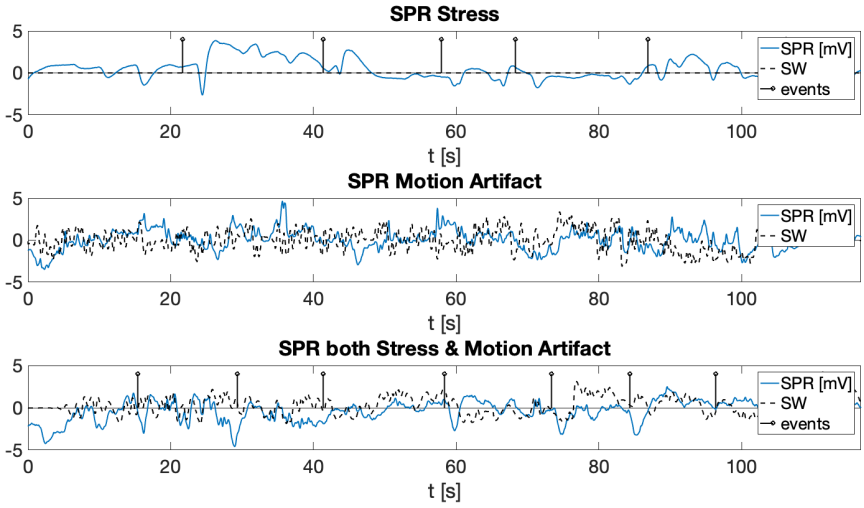


FIGURE 4.3: SPR preliminary measurements while: stress is induced with audio tracks (top); the subject is moving the SW (middle); stress is induced with audio tracks while the subject is moving the SW (bottom)

4.2 Adaptive filtering for Motion Artifact removal and SNEO peaks based Stress detection system

In this Section a system designed to remove the Motion Artifact and highlight the stress responses in SPR signal is presented. The results have been published in [5]. The system is depicted in Figure 4.4. It consists in the following main parts:
- Sensor
- Steering Wheel
- Motion Artifact removal
- Processing
- Detection and Display

all the system components will be described in the following sections.

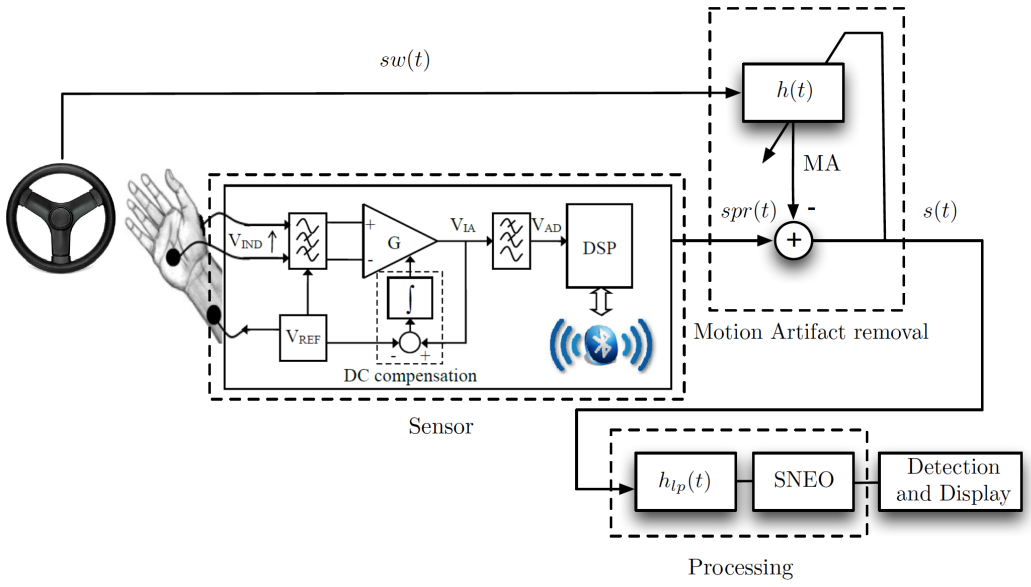


FIGURE 4.4: Block diagram of the proposed scheme.

4.2.1 Sensor and Steering Wheel

The sensor and the steering wheel are the ones we have employed and described respectively in Section 4.1.1 and Section 4.1.2 to which the reader can refer for the description.

4.2.2 Motion Artifact removal

The Motion Artifact removal block is composed by an adaptive filter, whose functioning is described below. The inputs of MA removal block are the SPR measurements and the Steering Wheel excursion.

Adaptive filters are used to predict the behavior of unknown and time varying systems. The adaptive filter algorithm minimizes the error between reference and output signals adapting its frequency response. The behavior of adaptive filters is strongly non-linear during the adaptation phase, but when the filter converges, its behavior resembles the behavior of a linear Finite Impulse Response (FIR) filter. An explanation of the variants employed in our work follows. For a detailed description the reader can refer to [53], [98].
In our application, the input of the adaptive filter $x(t)$ is represented by the recorded Steering Wheel angle $x(t) = sw(t)$. The reference signal $d(t)$ is represented by the measured Skin Potential Response $d(t) = spr(t)$. Assuming that Motion Artifact is correlated to the Steering Wheel excursion, we propose to model the acquired SPR signal as a linear regression:

$$spr(t) = sw * h(t) + s(t) \tag{4.1}$$

Least Mean Square Adaptive Filter

Least Mean Square (LMS) algorithm [53] adapts its impulse response $\boldsymbol{h}$ in order to minimize the expected Mean Square Error (MSE) between the filter output $\boldsymbol{x}_t * \boldsymbol{h}_t$ to the input $\boldsymbol{x}_t = \boldsymbol{sw}_t$, and the desired response $d(t) = spr(t)$. The input vector $\boldsymbol{x_t} = [x(t), \ldots, x(t-N+1)]^T$ is composed by the last N consecutive samples of the input signal; $\boldsymbol{h}$ is the time variant impulse response, whose value at time t is $\boldsymbol{h_t} = [h_{1,t}, \ldots\ldots, h_{N,t}]^T$. The input signal $x(t)$ then, undergoes to a time variant convolution. The coefficients of h_t are updated to minimize the MSE of the error $e(t) = d(t) - h_t$, trough the gradient descent method:

$$\boldsymbol{h_{t+1}} = \boldsymbol{h_t} - \frac{\mu}{2}\nabla E[e^2(t)] \tag{4.2}$$

In the implementation of the LMS algorithm, the filter update is calculated as $\boldsymbol{h_{t+1}} = \boldsymbol{h_t} + \mu e(t)\boldsymbol{x_t}$,where $\mu e(t)\boldsymbol{x_t}$ is an approximation of the gradient based filter coefficient correction. The filter length N and the step μ must be chosen in order to obtain the desired output and guarantee the stability. Then, μ cannot exceed the limit which depends on the eigenvalues of the input signal covariance matrix.

A typical choice is [53]:

$$\mu = \frac{\alpha}{\lambda_{max}} \tag{4.3}$$

where λ_{max} is the greatest eigenvalue and α is a constant whose valuee must be $0 < \alpha < 1$, in order to assure the filter convergence. In case of strongly non-stationary signals, the eigenvalues could spread, deteriorating the speed and the performance of the filter when they depends on the smaller eigenvalues. The trace of covariance matrix, instead of its eigenvalues, can be used for choosing μ. This approximation is not the optimum one but it assures stability and diminishes computational cost when applying the algorithm.

Recursive Least Squares Adaptive Filter

The Recursive Least Squares (RLS) algorithm updates the filter coefficients basing on the errors of the previous time instants [98]:

$$\epsilon_t = \sum_{i=0}^{t} \lambda^{t-i} e^2(i) \tag{4.4}$$

where $e(i)$ is the error at time i. λ is the forgetting factor that weights the proportion of past data which contribute to the filter update. The closer is λ to 1, the more the past samples contribute to the output at time t.

By minimizing 4.4 with respect to the filter coefficients, one obtains a system of equations with solution $\boldsymbol{h_t} = \boldsymbol{P}_t \boldsymbol{r}_{dx,t}$, where $\boldsymbol{P}_t$ is the inverse of the weighted input covariance matrix, and $\boldsymbol{r_{dx,t}}$

is the vector of weighted cross-correlation coefficients between $d(t)$ and the input. By using the matrix inversion lemma, it is possible to solve the system of equations recursively, by writing $\boldsymbol{P}_t$ and $\boldsymbol{r}_{dx,t}$ as functions of $\boldsymbol{P}_{t-1}$ and $\boldsymbol{r}_{dx,t-1}$.

The hyperparameters of RLS algorithm are the filter order N and the forgetting factor λ. λ must be $0 < \lambda < 1$ to avoid the input matrix to became singular, lead the system to instability and make the problem ill-conditioned. One also initializes $\boldsymbol{P}_0$ as $\boldsymbol{P}_0 = \delta^{-1}\boldsymbol{I}$, where the influence of δ becomes negligible at steady state if $\lambda < 1$. The RLS algorithm can keep track of the signal in a better way, but requires more computational resources than LMS. Moreover, it has been demonstrated that, for non-stationary signals, RLS filtering can give worse performance than the LMS algorithm [17].

Kalman Adaptive Filter

Kalman filtering is a general technique used to estimate the unknown state of a linear system from noisy observations and a model of state uncertainties. We can cast the problem of adaptive filtering as the state estimation problem for the system:

$$\begin{cases} \boldsymbol{h_t} = \boldsymbol{h_{t-1}} + \boldsymbol{q_t} \\ d(t) = \boldsymbol{h_t}^T \boldsymbol{x_t} + v(t) \end{cases} \tag{4.5}$$

,where $\boldsymbol{q}_t$ is the process noise and $v(t)$ is the observation noise.

The algorithm consists in two steps. In the prediction step, the Kalman filter estimates the state variables at the current time, as well as their uncertainty by means of the estimate covariance matrix. Once a new measure is acquired, the state estimate, represented by the covariance matrix a priori estimate P, is updated, with more weight being given to estimates with less uncertainty [53]. The algorithm is recursive and utilizes the current measurement and previously calculated state estimates.

To set Kalman filter properly, correct estimations of the hyperparameters, represented by the variance σ_v^2 of the observation noise, and the covariance matrix $\boldsymbol{Q}_t$ of the process noise, are

necessary. Moreover, one initializes the a priori estimate covariance matrix as $\boldsymbol{P}_{0|0}$. The algorithm can become unstable or ill-conditioned if the system described by Equation 4.5 becomes non-observable [53].

4.2.3 Processing

The output $s(t) = spr(t) - sw * h(t)$ of the MA removal block, is low-pass filtered in order to remove high frequency interference, which is not relevant for emotional and stress event detection. As it will also be apparent from the experiments in Section 4.2.8, emotional events typically have a few seconds duration. Therefore, the cutoff frequency of the low-pass filter is set to 0.5 Hz. Starting from signals sampled at 200 Hz, the filter is implemented as a cascade of multistage decimation linear phase filters.
The output of the filter is then processed via the Smoothed Nonlinear Energy Operator (SNEO), in order to enhance possible stress events [78, 23]. In particular, SNEO is a smoothed version of the nonlinear energy operator:

$$\psi[s(t)] = s^2(t) - s(t-1)s(t+1) \tag{4.6}$$

4.2.4 Detection and Display

By modelling $s(t)$ as a combination of emotional events spikes and noise, it is shown in [78] that $E[\psi[s(t)]]$ can reveal spike locations, since they are associated to sudden raising energy. The expectation operator for SNEO is approximated by convolving it with the five-point Bartlett window (2, 4, 6, 4, 2).

In the current implementation of our system, we process the signal offline and, for the purpose of visualization, we simplify and display the emotional information by subdividing the absolute value of the SNEO signal into blocks with duration of about 10 s. For each of the blocks, we select one event by approximating the signal with a single Gaussian shaped impulse. This can be

easily done by projecting the signal block into a dictionary of Gaussian-shaped atoms, with different variances and time position inside the block that represents the time interval and selecting the atom corresponding to the largest projection coefficient [34]. The stress is detected finding the peaks of the Gaussian SNEO reconstruction, which are greater than a fixed threshold.

4.2.5 Experimental setup

In order to assess the influence of MA and compare the different adaptive filter types, we preliminarily consider a test where a subject watches a video with different driving situations (curves, cone slaloms, normal driving), and moves the steering wheel accordingly. Figure 4.5 shows a picture of the lab with the subject taking the test. The steering wheel is connected to a computer which records the steering wheel rotation angle $sw(t)$, while the SPR signal is measured by the sensor and transmitted via a Bluetooth interface, as introduced in Section 4.1.1. We do not induce stress events in this case.

FIGURE 4.5: A picture of the test setup.

We then perform tests where we actually induce acute stress events, in order to unambiguously define a ground-truth for the

proposed stress recognition system. Tests were devised with the aim of evaluating the stress induced on subjects during a simulated driving experience in a controlled experimental setup. The kind of stimuli we consider here may not fully represent "driving stress" which may arise from different causes, like prolonged fatigue. We believe however that the proposed setup can be representative of the kind of responses that we can detect in the SPR signal.

We tested 15 subjects, 3 females and 12 males, in healthy conditions and with average driving experience, drawn from a pool of students from the University of Udine, Italy, with ages in the range 20–28 years. Stress is induced at random times during the experiment through sudden sounds and/or video interactions. This allows identifying unambiguously the events of interest. Notice that this setup, which targets stress possibly caused by unexpected events, could be a good approximation of real situations, as well as of emotional peaks due to sudden awakenings when the driver is in danger to fall asleep at the wheel.

The tests are carried out in two different scenarios described as follows:

Scenario 1: audio sequence with sudden random sound effects. In the first scenario the subjects are asked to move the steering wheel, normally using two hands and with the sensor in one hand, while listening to an audio sequence which contains both sudden random sound effects (used as triggers to induce stress) and a metronome ticking sound. They must turn the steering half a turn left or right for each tick. The frequency of the metronome is changing through the experiment (from 1 to 2.5 Hz). We used multiple metronome frequencies instead of a single frequency because the motion artifact during driving conditions may have several spectral components depending on the road characteristics (different curve radius, sudden manoeuvre, etc.). Using different frequencies, we verify that the adaptive filter may respond effectively to different spectral components. The audio sequence is roughly 17 min long. The purpose of the experiment is to evaluate the influence of MA on the recorded SPR signal, and to measure the ability of the proposed scheme

to detect induced stress in correspondence of the sound effects. In order to assess the influence of possible residual artifacts in the processed signals, we also consider an experiment with one additional male student subject (Subject 0, outside the group of 15) where two sensors are used, one for each hand.

The subject is asked to use only one hand to steer the wheel while the other hand is kept resting during the test, differently from the other subjects, in the same test, who steer the wheel with both hands. Variations in the recorded signals due to induced stress events should appear in both signals, while possible artifacts should appear in the signal recorded from the moving hand.

Scenario 2: audio and video sequence with a car accident event. In the second scenario the 15 subjects are watching a video of a pilot driving on a racing track at normal pace, as seen from the driver's helmet. Subjects must steer the wheel in order to follow the track curves, as in a realistic driving session, so the recorded SPR signal is possibly affected by MA. They are listening to peaceful music in the background. We then induce a single stress episode by suddenly reproducing a car accident video frame and turning up the audio. After several seconds, the video and music start again from where they were interrupted, so we can keep recording the SPR after the "accident". The audio/video sequence is roughly 4 min long.

For each of the 15 subjects in both scenarios, we can count the True Positives (TP), i.e., the number of correctly detected stress peaks after the trigger, and the False Negatives (FN), i.e., the number of missed triggers. We believe it is not significant to count the False Positives (FP) or True Negatives (TN), because we cannot verify when stress should not be present in a subject, because a person can be stressed for reasons that we don't know and that we cannot measure. Moreover, it is well known that EDA signals can exhibit nonspecific responses which cannot be traced to any stimulation [21].

We then compute the Recall figure (also known as Sensitivity), which is defined as:

$$Recall(\%) = \frac{TP}{TP + FN} * 100 \tag{4.7}$$

where $TP + FN$ amounts to the total number of stress triggers. [10]LMS

4.2.6 Adaptive Filters comparison

To test Adaptive Filters and tune their hyperparameters, we set up a simple experiment where the input signal $x(t)$ is white gaussian with unit variance. The measurements $d(t)$ are computed as $d(t) = h^o * x(t) + n(t)$, where $h^o(t)$ is a 100-tap FIR filter whose coefficients are chosen at random, independently and with a uniform $[0, 1]$ distribution. The filter energy is then normalized to unity. The noise $n(t)$ is white gaussian with variance 0.1. The filter parameters were chosen as follows. For the LMS filter, we chose $\mu = 0.002$. The RLS hyperparameters were chosen as $\lambda = 1 - 10^{-4}$ and $\delta = 10$ for the RLS filter, while we set $\boldsymbol{Q}_t = 10^{-7}\boldsymbol{I}, \sigma_v^2 = 0.1$ and $\boldsymbol{P}_{0|0} = 0.1\boldsymbol{I}$ for the Kalman filter. The adaptive filter length is set to $N = 100$ and the initial filter coefficients $\boldsymbol{h}_0$ are set to zero. Figure 4.6 shows the learning curve $|\boldsymbol{h}_t - \boldsymbol{h}^0|$, in dB, for the different adaptive filters in one trial.

As expected, the Kalman and RLS filters exhibit the fastest response, with a similar performance. The LMS filter, with the given μ value, exhibits a slightly slower response and a slightly larger error. The time-averaged error energy $E[e^2(t)] = E[(d(t) - \boldsymbol{h}_t^T \boldsymbol{x}_t)^2]$, in 50 trials with different random $h_o(t)$ filters and signal samples, was about 0.11, 0.11, 0.13 for the Kalman filter, the RLS filter and the LMS filter, respectively.

As a further test, we performed experiments with real data, using the same filter parameters as before, by setting $x(t) = sw(t)$, the steering wheel angle recording, and $d(t) = spr(t)$, the acquired SPR signal. According to model 4.1, the residual error signal $s(t)$ represents SPR signal components that are not due to the motion artifact, and that may evidence stress events.

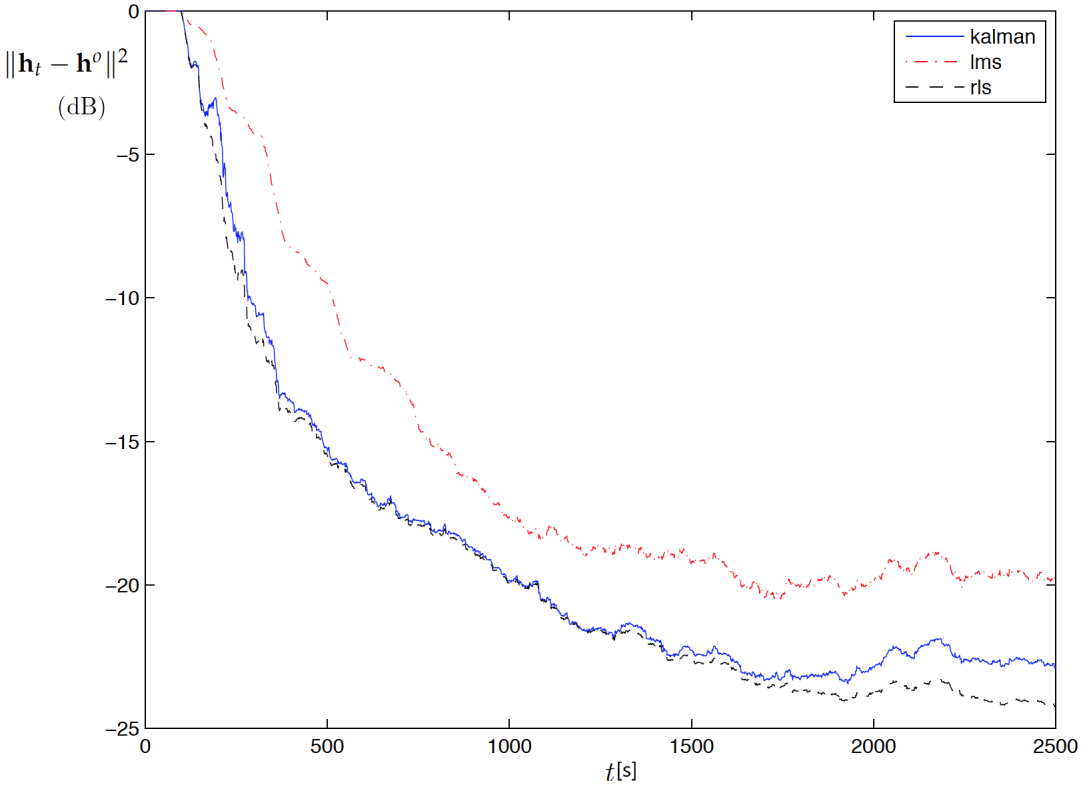


FIGURE 4.6: Learning curve for the adaptive filters.

Figure 4.7(a) and Figure 4.7(b) show $sw(t)$ and $spr(t)$, respectively.

The SPR signal is in general correlated with the steering wheel (SW) signal, and the MA removal procedure aims at removing the SW component from the SPR. As an example, Figure 4.8 shows the Pearson correlation between consecutive 5s blocks of the SPR and SW signals. The correlation for the event located at the sample interval between 13.000 and 15.000, where the SW is almost constant, has a smaller value $\rho = -0.22$.

As described in Section 4.2.5, the driver in this preliminary test moves the steering wheel following a test video, replicating typical driving behavior in different situations. The plots clearly show the steering wheel pattern in $spr(t)$, confirming the influence of motion artifacts. However, the regions evidenced by the red squared blocks show a rather clear deviation of $spr(t)$ signal from $sw(t)$, which may reveal emotional or stress events. Figure 4.7(a), 4.7(c) and 4.7(d) show the squared residual $s_2(t)$ after the Kalman and LMS filters, respectively. Although with

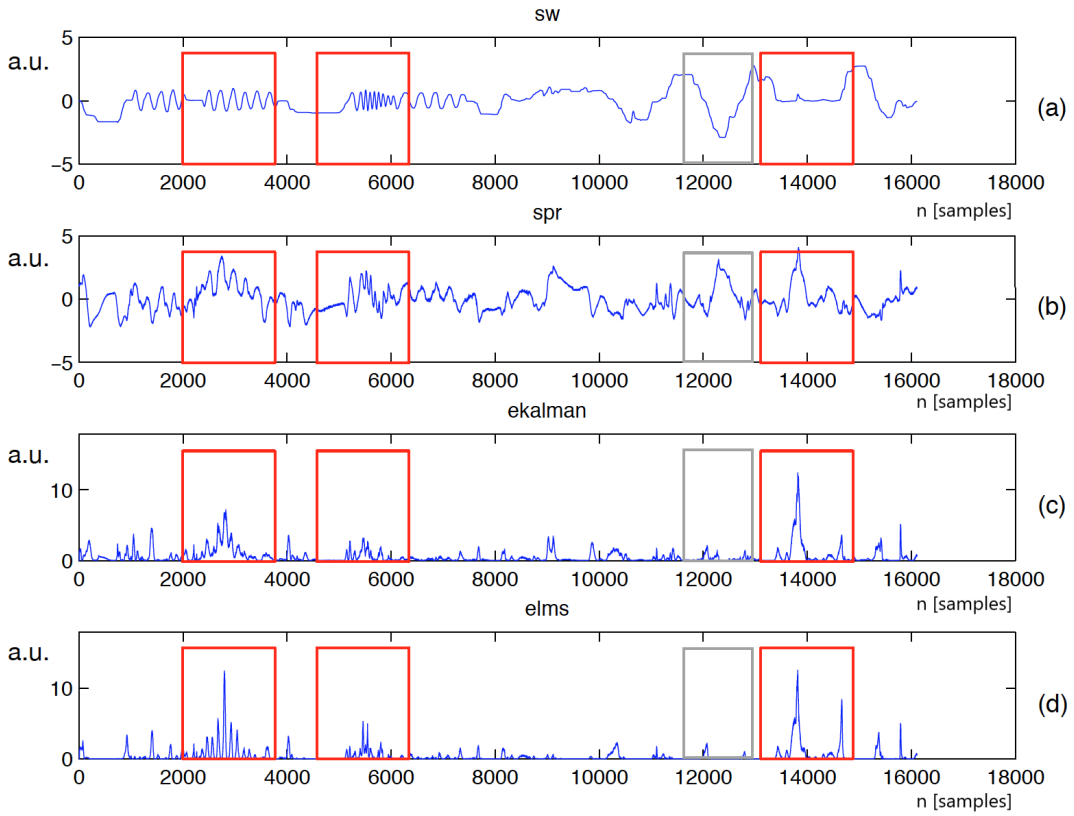


FIGURE 4.7: Adaptive filter performance with acquired signals. (a) Steering wheel angle sw(t), (b) Skin Potential Response spr (t), (c) squared residual $s^2(t)$ for the Kalman filter, (d) squared residual $s^2(t)$ for the LMS filter.

some differences, the qualitative behaviour of the filters is comparable. The time-averaged error energy $E[s^2(t)]$ in this example was 0.52, 0.86, 0.39 for the Kalman filter, the RLS filter and the LMS filter, respectively. Note that the signals in Figure 4.7(a) and 4.7(b) differ also at time instant 12,000, as evidenced by the gray rectangle.

However, the two signals are pretty much inverse correlated in this time interval, so that it is difficult to decide if the peak in $spr(t)$ signal is MA or if it is related to an emotional event. In any case, the adaptive filters compensate the peak, and the corresponding residuals have low energy. Note that the input signal $sw(t)$ can be nearly constant and narrowband in many time intervals. This may cause the input covariance matrix to become close to singular, and the system 4.5 to become nearly non-observable. When processing real data, we noticed that the

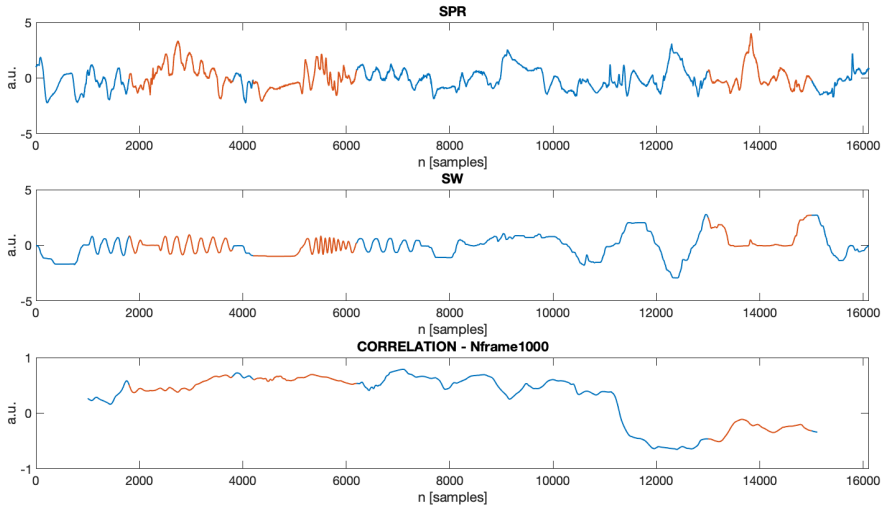


FIGURE 4.8: Top: Skin Potential Response; Middle: Steering Wheel; Bottom: correlation between SPR and SW over time, the moving window length is 1000 samples (corresponding to 5 s).

Kalman and RLS estimated filter coefficients can fluctuate significantly, even if this has little impact on the residual energy.

As a matter of fact, with a strictly narrow-band input signal, the adaptive filter frequency response in no-care regions can be arbitrary, without affecting the residual. For these reasons, and for its lower computational complexity, we chose the LMS adaptive filter for the experiments presented below.

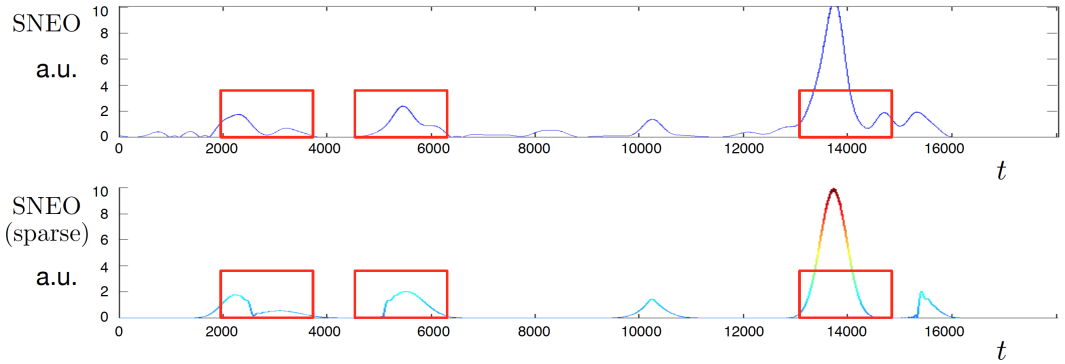


FIGURE 4.9: Absolute value of SNEO and its sparse representation.

The sparse approximation is however not necessary, and, in view of a real time implementation, we can take into account the original SNEO signal only.

4.2.7 Experimental setup

Scenario 1: Figure 4.10 shows the SPR acquired from Subject 0 in Scenario 1, for whom we also record the SPR signal from the still hand, and the corresponding steering wheel angle (Subplot 1, 250 s signal block). Sound occurrences are marked by the red vertical lines with a circle marker in the figure. We also plot lighter grey lines with a square marker to indicate the metronome frequency change. As it can be seen, the SPR has big peaks right after stress events (circle markers), and between them there are smaller variations which show a correlation to the wheel rotations. The stress signal, obtained after MA removal and lowpass filtering, is shown in Figure 4.10 Subplot 2 and compared with the SPR signal acquired from the still hand. It can be noticed from Subplot 1 that the raw SPR signal is correlated to the SW signal and from Subplot 2 that the obtained stress signal is consistent with the still hand signal, confirming the effectiveness of our procedure and its feasibility to reveal stress-related events.

As described in the previous sections, when the SPR signal is similar to the filtered SW signal, it can be inferred that a motion artifact due to the movements of the steering wheel and vibrations has occurred, and the stress component is negligible. On the other hand, if the signals are fairly different, this difference is due to the emotional or stress component. This is confirmed by looking at the figure, where it can be seen that the higher stress peaks happen right after each sudden sound effect. In order to make more evident the effect of MA removal, we show in Figure 4.11 a detail of the signals in the time interval 347–372 s. In particular, we show the raw SPR signal, the SPR signal after MA removal and low-pass filtering and the still hand signal.

We also show the raw SPR signal after low-pass filtering, but without MA removal, in order to evidence the effect of the adaptive filter. It can be seen that SPR signal, processed with the proposed procedure, is indeed consistent with the still hand signal, and that the raw SPR signal is indeed influenced by the motion artifact.

The stress signal is then processed to derive the SNEO and its sparse Gaussian reconstruction (as explained in Section 4.2.5). Figure 4.12 Subplot 1 shows the absolute value of SNEO of Subject 0 for the entire duration of the experiment. The stress identified by the peaks in the signal is now more evident and clearly appears after each sound event (circle markers). Smaller peaks also happen when the metronome frequency changes. In addition (see the sparse approximation in Subplot 2), for visualization purposes, we propose a possible way to segment stress level intensity using color thresholds, by dividing the maximum measured level (from blue for lowest/no stress, to red for highest stress).

Results for the 15 students in the group, using two hands for moving the steering wheel, are consistent with those of Subject 0, even if individual responses can have different characteristics. As an example, Figure 4.13 and Figure 4.14 show the signals and SNEO respectively for Subject 3 of the group, where the response under the same conditions is different. This is due to a variety of reasons. The measured SPR changes from person to person because of the differences in the skin properties, moreover each person might have a different emotional reaction to a particular sound. These variations can be seen in Figure 4.15 and Figure 4.16, relative to Subject 6, where there are lower or absent stress reaction peaks after some sounds, or more pronounced peaks during metronome frequency changes.

Scenario 2: Figure 4.17 shows the absolute value of SNEO and its Gaussian approximation for Subject 1 for the entire duration of the experiment. As before, the LMS algorithm is used to account for the MA signal. The stress response after the sudden simulated accident can be clearly seen. In the vast majority of subjects the stress peak right after the accident is higher than any other stress level during the rest of the video. In some subjects there have been similar stress peaks during other parts of the video (see also another subject reaction, Subject 8, in Figure 4.18). This might be due to the fact that some subjects may have some emotional reactions by just watching the video, or they expect a second accident after the first one in successive turns.

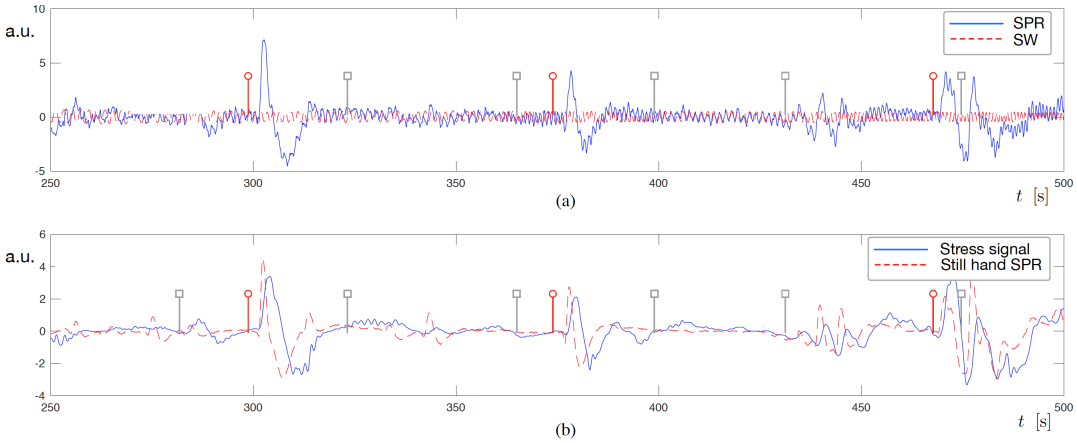


FIGURE 4.10: Subject 0 in scenario 1. Subplot 1: Skin Potential Response signal and Steering Wheel excursion; Subplot 2: Resulting stress signal after MA removal and lowpass filtering, compared with the still hand SPR signal. Red markers: induced stress events, Gray markers: metronome frequency changes.

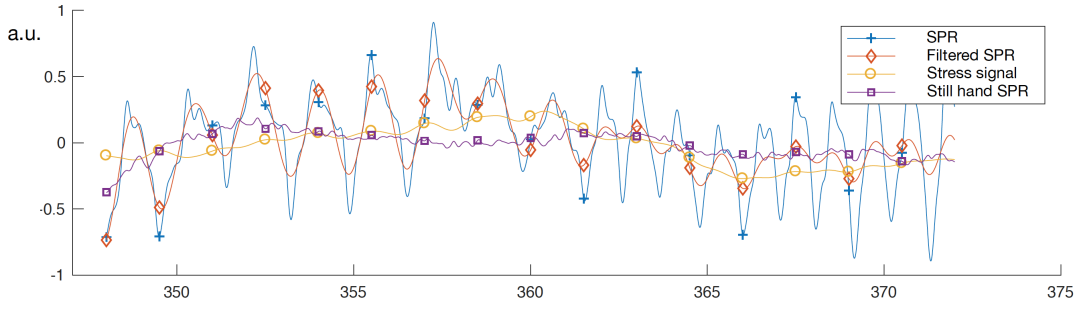


FIGURE 4.11: A detail of the signals in time interval 347–372 s.

4.2.8 Experimental results

In order to quantify the performance of our algorithm, we start from the hypothesis that each stress trigger (the sudden random sounds and the accident event) will cause a stress reaction in the subjects. We will consider a stress reaction as happening when the stress level will be greater than a threshold. This threshold is defined for each subject individually and, as a reasonable empirical choice, we set it to 75% percentile of the SNEO values obtained in a 5 min recording of the SPR signal in conditions similar to those of the test, before the actual experiment begins.

In a real world implementation, one would need to train the system for each particular subject. To collect the statistics of our

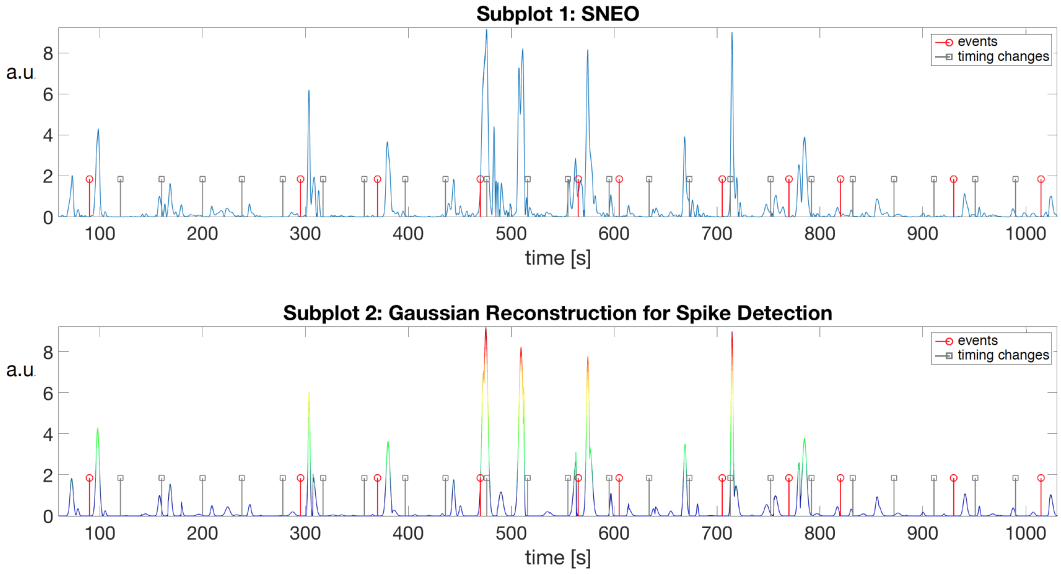


FIGURE 4.12: Subject 0 in scenario 1. Subplot 1: Absolute value of SNEO; Subplot 2: Sparse Gaussian reconstruction.

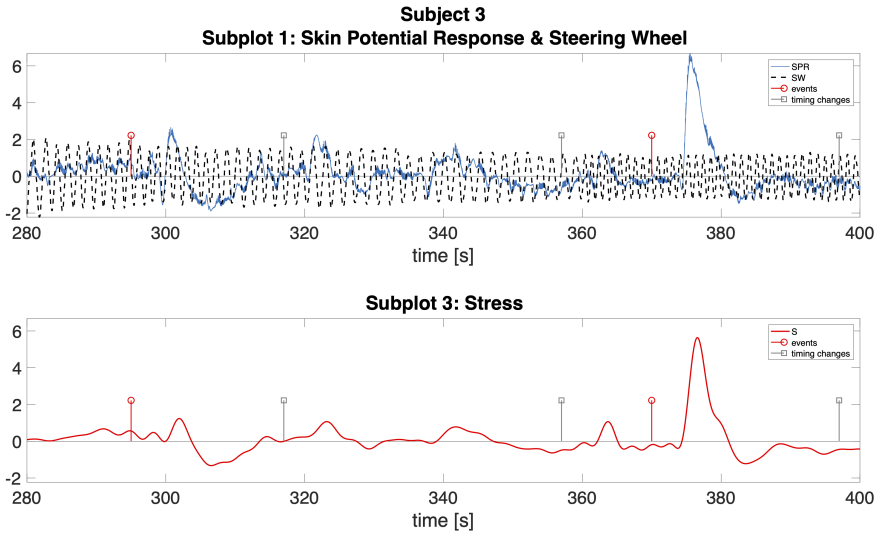


FIGURE 4.13: Subject 3 in scenario 1. Subplot 1: Skin Potential Response signal and Steering Wheel excursion; Subplot 2: Resulting stress signal after MA removal and lowpass filtering. Red markers: induced stress events, Gray markers: metronome frequency changes.

experiments, any value below the threshold will be ignored, and any occurrence above it will be counted as positive. To count the true positives, we consider only the stress happening in a ten seconds time window after each trigger event. If stress happens outside of this window we don't consider it as related to the

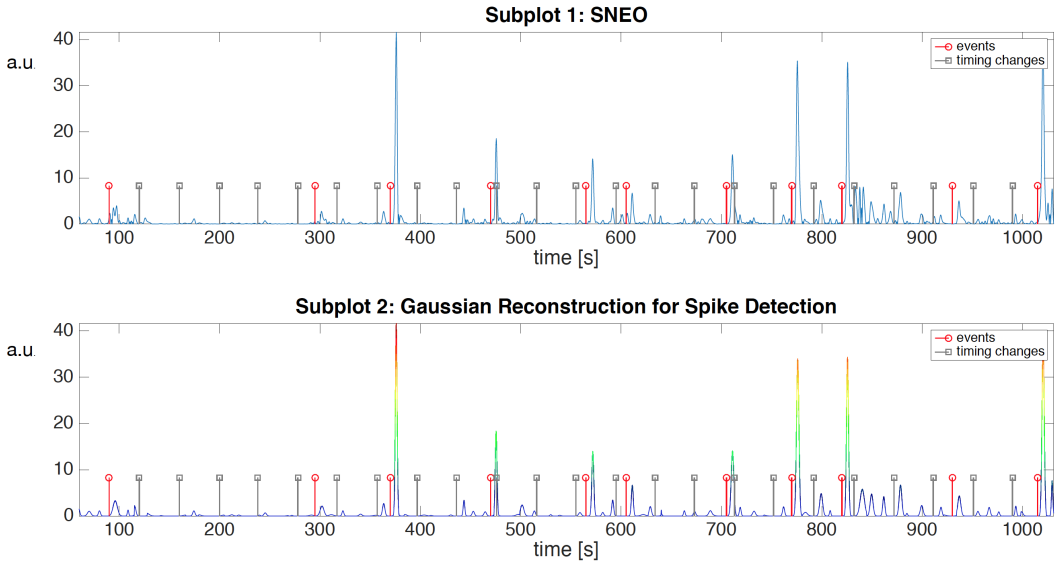


FIGURE 4.14: Subject 3 in scenario 1. Subplot 1: Absolute value of SNEO; Subplot 2: Sparse Gaussian reconstruction.

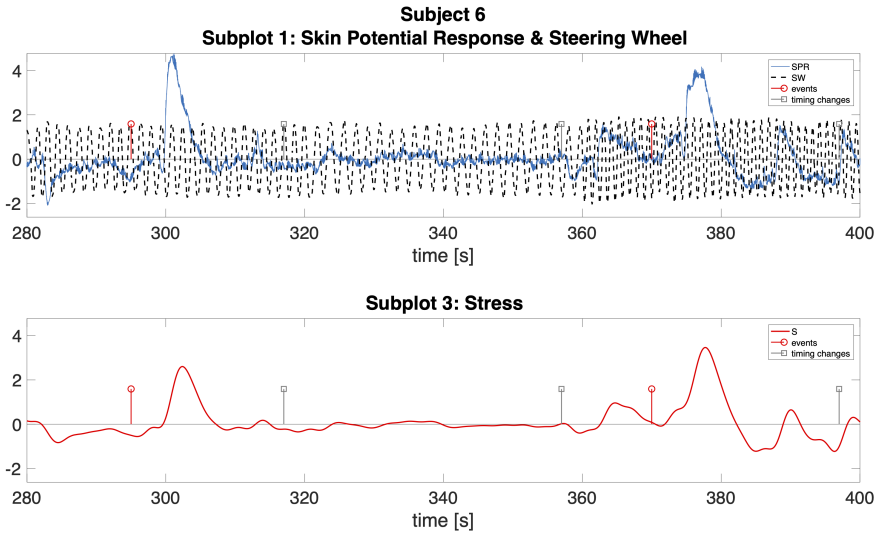


FIGURE 4.15: Subject 6 in scenario 1. Subplot 1: Skin Potential Response signal and Steering Wheel excursion; Subplot 2: Resulting stress signal after MA removal and lowpass filtering. Red markers: induced stress events, Gray markers: metronome frequency changes.

trigger event.

In the video experiment, the single stress event induced a stress reaction in all of the 15 test subjects, evidenced by a large peak of the SNEO signal in the 10 s window after the event, clearly

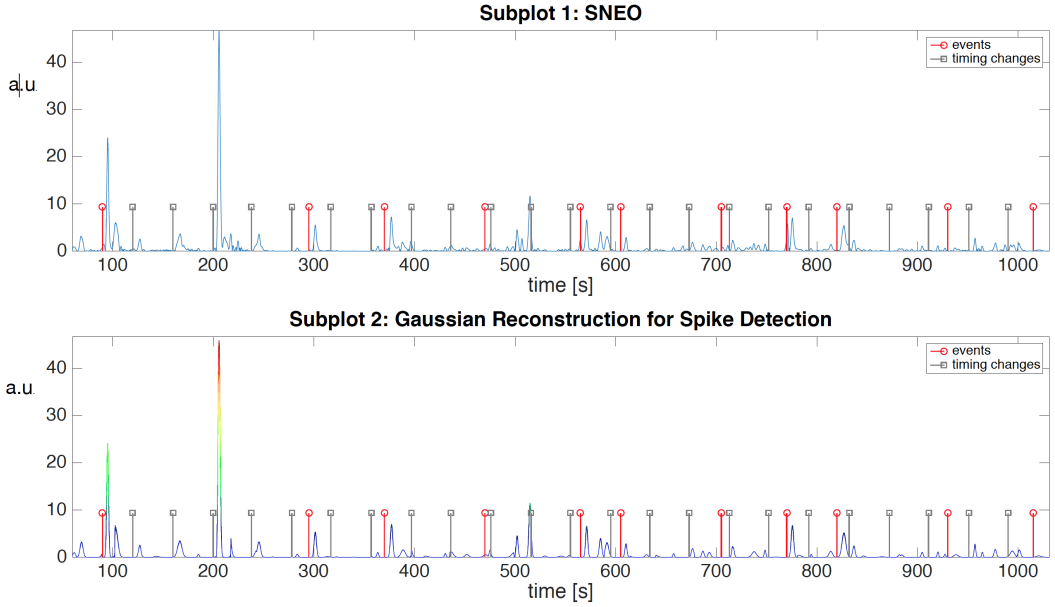


FIGURE 4.16: Subject 6 in scenario 1. Subplot 1: Absolute value of SNEO; Subplot 2: Sparse Gaussian reconstruction.

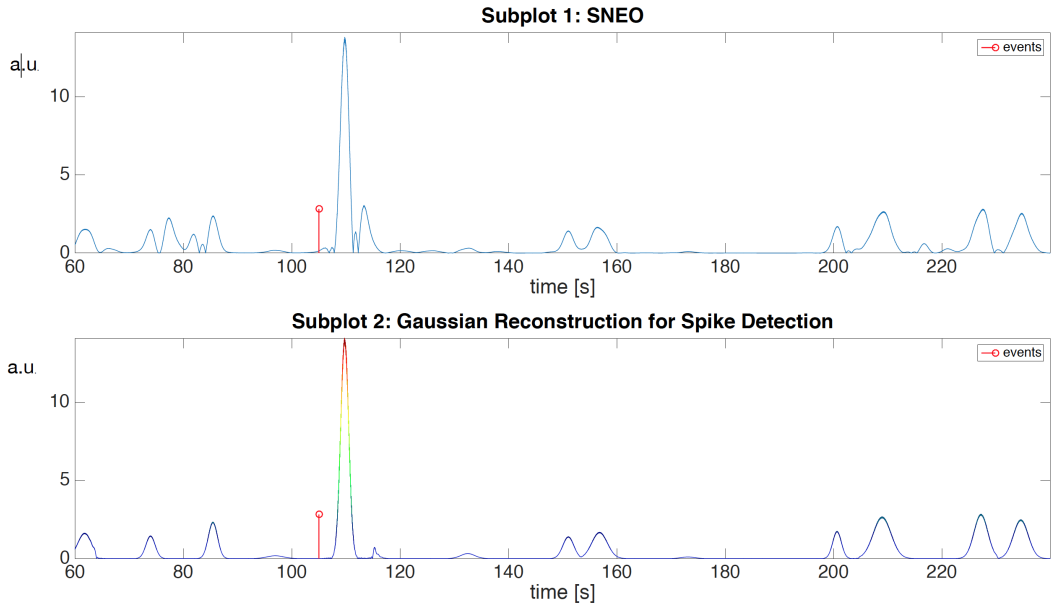


FIGURE 4.17: Subject 1 in scenario 2. Subplot 1: Absolute value of SNEO; Subplot 2: Sparse Gaussian reconstruction.

distinguishable from other smaller peaks (see Figure 4.17).

In the following, we therefore quantify the results for the audio events scenario only. The results are shown in Table 4.1, where we compute the Recall for each subject. We then average these Recall values in order to express the general performance

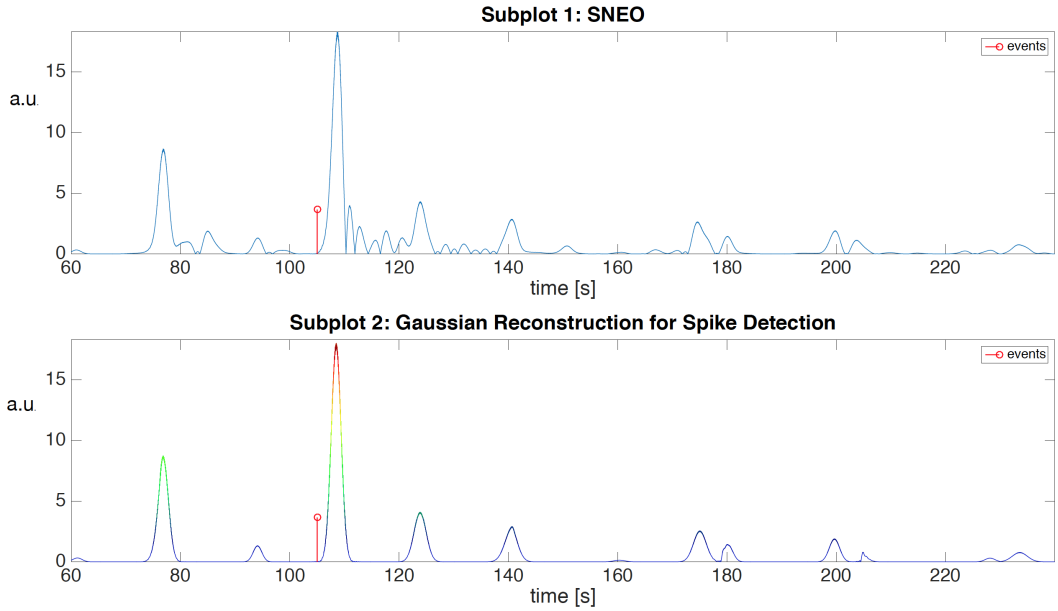


FIGURE 4.18: Another subject (Subject 8) in scenario 2. Subplot 1: Absolute value of SNEO; Subplot 2: Sparse Gaussian reconstruction. Red marker: induced stress event.

of the algorithm. The computed average turns out to be 95.15%, which confirms the good performance of the proposed scheme for stress event detection.

It is apparent from the experiments (see for instance Figure 4.16), that we can observe a certain number of peaks in the processed SNEO signal which are not related to the stress triggers. As mentioned, this can be due to changes of the emotional status of the subject that we cannot control. We observe however that many of these peaks are located in correspondence to changes in the metronome/steering wheel turning frequency, which may suggest they are related to artifacts still present in the processed signal. To better understand the presence of stress peaks which are far from triggers, as described in Section 3.2, we recorded on the first scenario with the audio sequence two signals for Subject 0, one from a resting hand and the other from the hand steering the wheel. In these conditions, we notice that SPR peaks, far from the trigger events, are still present in both recorded signals. This means that they are indeed probably due to the emotional or stress state of the subject and not, for instance, to residual

motion artifacts related to driving. Since they often happen after each metronome frequency change, it can be meaningful to analyze the algorithm performance in our original experiment (with the same 15 subjects measured data) in a different way, using a new hypothesis: not only each stress trigger, but also each metronome frequency change, will cause a stress reaction in the subjects.

We still count a stress reaction as happening when the stress level will be greater than the same thresholds we found before. For each subject we count the True Positives (TP), which this time are the number of correctly detected stress peaks after the sudden sounds and after the metronome frequency changes (within a ten seconds time window). We also count the False Negatives (FN), i.e., the number of missed peaks under the same setup. The results are shown in Table 4.2, where we compute the Recall for each subject. As before, we then average these Recall values, obtaining a result of 83.92%. This value is a little worse than before. This might be due to different reasons: the metronome frequency change is probably a less stressful event than the sudden sounds, the subjects may become used to the frequency changes, or slow to high frequency changes might be more stressful than high to slow ones. Nonetheless, the performance of the algorithm is still good, even under these new conditions.

For the sake of completeness, Table 4.3 shows the performance considering the metronome frequency changes only.

TABLE 4.1: TP, FN and Recall values for all the subjects (considering stress reactions after sudden random sounds).

Subject no.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
TP	11	10	11	11	10	10	10	11	11	11	10	10	11	10	10
FN	0	1	0	0	1	1	1	0	0	0	1	1	0	1	1
Recall (%)	100	90.91	100	100	90.91	90.91	90.91	100	100	100	90.91	90.91	100	90.91	90.91

TABLE 4.2: TP, FN and Recall values for all the subjects (considering stress reactions after sudden random sounds and metronome frequency changes).

Subject no.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
TP	33	28	28	28	26	29	27	26	30	32	27	28	28	28	30
FN	1	6	6	6	8	5	7	8	4	2	7	6	6	6	4
Recall (%)	97.06	82.35	82.35	82.35	76.47	85.29	79.41	76.47	88.24	94.12	79.41	82.35	82.35	82.35	88.24

TABLE 4.3: TP, FN and Recall values for all the subjects (considering stress reactions after metronome frequency changes).

Subject no.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
TP	22	18	17	17	16	19	17	15	19	21	17	18	17	18	20
FN	1	5	6	6	7	4	6	8	4	2	6	5	6	5	3
Recall (%)	95.65	78.26	73.91	73.91	69.57	82.61	73.91	65.22	82.61	91.30	73.91	78.26	73.91	78.26	86.96

4.3 Dual channel Electrodermal activity sensor and Motion Artifact removal through smoother selection algorithm

Although adaptive filters can follow the SPR deviations caused by the handling of the steering wheel, it has several limitations. While driving, people occasionally put the hands off the steering wheel. The Steering Wheel excursion then, cannot be correlated to the Motion Artifact when the driver is not holding it. This can lead to false stress detection. Another issue is represented by the artifacts caused by muscle stretching, which are not correlated to hand movements and can be wrongly detected as stress as well. The method developed to overcome the issue, considers SPR measurements taken on both hands. Relying on the fact that the emotional component is the same in both left and right SPR signals while the MA is not, the two SPR are processed through an algorithm whose output represents the common emotional component of SPR.

4.3.1 Sensor

The sensor architecture is shown in Figure 4.19(a). The dual channel SPR device acquires the electrodermal signals (named s_1 and s_2) on both hands and transmits the data via Wi-Fi to a laptop at a sample rate 200 Sa/s. The signals are acquired posing three Ag/AgCl electrodes on the palm, the back and the wrist of each hand. The sensor is battery operated with a single LiPo cell and is able to transmit data continuously for 8 hours. The sensor analog front end is shown in Figure 4.19(b). It is composed of two channels which amplify (with a gain G = 160) both s_1 and s_2 SPR signals whose amplitude is in the range $\pm$10 mV; in this way the output varies in the range $\pm 1.65V$ with respect to V_{REF}.

The bandwidth of the conditioning circuit is in the range [0.08, 40] Hz [2, 4, 3]. The analog to digital converter (12 bits resolution) on board of the DSP converts the data and sends them to the WiFi module. The analog front end of the sensor has been characterized in detail in [3], providing an accuracy in the measurement of the SPR signal in the order of 0.5%.

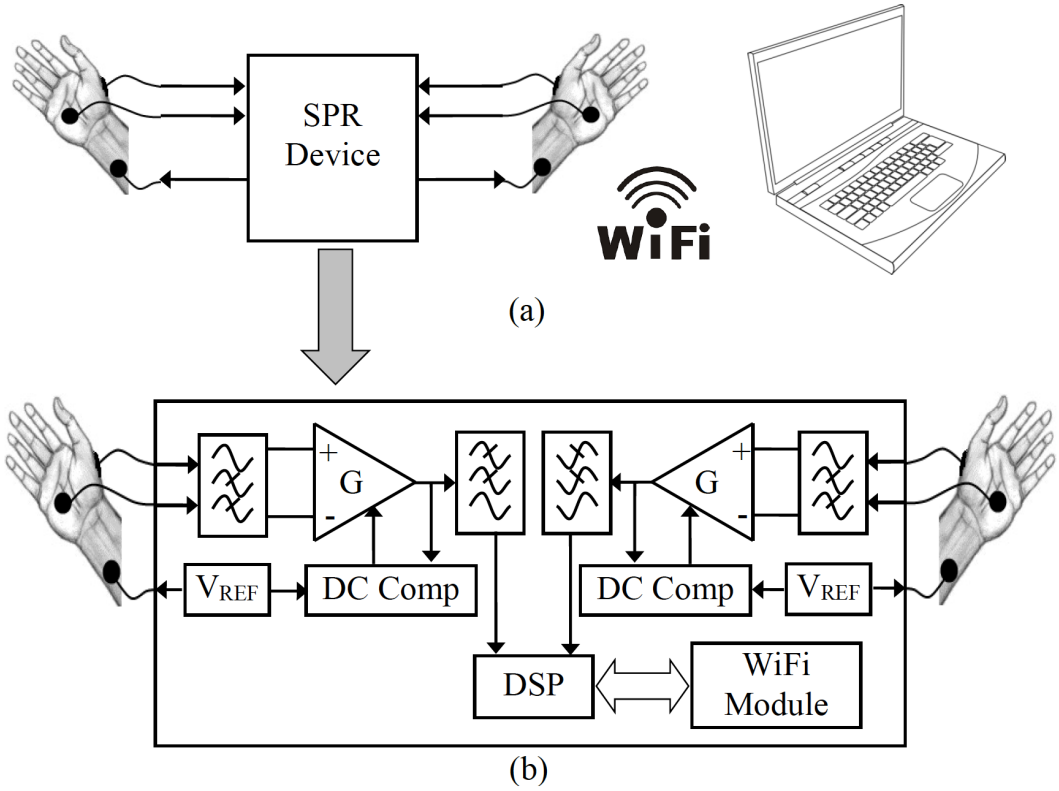


FIGURE 4.19: (a) Scheme of the developed system for EDA measure- ment; (b) block diagram of the dual channel SPR sensor.

4.3.2 Left and right hand SPR relation

In this section we show some measurements of SPR taken on the left and right hand with the sensor described in Section 4.3.1. These measurements put in evidence the crucial aspects between the two SPR signals on which the proposed method for MA removal has been based.

The two acquired SPR signals (named $s_1(t)$ and $s_2(t)$ as introduced before) are both governed by the sympathetic nervous

system, and thus they are almost identical if the hands are perfectly still. In Figure 4.20 it is possible to see, during a three minutes long recording, that the SPR signals from the two hands exhibit the same behavior in terms of peak amplitudes and variation rate.

As described in Section 4.1.5 and in Section 4.3, if SPR is measured while a person is driving, it will be affected by MA caused by hand movements and clasps. In Figure 4.21 we can see that because of the hands' motion, especially when the electrodes are mechanically stressed, the signals are strongly perturbed. Note that mechanical stress is typically induced by one of the two hands at a time, resulting in motion artifacts in one of the two signals only. For these reasons, a single channel EDA sensor could provide results that are not consistent with the real electrodermal activity due to stress or fatigue.

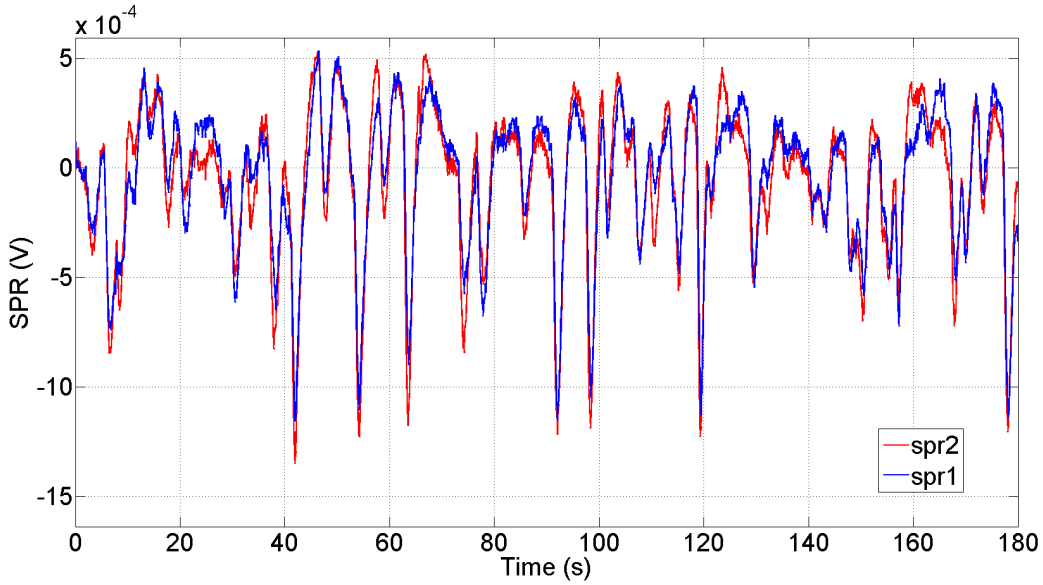


FIGURE 4.20: SPR raw signals acquired with both hands still for two minutes: with no motion, the signals are identical.

4.3.3 Motion Artifact removal

The first assumption of the proposed MA removal algorithm is that the motion artifact, if present, increases the local energy of the acquired signal, since there is a component due to EDA and

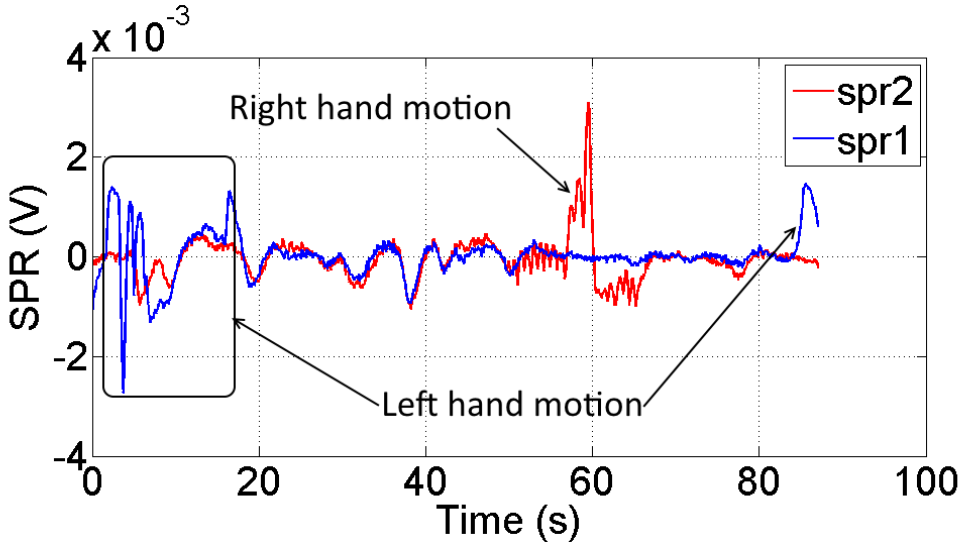


FIGURE 4.21: Top graph: motion artifacts on SPR signals due to hands movements; bottom graph: expected SPR signal when the hands are still.

a component due to motion. Moreover, we assume that motion artifacts are present in only one of the two signals s_1, s_2 at a time. This assumption is very often verified in real data, even if not always, but the proposed solution is a reasonable compromise to mitigate MA, as confirmed by the results which will be presented later in Section 4.3.4. We calculate first the RMS values σ_1 and σ_2 of the input signals s_1 and s_2 on a moving window whose width has been set to $N = 200$ samples, corresponding to a time duration of $1s$ with a sample rate of 200 Sa/s. At the k-th sample (with $k > N$), $\sigma_{1,2}$ will be:

$$\sigma_{1,2}(k) = \sqrt{\frac{\sum_{n=k-N+1}^{k} s_{1,2}^2(n)}{N}} \tag{4.8}$$

After RMS calculation, we choose a smooth threshold function:

$$g(x) = \frac{1}{1 + e^{-2(x-1)}} \tag{4.9}$$

Function $g(x)$ in (4.9) is such that $g(x) \approx 0$ when $0 \leq x \ll 1$ and $g(x) \approx 1$ when $1 \ll x \leq 2$. The threshold function $g(x)$ is used

to calculate the correction factor α used to remove the motion artifact. At the k^{th} sample, $\alpha(k)$ is:

$$\alpha(k) = \begin{cases} g\left(\frac{\sigma_1(k)}{\sigma_2(k)}\right) & \text{if } \sigma_2(k) \neq 0 \\ 1 & \text{if } \sigma_2(k) = 0 \end{cases} \tag{4.10}$$

From (4.9) and (4.10) it is possible to see that $\alpha \to 0$ when $\sigma_1 < \sigma_2$ and $\alpha \to 1$ when $\sigma_1 > \sigma_2$. Finally, we obtain the output of the proposed algorithm by linearly weighting the two input signals with the correction factor α obtained in (4.10). The output of the algorithm, at the k^{th} sample, is computed as:

$$out(k) = \alpha(k) \cdot s_2(k) + [1 - \alpha(k)] \cdot s_1(k) \tag{4.11}$$

Qualitatively speaking, the output in (4.11) follows the smoother signal, i.e., the input signal (whether s_1 or s_2) with lower energy content inside the moving window of 1 s duration.

4.3.4 Experimental setup and algorithm performances

In this section we compare the results obtained using our MA removal algorithm with two of the most commonly applied blind source separation techniques, namely Principal Component Analysis (PCA) and Independent Component Analysis (ICA) [93]. As a matter of fact, one can suppose that $s_1(t)$ and $s_2(t)$ are essentially obtained as the combination of a common component related to the Autonomous Nervous System activity, and of uncorrelated or independent components due to MA and noise. As well known, PCA allows to find statistically uncorrelated components, which do not necessarily represent independent sources, which are instead searched for by the ICA algorithm.

For the experiments in 4.3.4 and 4.3.4, the disturbances are planned. Then, we quantify the quality of signal extraction by evaluating the error in artifact removal as the difference $V_{ERR}(t) = SPR_{EXP}(t) - OUT_i(t)$, where $OUT_i(t), i = 1, 2, 3$, represents the

output of the PCA, ICA and of the proposed algorithm, respectively.

On the contrary, for the experiments described in Section 4.3.4 and Section 4.3.4, we do not have a ground truth since we do not have a still hand at any time. For this reason, we quantify the performance of the algorithm by means of spectral flatness as described in section 4.3.5.

Test on Artificial signals

We initially test the algorithms on artificially created signals that resemble disturbed SPR signals (s_1 and s_2). Figure 4.22 shows the artificial signals we used for comparison. In particular, we simulate the case of non-simultaneous MA appearing at different time instants in the SPR signals taken from the two hands, as well as the case of simultaneous MAs. In the figure, the upper plot shows ten seconds of a recorded original SPR signal, taken from one hand, where we superimpose artificial disturbances, mimicking MA, at time instant 5 s and 7.5 s. The plot in the middle shows the same signal with the disturbance at 7.5 s only, which we pretend to be acquired from the other hand.

In this simplified experiment, therefore, we model the signals from the two hands ("original spr1" and "original spr2 (a)" in the figure) as consisting of the same signal, with one non-simultaneous MA in one of the two hands, and one simultaneous MA in both hands. Figure 4.23 shows the signals obtained after processing with the PCA ("pca1" and "pca2" in the figure) and Fast ICA algorithms ("IC1" and "IC2" in the figure). It is clear from the figure that PCA does not remove the disturbances completely (see signal "pca1"), both the non-simultaneous and simultaneous ones, due to the lack of exact orthogonality between the disturbance and the common signal. It can be seen, however, that ICA is effective in removing the non-simultaneous disturbance (see signal "IC2", neglecting signal scaling). The bottom plot in Fig. 4.23 shows the signal obtained with the proposed algorithm, where the non-simultaneous disturbance is effectively removed.

Figure 4.24 shows the output of the PCA and ICA algorithms by using as input the signals "original spr1" and "original spr2 (b)" of Figure 4.22. In this simplified case, we pretend therefore that the signals from the two hands are identical, except from MA disturbances at about 5 s time instant in one hand, and at about 2.5 s in the other hand, again with a simultaneous MA at 7.5 s. It is clear from the figure that both PCA and ICA cannot completely remove the non-simultaneous artifacts. As a matter of fact, for ICA, we have in this case three components (the two non-simultaneous artifacts and the common signal) and only two measurements. Again, the bottom plot in Figure 4.24 shows the signal obtained with the proposed algorithm, where the non-simultaneous disturbances are effectively removed. All the methods, including the proposed one, fail to remove the simultaneous disturbance. This problem could be inevitable in real situations, where it is difficult to recognize a clear structural difference between MA and SPR signal morphologies, contrary to what happens in these artificial examples.

MA induced moving the hands and disturbing the electrodes test

The test is carried on in laboratory and consists in moving the hands and disturbing the electrodes during the SPR acquisition. In Figure 4.27 the artifacts related to each hand (spr_1 left, spr_2 right) can be noticed.

Figure 4.28 shows the output of the algorithm when the input data are the ones shown in Figure 4.27. It is evident that all the motion artifacts are removed and the output follows the signal which is not affected by disturbances. Comparing Figure 4.27 with Figure 4.25 and Figure 4.26, it is noticeable that PCA and ICA performances are limited as explained in Section 4.3.4 and that the proposed algorithm qualitatively performs much better than PCA and ICA.

Test on driving simulator

After the tests in laboratory, we tried the algorithm on a professional driving simulator; the simulated circuit is Jerez de la

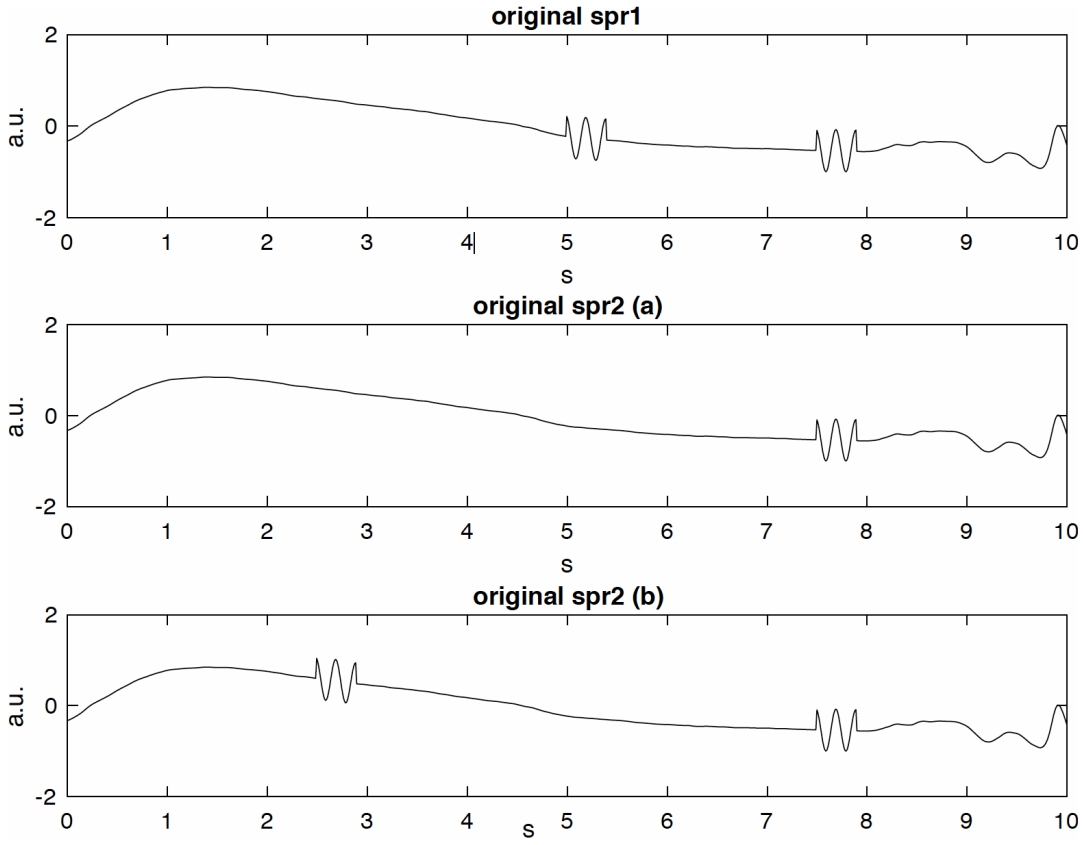


FIGURE 4.22: Artificial test signals. Top: original SPR signal with simulated MA at time 5 s and at 7.5 s. Middle: original SPR signal with simulated MA at 7.5 s. Bottom: original SPR signal with simulated MA at time 2.5 s and at 7.5 s.

Frontera. The driver (not a professionist) never drove on it before. It is evident that the algorithm follows the input with minimum energy when there is discordance between inputs. In Figure 4.29, one example of the signals acquired during the simulation is shown.

City driving test

As a last qualitative example, we show in Figure 4.30 the algorithm output when the signals are taken from a driver in a fifteen minutes drive in city traffic.

Since the number of SPR pulses is very high on a long acquisition, for the sake of clarity, in Figure 4.31 we show a zoom of Figure 4.30 in the time interval [400, 550] s. In this interval, two kinds of motion artifacts are clearly visible. The first one, a high frequency disturbance, is due to muscle contractions and

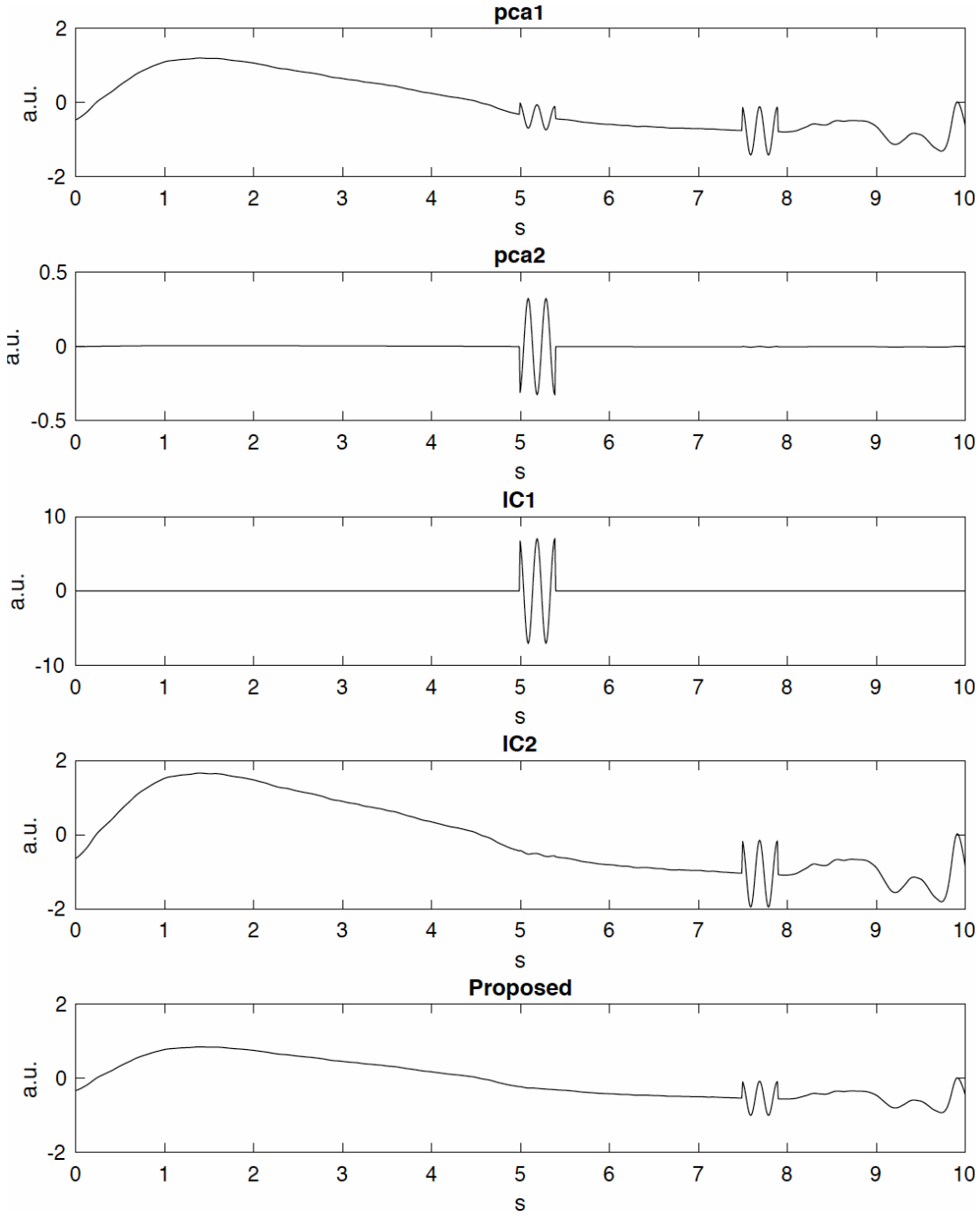


FIGURE 4.23: Results of processing signals “original spr1" and “original spr2 (a)" with PCA and ICA.

is present almost always, especially on the black line. The second one is a long term motion artifact due to a long curve, or a roundabout, appearing at the time instant $t = 450$ s and whose duration is about 5 s. Also in this case, it can be seen that the output of the algorithm is stable, presenting peaks only when they appear in both input signals.

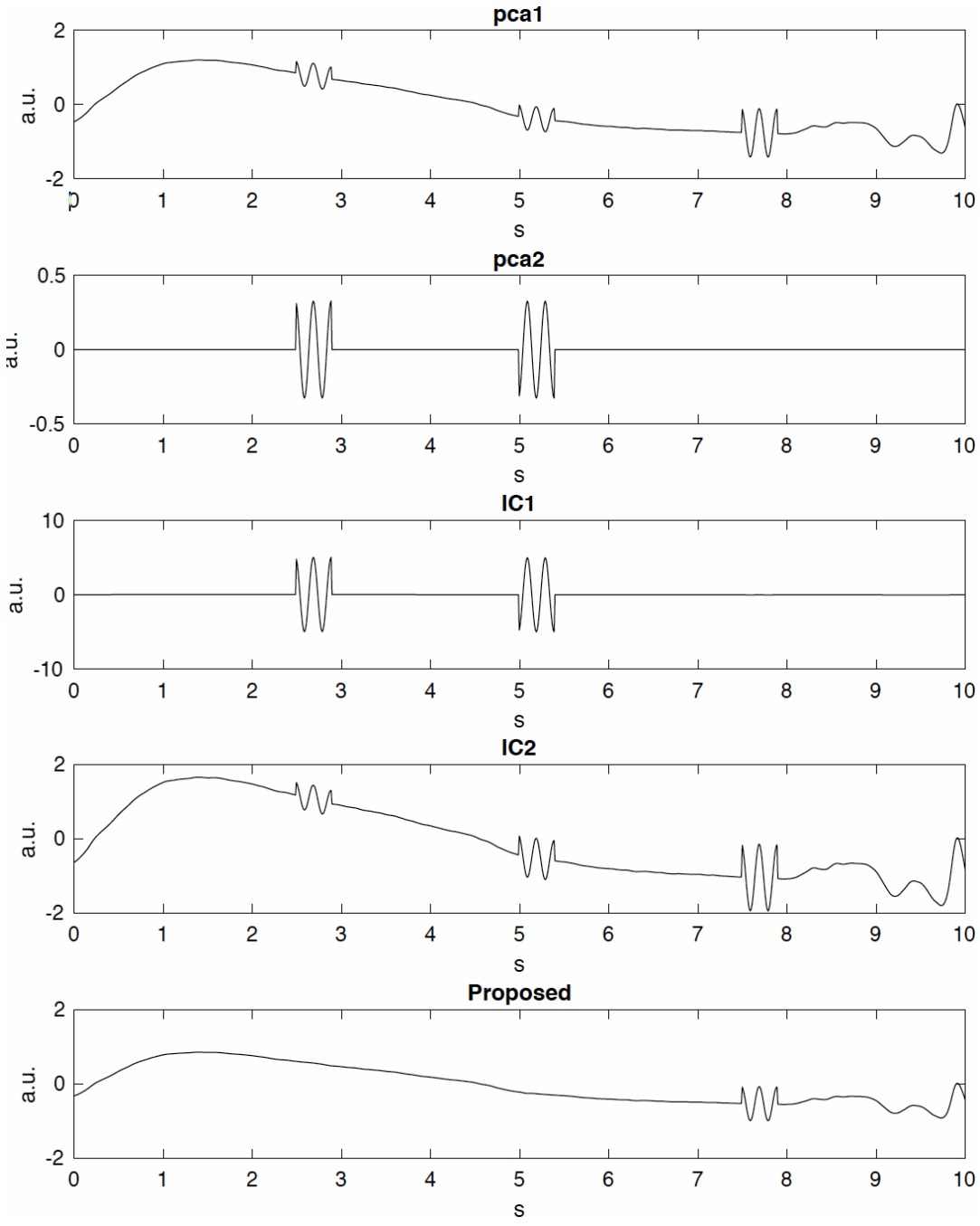


FIGURE 4.24: Results of processing signals “original spr1" and “original spr2 (b)" with PCA and ICA.

4.3.5 Spectral flatness

To quantify the smoothness of the output provided by our algorithm in comparison with PCA and ICA, we evaluate the Wiener entropy (or Spectral Flatness, SPF) [59], defined as:

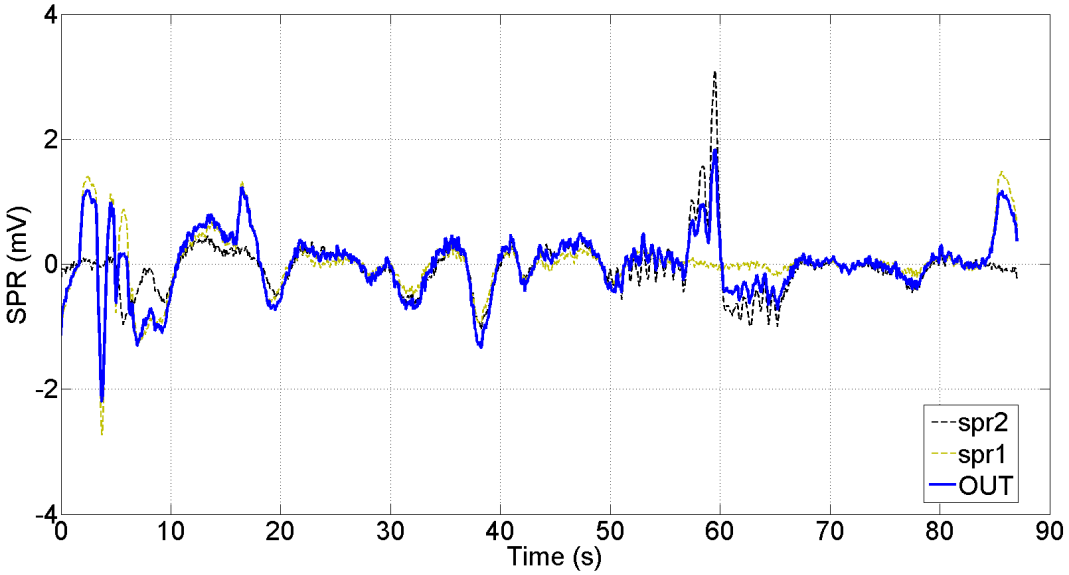


FIGURE 4.25: Raw signals versus Principal Component Analysis (blue line). It is noticeable that the algorithm does not remove all the motion artifacts.

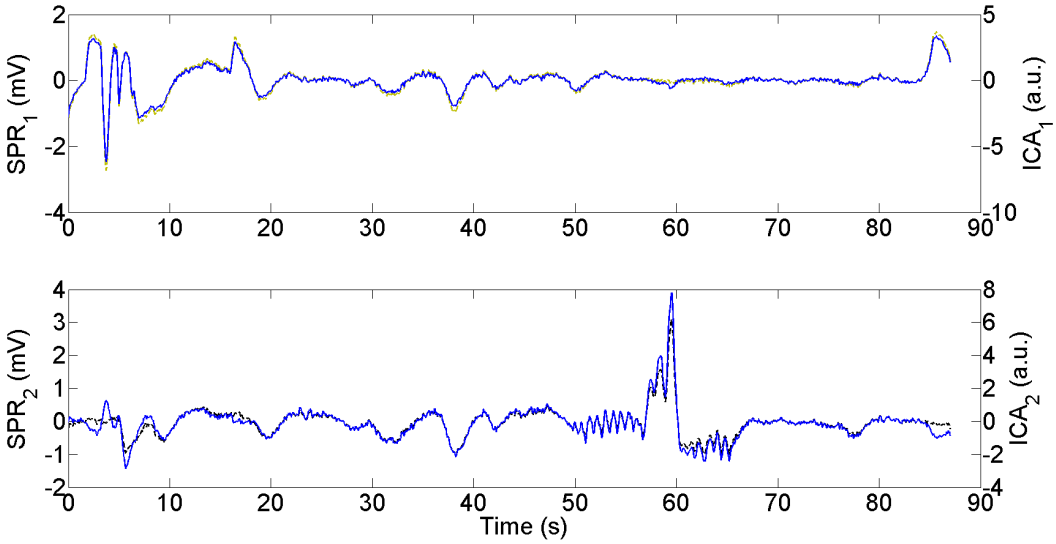


FIGURE 4.26: Raw signals versus ICA algorithm. Top plot: SPR_1 (green dashed) and ICA_1 (blue solid); bottom plot: SPR_2 (black dashed) and ICA_2 (blue solid). It is noticeable that ICA algorithm does not remove the motion artifacts.

$$\text{SPF} = \frac{\sqrt[N]{\prod_{k=0}^{N-1} Y(k)}}{\frac{1}{N}\sum_{k=0}^{N-1} Y(k)} \tag{4.12}$$

where $Y(k)$ is an N-point estimate of the Power Spectral Density

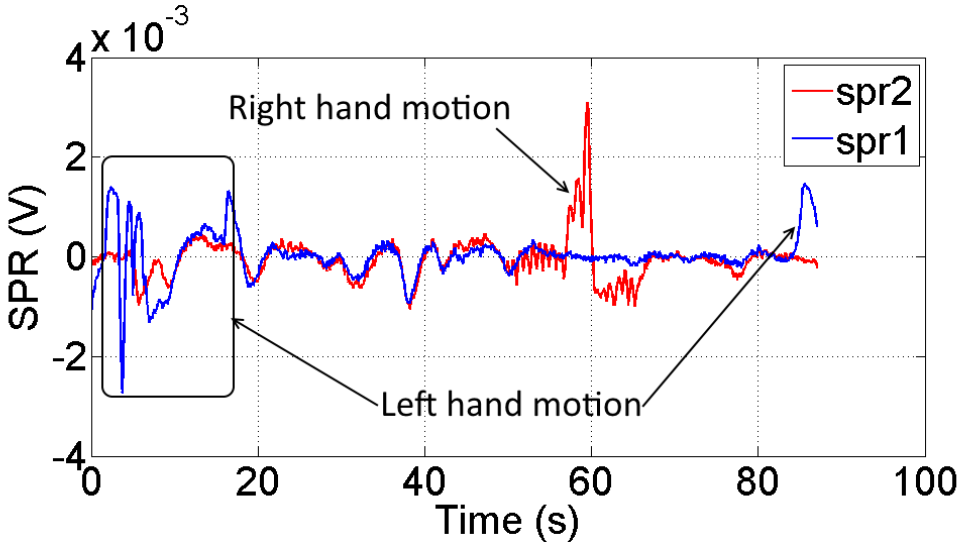


FIGURE 4.27: Motion artifacts on SPR signals due to hands movements

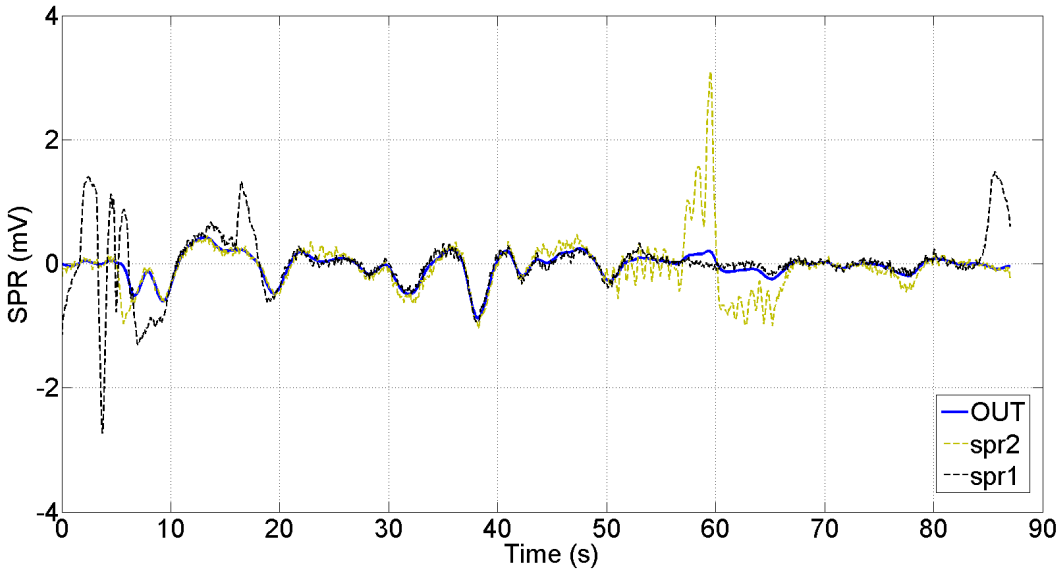


FIGURE 4.28: Raw signals versus the proposed algorithm: the output (blue line) is not affected by motion

of the input. Low SPF values indicate that the signal energy is concentrated in a narrow band (this typically happens for electrodermal signals), while high values of the SPF imply that frequency components are spread in the spectrum (e.g., when random or impulsive components are added to the electrodermal activity).

In Table 4.4 the values of SPF are reported. The proposed algorithm has lowest SPF in almost all the tests, and in some cases it

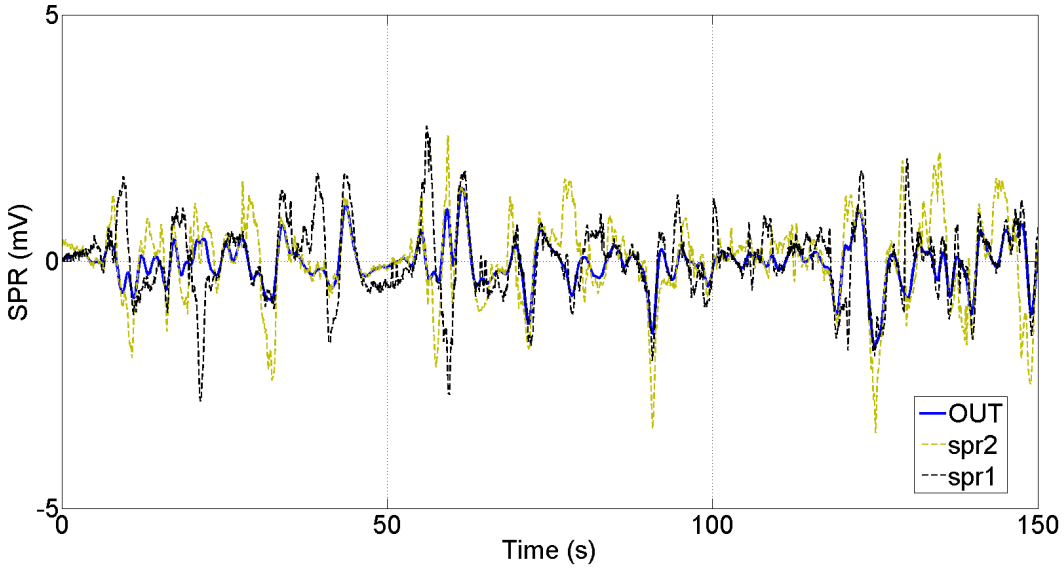


FIGURE 4.29: Signals acquired during a track lap on a driving simulator

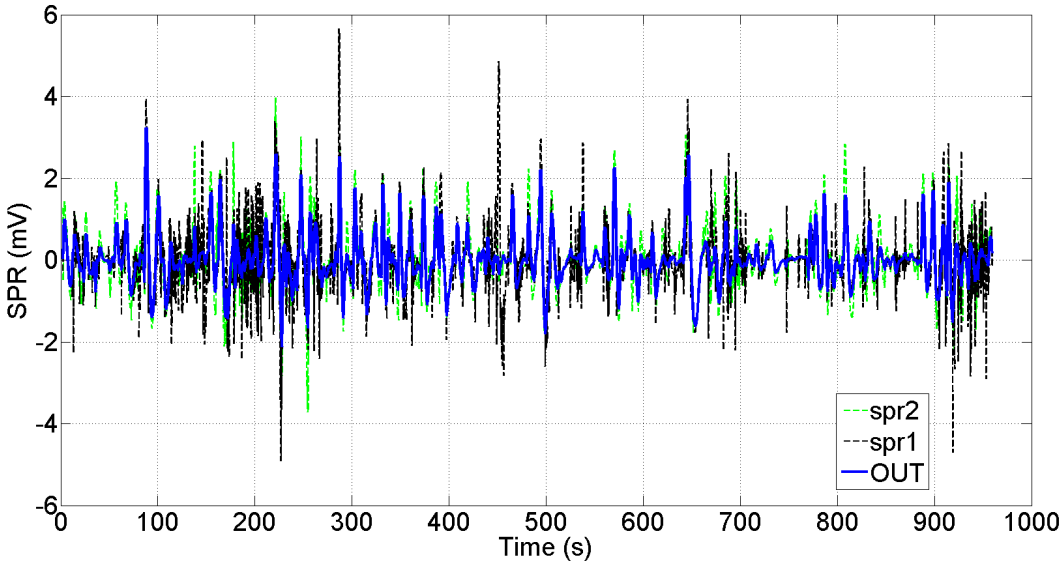


FIGURE 4.30: Signals acquired during a drive session in traffic

is several orders of magnitude lower than using the other methods.

4.4 Real time implementation

The proposed system could be implemented in real time. In particular, the proposed motion artifact removal algorithms only

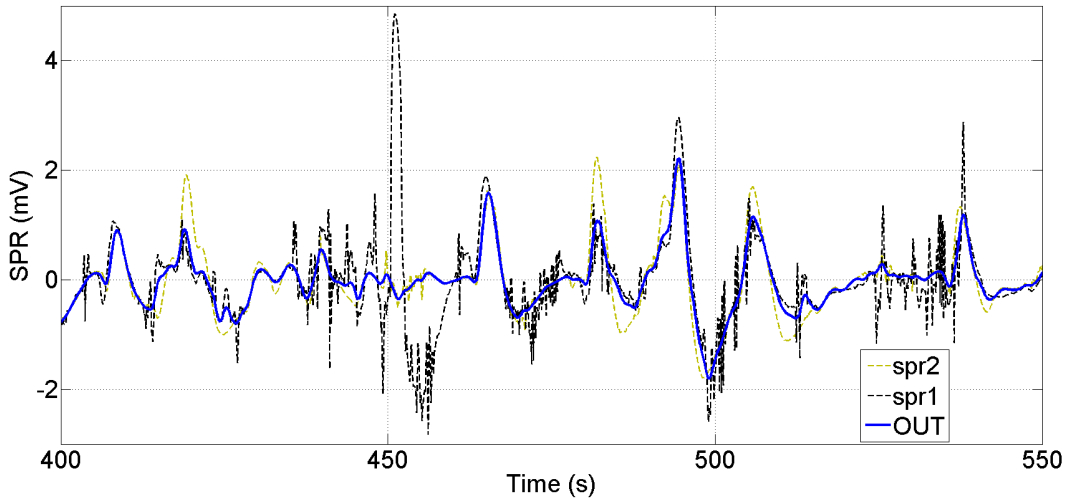


FIGURE 4.31: Signals acquired during a drive session in traffic (zoom); two kinds of artifacts removed: muscle activity and disturbance to electrode (at $t = 450$ s)

SPF $(\cdot 10^{-5})$						
	SPR_1	SPR_2	PCA	ICA_1	ICA_2	Proposed
Lab test$_1$	105.6	9.8	62.7	95.2	8.9	0.4
Lab test$_2$	1.5	4.6	0.5	1.6	1.4	0.02
Lab test$_3$	0.7	3.9	1.2	0.08	2.2	0.03
Lab test$_4$	2.1	1.9	1.9	2.1	0.7	0.6
Simulator test$_1$	4.5	1.5	3.8	1.3	10.7	0.06
Simulator test$_2$	2.9	0.5	1.9	5.6	0.6	0.08
Simulator test$_3$	2.6	1.5	5.5	5.1	1.4	1.2
Simulator test$_4$	0.3	0.7	1.1	1.1	0.3	0.4
Simulator test$_5$	1.6	5.3	2.2	0.4	3.3	0.04
Simulator test$_6$	0.3	0.6	0.4	0.1	2.3	0.5
Simulator test$_7$	0.8	0.4	0.1	0.2	0.3	0.0001
Simulator test$_8$	0.9	1.8	1.6	0.6	5.9	0.8
Simulator test$_9$	2.7	1	1.8	1.8	1.9	0.9
Simulator test$_{10}$	19.9	8.8	40.2	46.5	13.2	2.7
Simulator test$_{11}$	0.09	0.6	0.1	0.6	0.2	0.01
Traffic	0.9	6.3	1.3	7.1	4.4	0.1

TABLE 4.4: Wiener entropy evaluation of the raw SPR signals, PCA, ICA, and proposed method

require the past samples contained in a one second long sliding window to compute the output that follows the smoother SPR signal (we then shift by one sample and do the computation again). Since we collect 200 samples per second for adaptive filtering and 100 samples per second for smoother, there is a 1/200 s-1/100 s interval between each sample, so the constraint

is that the computation must be completed within 1/200 s and 1/100 s respectively.

4.5 Conclusions

This chapter is dedicated to the development of MA removal systems.

The first approach is based on adaptive filtering, exploiting the correlation between hand movements, represented by Steering Wheel angle excursion and Motion Artifact. The system has been tested during an experiment carried out in laboratory while the subjects move the wheel at different speeds. The results showed that the system is able to remove Motion Artifact and allow the detection of stress reactions.

More realistic experiments put in evidence the problems of tracking Motion Artifact from the Steering Wheel due to the driving behavior. Drivers occasionally take the hands off the Steering Wheel (at least one hand per time) and stretch the hand muscles during particular maneuvers and/or operations.

To handle this kind of situations another approach has been developed. The method is based on measuring SPR on both the hands and processing them in order to get only the common part, which is related to mental state, in the output. Although the proposed method cannot clear Motion Artifact when both the hands suddenly simultaneously move in the same way, we rely on the fact that these events barely occur in real driving scenarios. Since our hypothesis is based on drivers behavior, we set several experiments of increasing realism. The tests confirmed our hypothesis and the performances of Motion Artifact removal method outperformed PCA and ICA algorithms. Ultimately, the experiments have shown that Motion Artifact is strongly invasive in driving context and a perfect method for its removal cannot exist; but if the context is properly taken into account we can obtain good results and obtain a Stress (S) signal which can be correctly interpreted and processed to detect

stress. In the next chapter, signal features behavior will be analyzed, then machine learning techniques will be used to classify the drivers' mental state.

Chapter 5

Features extraction and data classification

In this chapter, we process the Stress signal obtained from the Motion Artifact removal system and the RR peaks and Heart Rate (HR) obtained from the Electrocardiogram of the subjects under test, in order to extract features that represent the presence or the absence of mental stress. The features then, are processed and classified with Machine Learning techniques which give an evaluation about the subjects' state.

This chapter is structured in the following way: Section 5.1 is dedicated to the theory and the functioning of the machine learning techniques that will be used to classify the subjects' state. The next sections regard the experiments in which the classifiers will be applied. In Section 5.2 the feature extraction and classification procedure is carried out on the Stress, RR and HR signals. In Section 5.3 and Section 5.4 are presented new experiments in realistic car driving scenarios, carried out on a professional simulator. Support Vector Machines (SVM) and feed-forward Neural Networks (NN) are then employed to classify the collected features. Section 5.5 discusses a possible real time implementation of the proposed systems. Finally in Section 5.6 the conclusions are given.

5.1 Machine learning

To evaluate the mental state of the subjects under test, we use machine learning techniques. The classifiers we employ in our work are the Support Vector Machine and the Neural Network. In this section, the functioning of the classifiers is explained in order to understand how the features of the physiological signals we measure are processed to detect drivers' stress.

5.1.1 Support Vector Machine

The idea behind the Support Vector Machines (SVM) classifier is to split the training data set, in order to classify new data, on the basis of their position in the N-dimensional hyperspace in which the feature values lie. Figure 5.1 shows a 2D representation of such a hyperspace. The hyperplane $h(\boldsymbol{x})$ separates the hyperspace in two regions, assigning to the features a binary label.
An explanation of binary SVM functioning follows.

Given the training data in the form:

$$S = \{(y_i, \boldsymbol{x}_i) | y_i \in \{+1, -1\}, \boldsymbol{x}_i \in \Re^N \forall i = 1, ..., l\} \tag{5.1}$$

where $\boldsymbol{x}_i$ is the i-th training sample and y_i is the label associated to $\boldsymbol{x}_i$. $\boldsymbol{x}_i$ is a vector of dimension $Nx1$, y_i is a scalar whose value can be -1 or +1. Each label assign a sample to a specific class.

A decision function $f(x)$ must be found to separate the hyperspace in regions in which all the samples have the same label. This hypothesis is verified if the decision function has the following property:

$$f(x) = \begin{cases} +1 & when \quad h(\boldsymbol{x}) > 0 \\ -1 & when \quad h(\boldsymbol{x}) < 0 \end{cases} \tag{5.2}$$

The separating surface, which represents the boundary between positive (+1) and negative (-1) labeled regions, is defined as follows:

$$\mathcal{H} = \{\boldsymbol{x} | h(\boldsymbol{x}) = 0\}. \tag{5.3}$$

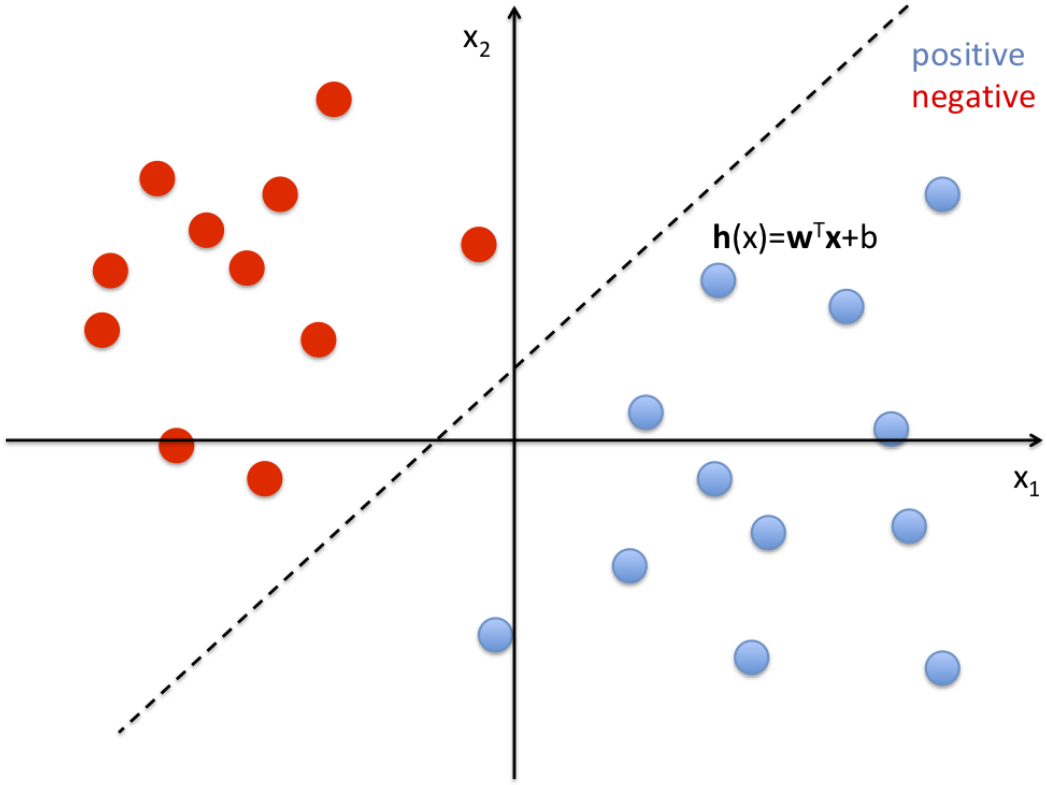


FIGURE 5.1: 2D space with linearly separable data separated by a hyperplane.

In first instance, we consider the case of a linearly separable scenario as depicted in Figure 5.1. In this scenario a hyperplane defined as $h(\boldsymbol{x})$ represented by a linear function:

$$h(\boldsymbol{x}) = \boldsymbol{w}^T \boldsymbol{x} + b \tag{5.4}$$

where $\boldsymbol{w}$ and b are the hyperplane parameters of NxN and $Nx1$ dimensions, is able to separate positive and negative regions.

Infinitely hyperplanes that can correctly separate the positive and negative instances may exist; the optimal hyperplane is the one with the largest gap between the classes because it will be more robust to any perturbation of the training data.

Since the margin could be measured in different ways, we formally define the distances of a feature x to the set $\mathcal{H}$

$$Distance(\boldsymbol{x}, \mathcal{H}) = \min_{\boldsymbol{p} \in \mathcal{H}} (\|\boldsymbol{x} - \boldsymbol{p}\|) \tag{5.5}$$

where $\boldsymbol{p}$ is the point in the hyperplane with the smallest euclidean distance from the closest element of the dataset; this setting is called **Maximum Margin Classifier**, where the margin is defined as the distance between the hyperplane and the closest dataset element.

If H is a hyperplane we can apply Theorem 2.2 in [72] and write Equation 5.5 in closed form:

$$Distance(\boldsymbol{x}, \mathcal{H}) = \frac{|\boldsymbol{w}^T \boldsymbol{x} + b|}{\|\boldsymbol{w}\|^*} \tag{5.6}$$

Different norms can be used, we choose the dual norm $\|\|^*$.

The margin, which measures the gap between the separating surface H and the nearest instance, is defined as follows:

$$M = \min_i \frac{|\boldsymbol{w}^T \boldsymbol{x}_i + b|}{\|\boldsymbol{w}\|^*} \tag{5.7}$$

When the instance x_i lies on the correct side of the hyperplane (i.e. if a positive labeled instance lies on the sector of the hyperplane that belongs to positive instances), the numerator of M could be simplified as follows:

$$|\boldsymbol{w}^T \boldsymbol{x}_i + b| = y_i \left(\boldsymbol{w}^T \boldsymbol{x}_i + b\right) > 0 \tag{5.8}$$

Thus, to find a maximum margin classifier, we can instead solve the following problem:

$$\max_{\boldsymbol{w}, b, M} M, \quad s.t. \quad \frac{y_i \left(\boldsymbol{w}^T \boldsymbol{x}_i + b\right)}{\|\boldsymbol{w}\|^*} \geq M \tag{5.9}$$

It could be verified that any non-zero multiple of the optimal solution $(\bar{w}, \bar{b})$ is still an optimal solution. Therefore, we could set:

$$M\|w\|^* = 1 \tag{5.10}$$

than, the problem can be written in the equivalent form:

$$\min_{w,b} \|w\|^*, \;\; s.t. \;\; y_i \left(w^T x_i + b\right) \geq 1, \;\; \forall i = 1.....l \tag{5.11}$$

In real scenarios however, we do not have linearly separable regions. Therefore, the model should also be able to handling non separable data. Figure 5.2 shows a non-linearly separable dataset. To find the optimal hyperplane then, condition 5.11 must be violated and a loss function $\xi(\cdot)$ must be introduced:

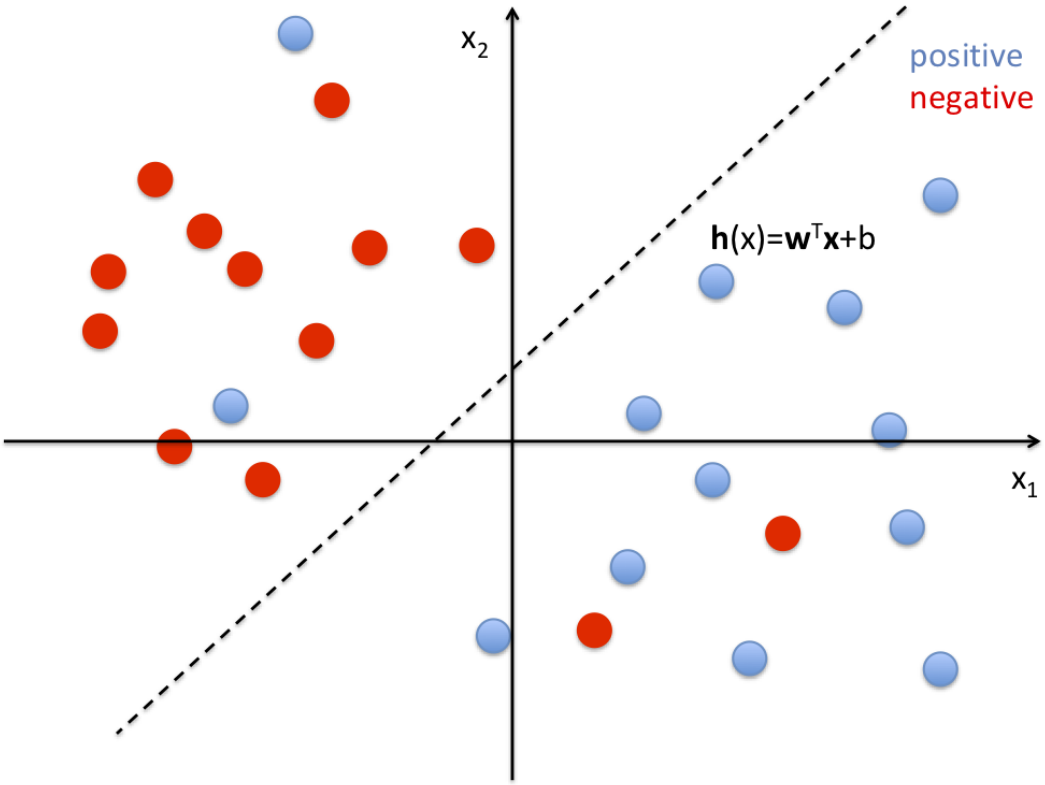


FIGURE 5.2: 2D space with non-linearly separable data separated by a hyperplane.

$$\xi_{L2}\left(w, b; y_i, x_i\right) = max\left(0, 1 - y_i\left(w^T x_i + b\right)\right)^2 \tag{5.12}$$

The goal of the loss function is to contrast the optimal hyperplane objective function, penalizing the elements having values that significantly differ from the other elements of dataset which belong to the same class, known as loss terms.

We need to find the optimal hyperplane which excludes the loss terms and, at the same time, maintains a large margin. Thus, there are two objectives, and there may be multiple ways of balancing them.

For example, one can have a smaller margin with one violated instance. Alternatively, one can have a larger margin with more violated instances. Since it is more difficult to minimize multiple objectives simultaneously, some tradeoffs between $\|\boldsymbol{w}\|^*$ and the loss terms are introduced. If the distance is defined by the Euclidean norm, and the L1 hinge loss is chosen, we have the classic SVM formulation in [20]:

$$\underset{\boldsymbol{w},b}{min} \quad \tfrac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + C\sum_{i=1}^{l}\xi_{L1}\left(\boldsymbol{w}, b; y_i, \boldsymbol{x}_i\right) \tag{5.13}$$

composed of two parts: a norm from the maximum margin objective, and the loss terms from the violation of the models. The maximum margin part is also called the Regularization Term, which is used to control the model complexity with respect to the violation of the models. From the perspective of numeric stability, the regularization term also plays a role in the balance between the accuracy of the model and the numeric range of $\boldsymbol{w}$. It ensures that $\boldsymbol{w}$ does not assume extreme values.

Equation 5.13 could be written in the following form:

$$\underset{\boldsymbol{w},b,\xi}{min}\tfrac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + C\sum_{i=1}^{l}\xi_i, \quad y_i\left(\boldsymbol{w}^T\boldsymbol{x}_i + b\right) \geq 1 - \xi_i, \geq 0 \quad , \forall i = 1, ..., l \tag{5.14}$$

defining the scalar $\xi_i = \xi_{L1}\left(\boldsymbol{w}, b; y_i, \boldsymbol{x}_i\right)$.

Different from the original non-differentiable problem, this formulation is a convex quadratic programming problem over some linear constraints. Since this problem is a constrained optimization problem, a Lagrangian relaxation approach may be used to

derive the optimality condition [40]. Following the procedure in Section 7.4 of [7] the Lagrangian relaxation method is applied to our problem.

In several cases, a hyperplane cannot be a valid surface to define the regions properly; or maybe, it is just more convenient to use a different kind of separating surface rather than a hyperplane.

Equation 5.4 can be generalized as follow:

$$h(x) = \boldsymbol{w}^T \phi(x) \tag{5.15}$$

where ϕ is a non-linear kernel.

Specifically, if we apply a feature transformation described in [7][Section 7.5], the data set becomes linearly separable in the mapped space. A widely used kernel is the Radial Basis Function (RBF), also known as Gaussian kernel

$$k\left(\boldsymbol{x}, \boldsymbol{z}\right) = exp\left(-\gamma \|\boldsymbol{x} - \boldsymbol{z}\|^2\right) \tag{5.16}$$

where $\boldsymbol{x}$ and $\boldsymbol{z}$ are two samples that belong to the features dataset, $k\left(\boldsymbol{x}, \boldsymbol{z}\right) = \phi\left(\boldsymbol{x}\right)^T \phi\left(\boldsymbol{z}\right)$ and γ is a defined hyper-parameter.

Another widespread kernel type for $\phi(x)$ is the polynomial function of a certain defined order.
Figure 5.3 shows an example of a non-linear separation boundary.

The reader can refer to [7], for a detailed explanation about SVM theory, including the Lagrangian relaxation approach used to derive the optimality condition, other kernel tricks, the functioning of algorithms employed by the solvers and multiclass SVM classifiers.
For a practical guide to employ SVM, the reader can refer to [14].

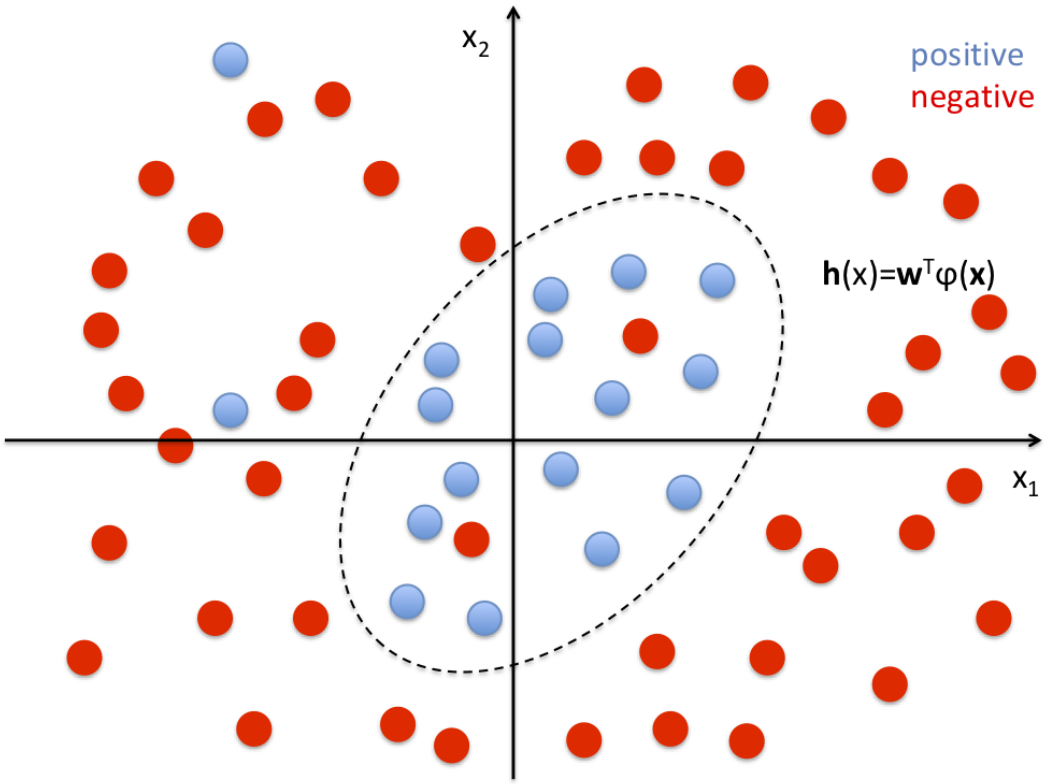


FIGURE 5.3: 2D space with data separated by a surface obtained by RBF kernel.

5.1.2 Neural Networks

An artificial neural network (ANN) or neural net is a graph of connected units representing mathematical models of biological neurons. Those units are sometimes referred to as processing units, nodes, or simply neurons. The units are connected through unidirectional or bidirectional arcs with weights representing the strength of the connections between units. This is inspired from the biological model in which the connection weights represent the strength of the synapses between the neurons, inhibiting or facilitating the passage of signals.

The neural network processes the input data through several steps. Each neuron takes the input from the neurons which belong to the previous layer and transmits its output to the neurons of the next layer.

For classification problems, a neural network is characterized by the characteristics listed below:

• The neuron model or the mathematical model of a neuron that describes how a unit in the network produces an output from its inputs and the role it plays in the network (input unit, output unit, or computing unit).

• The architecture or the topology that outlines the connections between units, including a well-defined set of input and output units.

• The data encoding policy describing how input data or class labels are represented in the network.

• The training algorithm used to estimate the optimal set of weights associated with each unit.

The details of the following subsections are well-described in [7].

Neuron model

In Figure 5.4 the structure of a neuron is depicted. It consists of:
• N gains w_i, also called synaptic weights, on which each of the N inputs x_i is weighted
• a net value function ξ, which utilizes the synaptic weights (or parameters) w of the unit to summarize input data into a net value, ν, as:

$$\nu = \xi(\boldsymbol{x}, \boldsymbol{w}) \tag{5.17}$$

with $\boldsymbol{x} = (x_1, x_2, x_3, x_4, x_5)$, $\boldsymbol{w} = (w_1, w_2, w_3, w_4, w_5)$

• an activation function, or squashing function, ϕ, that transforms net value into the unit's output value o as:

$$o = \phi(\nu) \tag{5.18}$$

The net value function mimics the behavior of a biological neuron, as it aggregates signals from linked neurons into an internal representation. Typically, it takes the form of a weighted sum, a distance, or a kernel. The activation function simulates the behavior of a biological neuron as it decides to fire or inhibit signals, depending on its internal logic. The output value is then

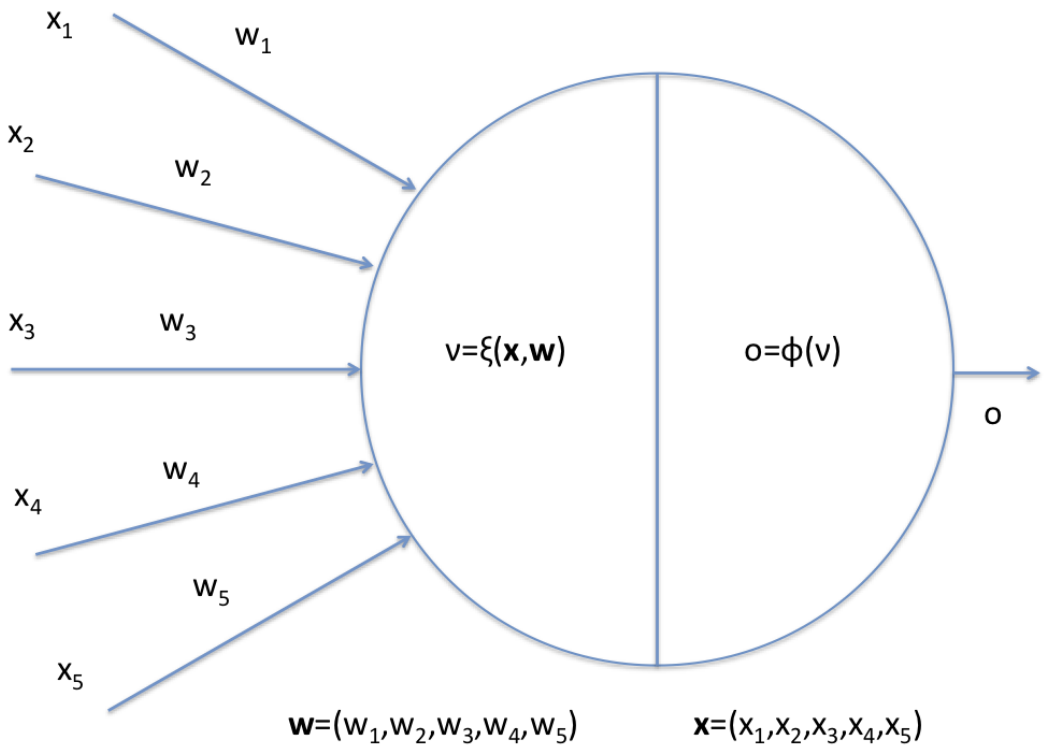


FIGURE 5.4: Neuron structure

dispatched to all receiving units as determined by the underlying topology. Among activation functions, the most widely-used ones include linear, step, threshold function, sigmoid and hyperbolic tangent functions. Table 5.1 lists several types of commonly used units.

TABLE 5.1: Type of units commonly used in Neural Network

Type of unit	net function	activation function
linear threshold unit	weighted sum	step or sign function
linear unit	weighted sum	linear or piecewise linear
sigmoidal unit distance unit	weighted sum distance	sigmoid or tanh linear or piecewise linear
gaussian unit	distance	gaussian kernel

Architecture

In Figure 5.5 an example of a standard feed-forward Neural Network architecture is shown. It is composed by several layers which, in turn, are composed by a certain number of neurons. The layers are connected each other through their neurons.

Here we refer to a standard feed-forward NN whose the architecture is composed of an input layer, one or more hidden layers and an output layer, which are connected in cascade as we can see in Figure 5.5. The intermediate layers gradually extract the information from the input to the output layer through the processing carried out by their neurons.

In other types of NN, many kinds of architectures and connections can be found. To deepen the topic, the reader can refer to [35, 7].

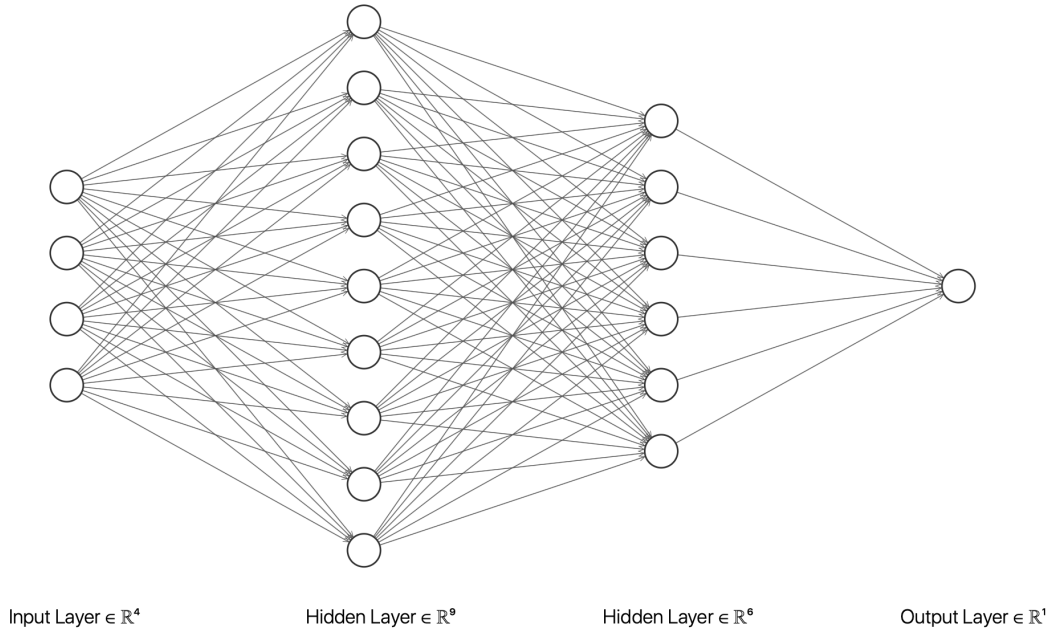


FIGURE 5.5: Architecture of a Neural Network with 2 hidden layers

Data encoding policy

At the unit level, computation is a two-stage process, in which the unit computes the net value using the net value function, and then computes its output using the activation function. At the network level, computation starts with a presentation of the data to the input units in the network.

The input units then transmit the data (without modifications) to the connected successor units. These units are then activated, each one computing its net value and output, and in turn transmitting that output to the connected recipient units. This process is called spreading activations or forward propagation, as

activation values spread from the input layer up to the output layer in a sequential fashion. Depending on the network connectivity, the spreading activation process can be implemented in a Parallel Distributed System (PDP) [95].

In networks without feedback, the process of activation and propagation ends in the output layer. In networks with feedback, such as recurrent networks, the activation propagation never ends, instead following a dynamic trajectory through the state space, as units and layers are continuously updated. The state of the network is represented by the values of its synaptic weights.

Training algorithm

The algorithm by which a Neural Network is trained, consists of the backward propagation of errors, also known as Backpropagation. Basically it consists of adapting the weighting factors using a gradient-based procedure, in order to obtain the desired output for the training dataset. Initially, training data inputs are processed as usual as in a feed-forward network, from the input to the output layer. Then, the gradients are calculated beginning from the last till the first layer through the Backpropagation algorithm. Then, the synaptic weights are updated as

$$w \leftarrow w_0 + \nabla w \tag{5.19}$$

$$\nabla w = -\eta \frac{\partial E}{\partial w} \tag{5.20}$$

where w_0 is the previous value of w, E is the error between the output and the desired output (also called target vector). In the output layer, the target vector corresponds to the label vector $\boldsymbol{y}$, while in the hidden layers it is computed through the optimization procedure.

Using the mean square error, E is defined as:

$$E = E_{mse}\left(\boldsymbol{x}; \boldsymbol{w}\right) = \sum_{k=0} K \left(o_k - t_k\right)^2 \tag{5.21}$$

Specifically, if w_{jk} is the weight connecting the j-th unit in the last hidden layer and the k-th output unit with net value ν_k and output $o_k = \phi(\nu_k^{(L)})$:

$$\frac{\partial E}{\partial w_{jk}} = \frac{\partial E}{\partial \nu_k^{(L)}} \frac{\partial \nu_k^{(L)}}{\partial w_{jk}} = \delta_k z_j^{(L-1)} \tag{5.22}$$

where $z_j^{(L-1)}$ is the output of the j-th unit in the last hidden layer H_{L-1} and where we define $\delta_k = \frac{\partial E}{\partial \nu_k^{(L)}}$.

Similarly, the error gradient relative to weight $w_{hj}^{(l)}$ between the h-th unit in hidden layer $l-1$ and the j-th unit in hidden layer l is computed as:

$$\frac{\partial E}{\partial w_{hj}^{(l)}} = \frac{\partial E}{\partial \nu_j^{(l)}} \frac{\partial \nu_j^{(l)}}{\partial w_{hj}^{(l)}} = \delta_j^{(l)} z_h^{(l-1)} \tag{5.23}$$

By referencing the output layer as H_L and input layer as H_0, the weight update is simply:

$$\nabla w_{ij}^{(l)} = -\eta \delta_j^{(l)} z_i^{(l-1)} \tag{5.24}$$

for any given weight $w_{ij}^{(l)}$ with $l \in \{1, ...L\}$.

To deepen the topic, the reader can refer to [7] and [35], which provide a complete explanation of NN variants along with application examples.

5.2 Feature extraction and classification 1

In this section, a classification procedure is implemented on the stress signal S obtained during the experiment undertaken in Section 4.2, which has been described in 4.2.5. Figure 5.6 shows the system.

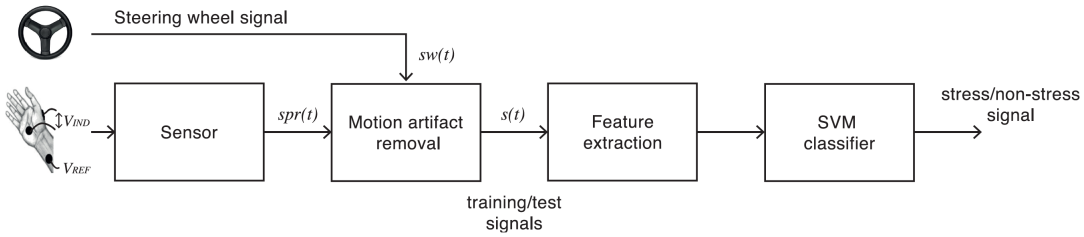


FIGURE 5.6: Block diagram of the proposed system

TABLE 5.2

Electrode positioning	Electrode 1: Hand Palm (sweat glands present) Electrode 2: Hand Back (sweat glands not present) Electrode 3: Wrist (for voltage reference)
Battery	Lithium Polymer
V_{REF}	$1.65V$
Input impedance	100 MΩ
Gain	80
Range	± 20 mV
Frequency Band	[0.08 40] Hz
Sampling Frequency	200 Hz
Transmission mode	Bluetooth
ADC	12 bit

5.2.1 Sensor and Steering Wheel

The Sensor and the Steering Wheel are the ones described in Section 4.1.1 and Section 4.1.2 respectively and which have been depicted in Figure 4.1.

We briefly summarize the sensor characteristics in Table 5.2; for details, see Section 4.1.1 and [4]:

Regarding the Steering Wheel, we briefly remind that it is a Logitech G Driving Force GT Steering Wheel which gives an output that ranges from [-1 to 1], representing an excursion within the range of [-90 90]°. The reader can refer to the relative section for details.

5.2.2 Motion Artifact removal

As for the Sensor and Steering Wheel blocks, Motion Artifact Removal blocks are the ones described in the experiment in Section 4.2. See Section 4.2.2 for a complete explanation.

We assume that Motion Artifact is correlated with the motion of the steering wheel [6], modeling the acquired SPR signal as:

$$spr(t) = sw \star h(t) + \hat{s}(t). \tag{5.25}$$

In (5.25), $sw(t)$ is the Steering Wheel (SW) angle excursion, simultaneously recorded with the SPR signal, while $\hat{s}(t)$ is the SPR signal component actually related to emotional events.

The LMS algorithm is used to estimate an adaptive filter $\hat{h}(t)$ which minimizes the expected MSE $E[s^2(t)]$, $s(t) = spr(t) - sw \star \hat{h}(t)$. In such a way, we remove from the recorded $spr(t)$ any component which is linearly correlated with the steering wheel angle and causes the Motion Artifact. The use of an adaptive procedure can take into account possible time-varying relations between Motion Artifact and its effect on the recorded $spr(t)$. According to model (5.25), the residual signal $s(t)$ is therefore a suitable estimate of the emotional component $\hat{s}(t)$, where Motion Artifact has been removed or at least reduced. The signal $s(t)$ is then further processed for classification.

Differently from the system employed in 4.2, the stress signal $s(t)$ is not processed through SNEO operator. This time the purpose is to classify the driver's stress relying on several signal characteristics and make the evaluation more reliable as explained in the following sessions.

5.2.3 Feature extraction

The residual signal s(t) is therefore analyzed by extracting seven statistical features from each time interval signal block, which we chose to be 15 s long, as it will be described in Section 5.2.6.

5.2.4 Feature labeling and classifier setup

To detect the stress component in a given time interval we use a classification algorithm. In particular, we use the Support Vector Machine (SVM) supervised learning algorithm. We ultimately want to label a given time interval in a binary way, as in "with stress" (labeled as "1") and "without stress" (labeled as "0").

The SVM classifier has been set up using the MATLAB routine functions (MATLAB 2017.a) with a Radial Basis Function kernel. The Bayesian optimization procedure has also been applied during the training procedure. The experimental setup is fully described in the following section. The experimental results are then presented in 5.2.6.

5.2.5 Experimental Setup

The data we used to test the classifier are derived from the testing scenario we considered in Section 4.2. We performed several experiments in which acute stress events were introduced throughout a drive test simulation.

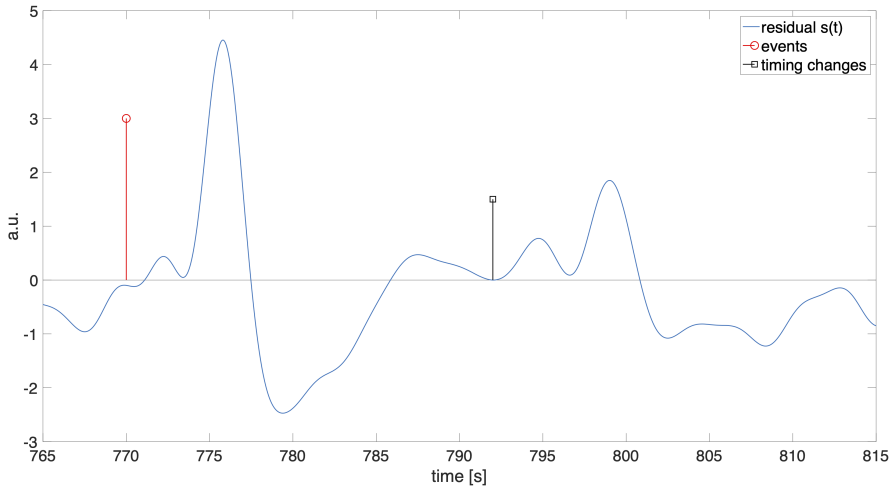


FIGURE 5.7: A portion of a subject residual signal

Figure 5.7 shows a 50 s long portion of the residual signal of a subject, after a sound stress event (the first vertical line with a

circle marker) and after a frequency metronome change (the second vertical line with a square marker). The characteristic stress peaks, as described in [5], are clearly visible after each trigger.

5.2.6 Experimental results

We used the data from 12 subjects out of the 17 to build the classifier. To populate samples of the "stress" class, we consider all the stress inducing events, i.e., all the sound stress triggers and the metronome frequency changes. For both, we selected the events where we noticed a stronger response, and we removed the ones where there was a weaker or no response. We kept a total of 11 stress inducing events. Signal blocks with 15 s duration, belonging to the "stress" class, are then extracted across the trigger time position.
In particular, in order to mimic the fact that in the actual classification task we are of course not synchronized with stress events, we choose three time intervals, which overlap the stress trigger event with different timings. If we call t_0 the time a given stress trigger or metronome change happens, the first interval starts from $t_0 - 5$ s to $t_0 + 10$ s, the second interval from t_0 to $t_0 + 15$ s, and the third interval starts from $t_0 + 5$ s to $t_0 + 20$ s (see also Fig. 5.8).

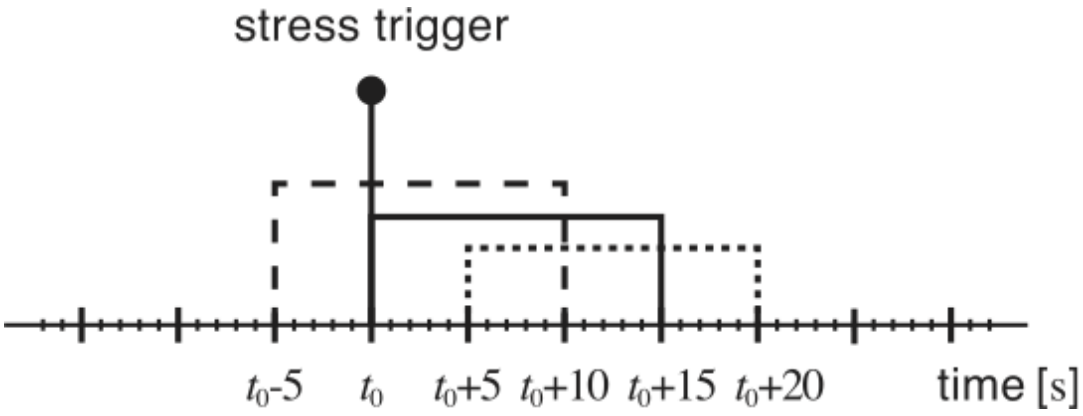


FIGURE 5.8: The overlapping intervals we consider for classification, for each stress trigger

All these intervals are labeled as "1", with stress. We assume the stress as non-existent ("0" class, with no stress) in an equal number of 15 s intervals, extracted far from the stress triggers and the

metronome changes time positions. We are now able to build a dataset of 11 (stress inducing events) $\times$ 3 (time intervals) $\times$ 12 (subjects) $=$ 396 feature vectors labeled as "1", with stress, and other corresponding 396 feature vectors labeled as "0", with no stress. Each feature vector is composed by 7 features, as introduced above.

The features, extracted from the stress signal, should be chosen to identify the stress signal characteristics, discerning when they are related to a stressed state and when not. In particular, we consider the energy (since the energy of the signal peaks due to stress is much higher than the energy of the stress signal in normal conditions), the mean absolute value, the variance, the mean absolute derivative and the max absolute derivative to identify the signal sudden raising and fall when a stress event occurs. We also consider the Peak to Average Power Ratio (PAPR) and kurtosis to help discriminate the shape of the peaks. Each feature, in all corresponding feature vectors, has been normalized to the range [0, 1].

In Figure 5.9 and Figure 5.10, the histograms of the features which have been extracted from the $s(t)$ signals belonging to the subjects under test are shown. The blue histograms are related to the features labeled as "1", the red histograms are related to the features labeled as "0". It can be noticed that the feature values corresponding to "0" (no stress) exhibit different histograms than in the case "1" (stress). Even if the histograms of PAPR and kurtosis alone do not show evident differences between the two cases, these features proved to be useful for classification when used together with the other features.

We use about 70% of this data for the training, and 30% for the test. This let us have almost 800 data samples for the training. This is a limited number, but considering the ratio between the training and the test data, this number can be acceptable [79]. In addition, a 10-fold cross validation has been used. We are able to count the True Positives (TP), the number of test intervals that the classifier correctly classified "stress"; the False Negatives

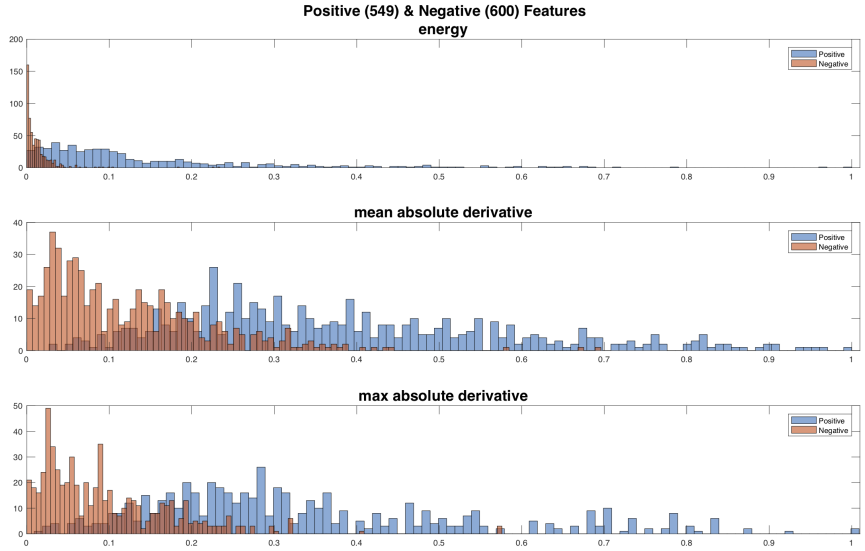


FIGURE 5.9: The overlapping intervals we consider for classification, for each stress trigger

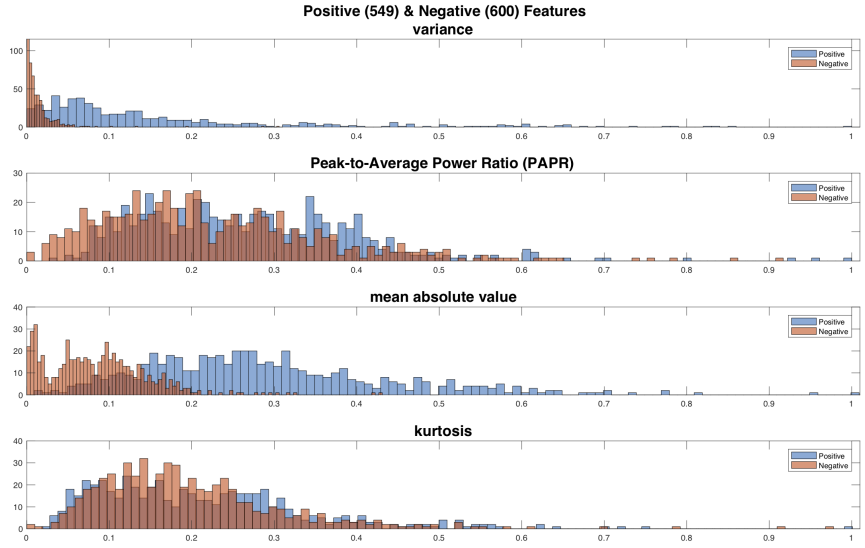


FIGURE 5.10: The overlapping intervals we consider for classification, for each stress trigger

(FN), the number of "non-stress" classified intervals which instead we expected as "stress"; the True Negatives (TN), the correctly classified "non-stress" intervals; the False Positives (FP), the number of incorrectly labeled "stress" intervals that instead we expected as "non-stress". With these values, we can create the confusion matrix of the system, which is shown in Fig. 5.11. The total average accuracy resulted to be 87.40%.

actual \ predicted	stress	non-stress
stress	84.03%	15.97%
non-stress	9.24%	90.76%

FIGURE 5.11: Confusion matrix for the two classes classifier ("stress"/"non-stress") with total average accuracy of 87.40%

We used the data of the remaining 5 subjects to test the classifier in an actual experiment, under the hypothesis that we do not know anything about the stress triggers, or their timing, so to evaluate the classifier stress detection capability in a more realistic situation. We will consider both the trigger sounds and the metronome frequency changes as stress inducing factors. The classifier will start evaluating the signal from the beginning, considering 15 seconds long intervals. The time window length has been chosen on the basis of the stress impulse duration which, as it can be seen in Figure 5.12, is on average 15 s. A new evaluation will start every 5 seconds. The classification result will be displayed at the end of each interval. So, every 5 seconds we will have an indication of the stress presence of the previous 15 seconds interval.

We count a True Positive detection as happening if at least 1 out of the 4 subsequent intervals after the stress trigger, is labeled as positive by the classifier. If all the 4 intervals are labeled as negative, we consider it as a False Negative detection. As in [5], we do not consider the False Positives (FP) and True Negatives (TN) detections, because we cannot truly verify when stress should not be present in a subject at all, and more importantly because we ultimately want to evaluate the classifier performance in detecting the stress component in proximity of the controlled stress trigger events. We then compute the Recall (also known as Sensitivity), which is defined as:

$$\text{Recall } (\%) = \frac{\text{TP}}{\text{TP+FN}} \cdot 100 \tag{5.26}$$

where TP+FN represents the total number of stress triggers. The

TABLE 5.3: **Top:** TP, FN and Recall values for the five evaluated subjects, considering stress reactions after sudden random sounds. **Bottom:** TP, FN and Recall values for the five evaluated subjects, considering stress reactions after sudden random sounds and metronome frequency changes

Subject no.	1	2	3	4	5
TP	11	11	11	11	10
FN	0	0	0	0	1
Recall (%)	100	100	100	100	90.91

Subject no.	1	2	3	4	5
TP	29	30	31	33	33
FN	5	4	3	1	1
Recall (%)	85.29	88.24	91.18	97.06	97.06

results are presented in Table 5.3 (a), where we have computed the Recall for each of the five evaluated subjects, considering only the sound triggers (which we expect to provide a stronger response) without the metronome. Then, we average these Recall values to obtain the overall performance, which equals 98.18%. Table 5.3 (b) shows instead the Recall value for each subject, but considering both the sound triggers and the metronome changes as stress inducing factors. Again, we average these Recall values to obtain the overall performance, which equals 91.77%. This confirms the ability of the proposed system to detect the induced stress events.

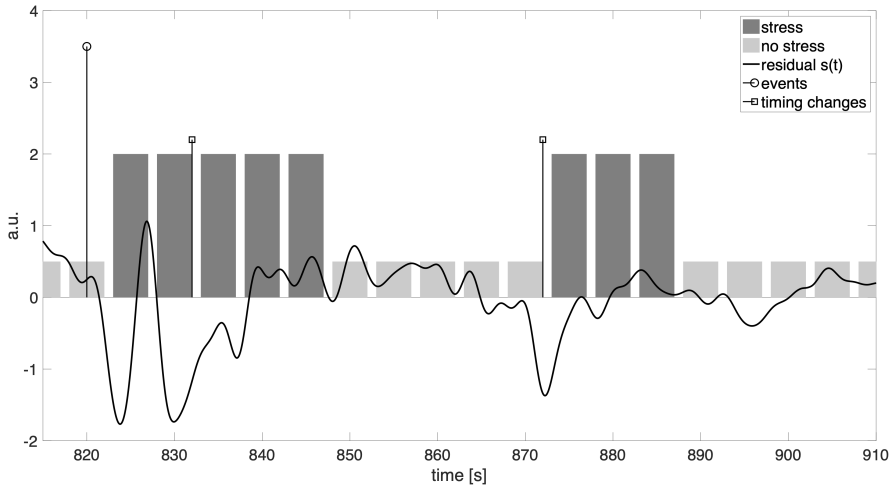


FIGURE 5.12: A visualization of the classification results for a tested subject. In this part of the acquired signal there are three different stress inducing episodes

Figure 5.12 shows the output labels of the classifier, in a 95 s long portion of the residual signal for one of the five tested subjects. A tall dark bar corresponds to stress detection, whereas a light short bar corresponds to non-stress detection. It can be noticed that the classifier detects the presence of stress in proximity of the stress triggers.

5.3 Feature extraction and classification 2

In this section we continue the activity described in 5.2, where we have tested SVM performance to classify the subject mental state. We set up a new experiment in collaboration with a company which specializes in driving simulator equipment. The experiment is a driving simulation running on a motorized platform which allows to recreate a more realistic scenario respect to the one described in 5.1. In addition to SPR, we measure and process Electrocardiogram (ECG) too, in order to have more features and improve the stress detection performance.

5.3.1 Sensor

Double channel SPR: The sensor architecture is similar to the one that we presented in Section 4.3.1 and is depicted in Figure 4.19; it is a dual-channel sensor that measures SPR from both the hands. In table 5.4 is a summary of sensor characteristics.

Triple channel ECG: The analog front end for ECG channels is basically a band-pass differential amplifier. The maximum input range for each channel is supposed to be $\pm 5\,\text{mV}$. The signals must be therefore amplified to obtain the output range $\pm 1.65\,\text{V}$ with respect to V_{REF}. The bandwidth of each channel is $[0.03, 160]$ Hz. The input impedance of each channel is $100\,\text{M}\Omega$, in order to reduce the load error to less than 1%. Finally, accordingly to the A/D specifications, the resolution of the ECG signals acquired on the skin results in the order of 2.5 μV.

TABLE 5.4: (a) Scheme of the developed system for EDA measurement; (b) block diagram of the dual channel SPR sensor

Electrode positioning	Electrode 1: Hand Palm (sweat glands present) Electrode 2: Hand Back (sweat glands not present) Electrode 3: Wrist (for voltage reference)
Battery	Lithium Polymer
V_{REF}	$1.65V$
Input impedance	100 MΩ
Gain	160
Range	± 10 mV
Frequency Band	[0.08 40] Hz
Sampling Frequency	200 Hz
Transmission mode	wi-fi
ADC	12 bit

5.3.2 Motion Artifact removal

Motion artifact removal technique is the one described in Section 4.3.3. Being $s_1(t)$ and $s_2(t)$ the acquired SPR from left and right hand respectively, the output of the motion Artifact removal block is:

$$out(k) = \alpha(k) \cdot s_2(k) + [1 - \alpha(k)] \cdot s_1(k) \tag{5.27}$$

$$\alpha(k) = \begin{cases} \frac{1}{1+e^{-2(\frac{\sigma_1(k)}{\sigma_2(k)}-1)}} & \text{if } \sigma_2(k) \neq 0 \\ 1 & \text{if } \sigma_2(k) = 0 \end{cases} \tag{5.28}$$

$\sigma_1(k)$ and $\sigma_2(k)$ are the standard deviations of the least N samples of $s_1(k)$ and $s_2(k)$ respectively.

5.3.3 Feature extraction and classification

In the proposed system we use two different supervised learning algorithms for classification: a Support Vector Machine and an Artificial Neural Network. We want to classify each time interval in one of the two possible classes: with stress (or "1") and

without stress (or "0"). To do so, we consider the cleaned SPR signal S, resulting from the MA removal block HR and RR obtained from ECG.

The magnitude of the SPR signal, subsampled at 100 Sa/s, may be very different from one subject to another, and a normalizing preprocessing phase is necessary. We consider, for each sample, the SPR signal in the previous 5 minutes. We then calculate, for each subject, the mean and standard deviation of all the SPR values included in this interval, and finally we standardize each value of the SPR, by subtracting the computed mean value and dividing the result by the standard deviation.
This normalization procedure is the result of the evaluation of different other normalization algorithms, and different durations of the intervals to consider, preceding the one analyzed, have been tested as well. The chosen method allows us to obtain a signal with approximately zero mean and a standard deviation equal to 1. The 5 minutes long interval is a good compromise in relation to the amount of data to process and the time window length needed to apply the algorithm.
From the SPR signal we consider five statistical features from each time interval signal block, which we chose to be 15 seconds long. Blocks are overlapping by 10 s, thus considering a new block every 5 s. This allows detection with an acceptable low delay.

In ECG signal, we detect R-peak locations with the Pan-Tompkins algorithm [80]. The instantaneous HR signal is extrapolated from RR intervals at a sample rate of 100 Sa/s. The HR signal is again normalized using the same procedure of the SPR signal. We consider eight statistical features for each time interval signal block.

Features were chosen to represent the stress signal component characteristics. In particular, regarding the SPR signal, we consider the block variance, the energy, the mean absolute value, the mean absolute derivative and the max absolute derivative. Kurtosis and PAPR were not chosen, since they did not improve

accuracy, contrary to what observed in the experiment described in Section 5.2. In this more realistic case, the SPR signal presents some fluctuations even when no stress events occur, as evidenced by Figure 5.14. Regarding the ECG signal, we compute the mean value of normal-to-normal RR (also known as NN) intervals, standard deviation of RR intervals (SDNN), standard deviation of subsequent RR interval differences (SDSD), root mean square of subsequent RR interval differences (RMSSD), number of subsequent RR intervals differing more than 50 ms (NN50) and the corresponding relative value in percentage (PNN50), mean value of the Heart Rate (HR) and HR mean derivative value (in Table 5.5 we summarize the considered features). We also took into account frequency domain features, in particular Low Frequency (LF) and High Frequency (HF) power spectra and the ratio LF/HF, but they were not selected for classification since they did not improve performance.

Input	Extracted features
SPR	**Variance** **Energy** **Mean absolute value** **Mean absolute derivative** **Max absolute derivative**
ECG	**RR mean** **SDNN** SDSD **RMSSD** NN50 PNN50 **HR mean** HR mean derivative

TABLE 5.5: Features used in our classification algorithms. In bold we highlighted the features combination which gives the best results, for both classifiers.

Feature vectors are obtained by combining the features of both signals, and each feature is then normalized in the [0, 1] range. The SVM classifier has been set up using the Matlab routine functions (Matlab 2017.a) with a Radial Basis Function (RBF) kernel and the Sequential minimal optimization (SMO) solver. The RBF hyperparameters, box constraint and kernel scale, have

been set up through the Bayesian optimization procedure which has also been applied during the training procedure.

The Artificial Neural Network has been setup using Python, with the Keras library and the optimization package [16] which allow us to apply the Bayesian optimization procedure as well. The ANN has one input layer, two hidden layers and one output layer. The input layer has a number of nodes equal to the number of selected features. The number of nodes in the first and second hidden layer ranges from 16 to 128. To choose the number of nodes in both hidden layers we run the optimization procedure. It works by comparing all the nodes combinations and selecting the best number of total nodes to use for each layer. Finally, the output layer has a single node with a binary output, "1" for stress and "0" for non-stress situations. Thanks to the optimization package, we could also optimize the "batch-size", the number of samples in each subset considered during the cross-validation training procedure. We choose it in order to obtain the same number of subsets as in the cross validation procedure used in the SVM. We also let the optimization procedure set the "drop-out" percentage, that corresponds to the number of nodes discarded during the classification process.

The "leave-one-person-out" method has been used for both classifiers. In particular, we train each classifier using data of all subjects, leaving one out, which is the one the classifier is tested on (similarly to what is done in 5.1). This is repeated leaving out each one of the subjects. Performance is calculated by averaging the test results of all subjects.

Considering that the total numbers of features is relatively low (5 extracted from the SPR signal and 8 from the ECG signal) we try all the combinations of these features and select the one which provided the best accuracy. This combination happens to be the same for both classifiers, and comprises nine features (indicated in bold in Table 5.5). The following section introduces the experimental setup. Section 5.3.5 then presents the experimental results.

5.3.4 Signal Processing

Tests are carried out in collaboration with a company which designs driving simulators, both in hardware (static and moving cockpits) and in software. We use a dynamic driving simulator, in which the subject is sitting inside a reproduced car cockpit, which allows to drive and to have physical feedback of car motion, acceleration, and of some external events. 18 non-professional drivers are tested, 14 men and 4 women, with age in the range 19-35 years, coming from the University of Udine and the University of Padua.

Before sitting on the simulator, each subject has to wear the sensor-equipped vest, which measures the ECG, and the double sensor which measures the SPR signal from each hand (as introduced in Section 5.3.1). In particular, in order to minimize the effects of bumps of the hand on the steering wheel (which may give rise to signal artifacts), we positioned the palm electrode on the *Abductor pollicis brevis* muscle. The driving test consists in a 67 km long track, simulating a straight highway, where 12 different stress-inducing events happen, spaced at different locations. The total run time depends on the speed each subject drives, but is roughly 40 minutes long because the subjects are instructed to keep a constant velocity between 120 km/h and 130 km/h. Each subject is also instructed to drive freely and in a realistic way. The 12 events are the ones shown in Figure 5.13: Double lane change (right to left or left to right), Tire labyrinth, Sponsor block (from left or from right), Slalom (from left or from right), Lateral Wind (from left or from right), Jersey LR, Tire trap, Stop.

In Figure 5.14 we show the full cleaned SPR signal, obtained after motion artifact removal from s_1 and s_2, and the HR signal of one subject, during the test. The vertical lines with a circle marker denote the starting points of the stress events (for obstacles, it is set at the moment they become visible, at a distance of 800 meters). The vertical lines with a squared marker denote the end points (40 seconds after the obstacle). The typical stress peaks in S signal are distinctly noticeable. Also, the characteristic rise in the Heart Rate during the stress episodes is clearly

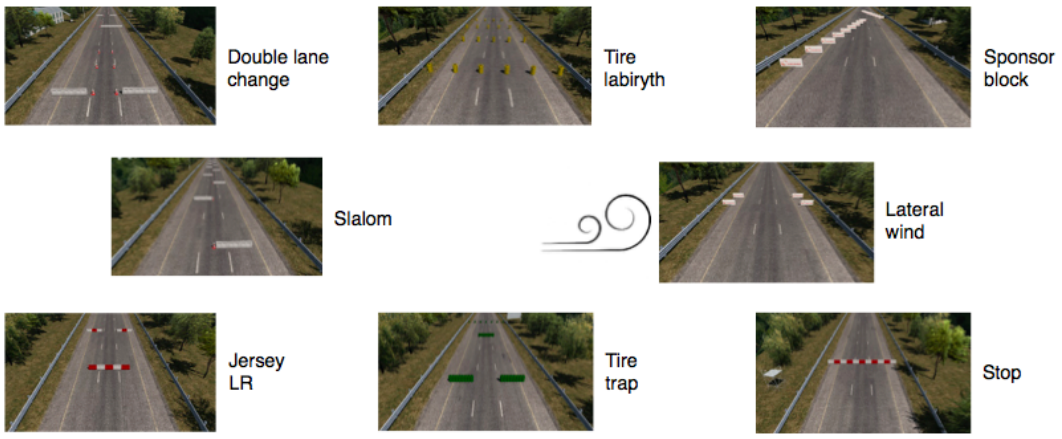


FIGURE 5.13: Stress events used in the simulation.

evident.

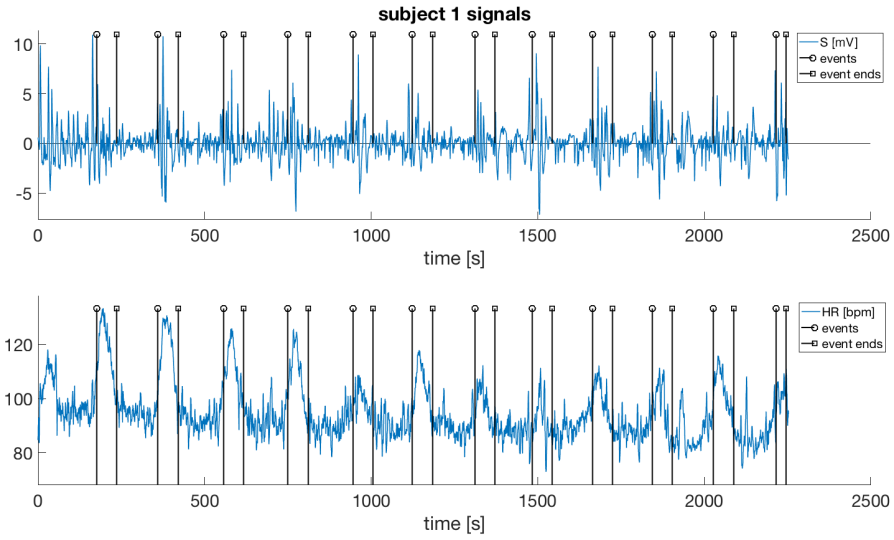


FIGURE 5.14: Stress (top) and HR (bottom) signals. The circled and the squared stems represent the beginning and the ending of stress events

5.3.5 Experimental results

As explained above, SPR and ECG features calculated from 15 s long blocks are used as the input to the classifier. We pick a new interval every 5 seconds, so there is an overlap between the intervals we consider. As explained before, an event begins when

an obstacle becomes visible and ends 40 seconds after the obstacle. So we know when and for how long the stress is supposed to happen in a subject. If an interval falls outside of these events, we then consider it to be "0", without stress, and if an interval, even partially, intersects a moment of stress, we consider it to be "1", with stress.

As introduced previously in Section 5.3.3, for classification we use the “leave-one-person-out" procedure, i.e., for the training process we use the data of all subjects, leaving one subject out, which is the one which the classifier is tested on. The subjects can perform the assigned tasks at different speeds and thus elapsed times, so each subject will have a different total number of stress intervals. For the training process we thus pick a number of stress intervals which can differ from subject to subject (variable from 148 to 169). We then extract, for each subject, a number of non-stress intervals equal to the number of the stress intervals we picked for that subject before, and we select these non-stress intervals randomly among the bigger pool of them (because our test track produced more non-stress intervals than stress intervals).

Therefore, for training and for each subject, we use a dataset in the range of [148 to 169] x 17 subjects = [2516 to 2873] intervals for each case, stress and non-stress. For testing, instead, the number of intervals to consider for each subject is 432, on average, with the majority (about two thirds) being non-stress ones. However, in calculating the performance indicators (that will be explained shortly in this section) we account for this unbalanced test dataset.

In the case of SVM classification, we use a 10-fold cross validation, and in the case of Neural Network classification we choose a specific "batch-size" in order to obtain 10 subsets in the cross validation procedure. In order to have 10 subsets, the batch size value must be approximately equal to 240. As introduced before, we use the optimal features combination for both SVM and

ANN classifiers, and to derive the performance measures we average the results for all the subjects.

We are able to count the True Positives (TP), which are all the test intervals which the classifier correctly labeled as “stress"; the False Negatives (FN), which are all the “non-stress" labeled intervals which instead we expected as “stress"; the True Negatives (TN), which are all the correctly detected “non-stress" intervals; and the False Positives (FP), which are all the incorrectly labeled “stress" intervals that instead we expected as “non-stress". We then compute the performance in terms of Accuracy, Sensitivity (also known as True positive rate, TPR) and Specificity (also known as True negative rate, TNR), defined as follows:

$$\text{Accuracy (\%)} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \cdot 100 \tag{5.29}$$

$$\text{Sensitivity (\%)} = \frac{\text{TP}}{\text{TP+FN}} \cdot 100 \tag{5.30}$$

$$\text{Specificity (\%)} = \frac{\text{TN}}{\text{FP+TN}} \cdot 100 \tag{5.31}$$

The accuracy expresses how effective our classifier is in general, computing how many correctly labeled stress and non-stress events are present among the total number of events. The sensitivity represents how effective our classifier is in correctly labeling stress events to the total number of actual stress events. The specificity represents how effective our classifier is in correctly labeling non-stress events to the total number of actual non-stress events. We also compute the Balanced Accuracy (BA), which is the arithmetic mean of sensitivity and specificity, and the Geometric Mean (GM) of sensitivity and specificity:

$$\text{BA (\%)} = \frac{1}{2}\left(\frac{\text{TP}}{\text{TP+FN}} + \frac{\text{TN}}{\text{FP+TN}}\right) \cdot 100 \tag{5.32}$$

$$\text{GM (\%)} = \sqrt{\frac{\text{TP}}{\text{TP+FN}} \cdot \frac{\text{TN}}{\text{FP+TN}}} \cdot 100 \tag{5.33}$$

These two additional indicators are significant when the test data are unbalanced like in our case.

	Acc (%)	Sens (%)	Spec (%)	Balanced Acc (%)	GM (%)
SVM	77.49	71.67	81.78	76.72	76.40
ANN	79.10	65.59	88.71	77.15	76.04

TABLE 5.6: SVM and Artificial Neural Network performance obtained using the optimal combination of features for both of them.

In Table 5.6 we show the values of these parameters obtained in our experiment. It can be seen that the accuracy and balanced accuracy are comparable between the SVM and ANN classifiers, with a slight advantage for the ANN. SVM sensitivity is higher than the ANN one, but SVM specificity is lower than the ANN one. The GM values are however very similar. The overall performance is fairly good, even if our hypotheses may have limited it. The first hypothesis is that we suppose that all the time intervals within a stress event are positive, while it may happen that a subject is not continuously stressed during the event. The second one is that a person should not be stressed outside of a stress event. This is not always true, because it may happen that a person becomes stressed at any time for reasons that we do not control.

To take these possibilities into account, in an attempt to detect stress more reliably, we apply a re-labelling procedure and consider the stress as present if we detect at least 4 consecutive positive stress labels. If there are less than 4 consecutive labels detected as stress we consider the stress to be absent (see Figure 5.15 that shows how we re-label the classifier's output).

In Table 5.7 we show the performance under these new assumptions, using the same indicators, which are slightly better than before. However, the sensitivity value is slightly worse (for both SVM and ANN classifiers), because we noticed that some of the correctly detected stress intervals are ruled out, since they were not consequential in groups of four or more.

In order to assess the capability of the proposed system to identify the stress inducing events, instead of single 15 s intervals,

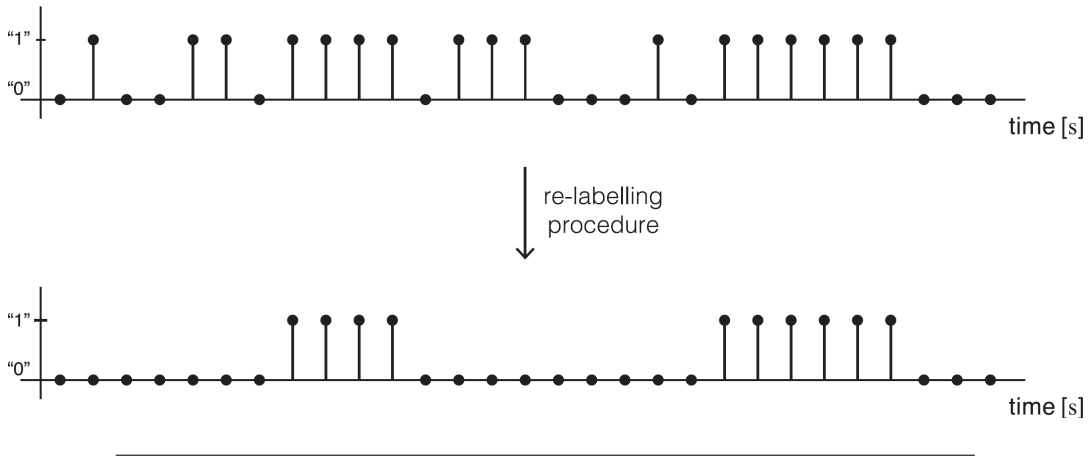


FIGURE 5.15: Relabeling procedure.

	Acc (%)	Sens (%)	Spec (%)	Balanced Acc (%)	GM (%)
SVM	80.21	69.55	87.93	78.74	78.04
ANN	80.78	63.32	93.19	78.26	76.49

TABLE 5.7: SVM and Artificial Neural Network performance after the re-labelling procedure.

we segment the signals into "stress" (starting 800 m before the obstacle and ending 40 s after) and "non-stress" blocks. For each subject there are therefore 12 stress blocks (corresponding to the 12 stress events) and 23 non-stress blocks. These are 23 because the time length of the non-stress blocks was longer than the stress ones, so we decided to divide them in half, making them roughly the same length. A block is marked as "stress" if we find inside it at least 4 consecutive positively classified intervals.

The performances under these assumptions, obtained averaging the results for all the subjects, are reported in Table 5.8. The accuracy and balanced accuracy are very high for both the SVM and ANN classifiers, with an advantage for the ANN. The SVM sensitivity is higher than that of the ANN, but the SVM specificity is lower than the ANN one, as in the two previous cases. The overall performance of this third case is very high, our hypothesis being basically that a stress detection in a block is considered positive if the subject is continuously stressed for at least 30 s, corresponding to 4 overlapping 15 s intervals, with 5 s displacement.

	Acc (%)	Sens (%)	Spec (%)	Balanced Acc (%)	GM (%)
SVM	87.62	95.83	83.33	89.58	89.21
ANN	91.27	93.98	89.86	91.92	91.73

TABLE 5.8: SVM and Artificial Neural Network performance considering stress and non-stress blocks (instead of single intervals).

5.4 Feature extraction and classification 3

Further driving simulations have been carried on, to test the classifier performance in a slightly different car driving context respect to the one in Section 5.3.

5.4.1 System, instrumentation and methods

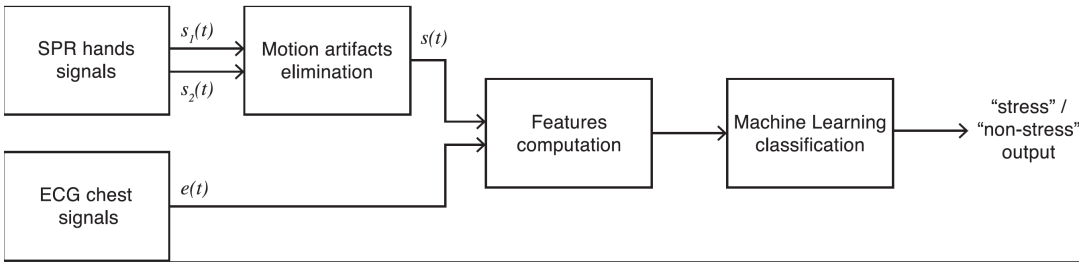


FIGURE 5.16: Scheme of the proposed system.

The scheme of our system is depicted in Figure 5.16. It is essentially the same of 5.3, two Skin Potential Response signals are given as input along with the Electrocardiogram signal. The SPR signals are then cleaned from the motion artifact and the result is a single combined SPR signal. Finally, through the use of a classifier, the recognition of stress in a given time interval is given as output.

The structure of the sensors are much the same of the ones used in 5.3; the Double channel SPR sensor and the triple channel ECG unit described in Section 5.2.1.

Also the Motion Artifact removal algorithm as well as feature extraction method and classifiers are the same of Section 5.3, described in Section 5.2.2, Section 5.2.3, Section 5.2.6 respectively.

5.4.2 Experimental setup

The tests are conducted in a company specialized in driving simulators. A professional dynamic driving simulator is utilized, which moves accordingly to the car movement. A total of 16 healthy individuals took part in the test, with an age in the 22-47 range, some attending the University of Udine and the University of Padua. They gave permission to have their physiological signals logged, and the principles of the Declaration of Helsinki were also respected during the tests.

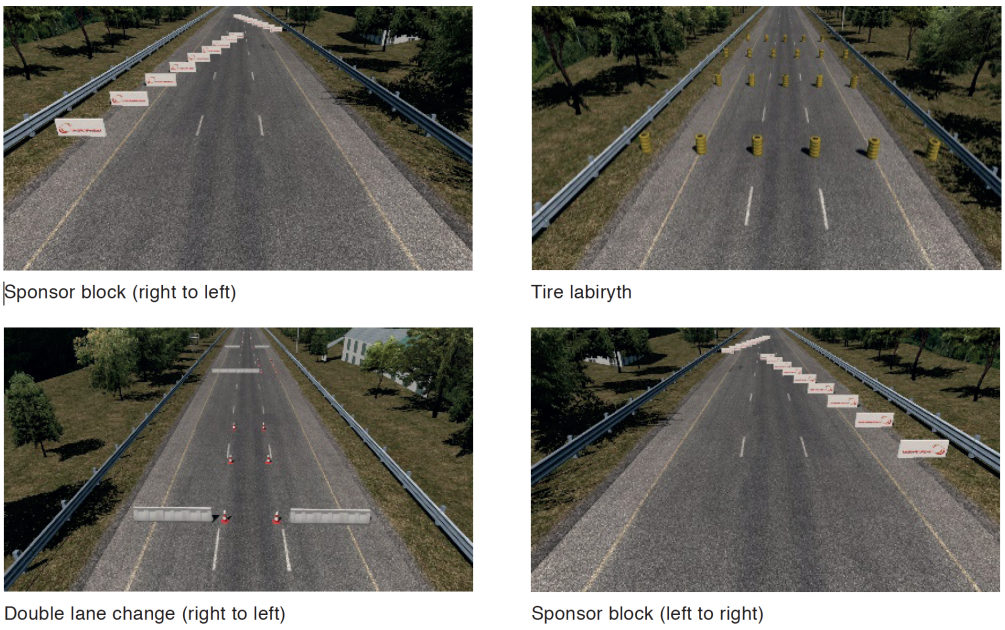


FIGURE 5.17: The obstacles placed along the track.

Each subject has to be prepared beforehand, wearing the sensors on the chest and on the hands. The experiment took place in a simulated 28 km long road, reproducing a highway, with four obstacles, requiring some effort to cross, positioned at different distances along the road. Figure 5.17 shows the four obstacles, which are: Sponsor block (from right to left), Tire labyrinth, Double lane change (right to left) and Sponsor block (from left to right).

Figure 5.18 shows, for a tested individual, the SPR signal once the motion artifact is removed, and the HR signal. The circle and square markers indicate respectively the onset and offset of

each stress event (the start is set when the individual sees the obstacle, at 800 meters distance, and the end is set 30 seconds after the obstacle is surpassed). In the figure the representative stress spikes in the SPR signal are clearly visible, as well as the typical increase in Heart Rate during the stress event.

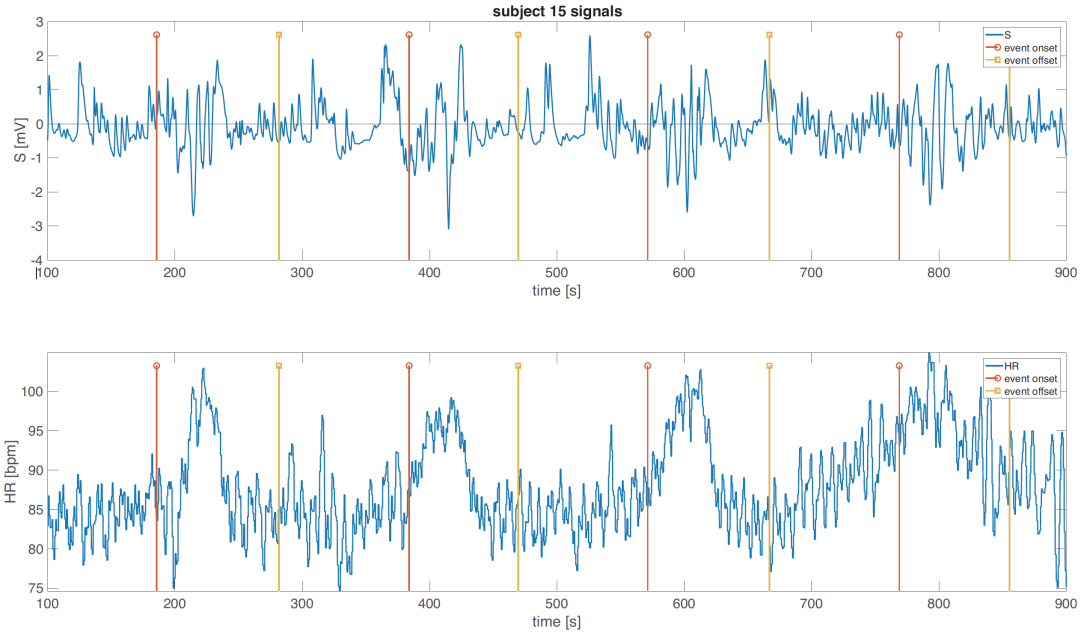


FIGURE 5.18: SPR (top) and HR (bottom) signals for an individual during the drive.

5.4.3 Experimental results

Both the cleaned SPR and the ECG signals of each subject are used to construct the classifier for stress detection in each 15 s time interval. The signal features are extracted in each of the intervals, which, as mentioned, overlap by 10 s.

All the intervals belonging to a road section with obstacles are supposed to be "1", with stress, and all the intervals belonging to a section without obstacles are supposed to be "0", without stress. As explained before, for the classification process, the "leave-one-person-out" method is used. Therefore, during the training phase, we consider 15 subjects at a time. This generates a number of intervals variable between 1176 and 1209, for both stress and non-stress classes, from which we extract the optimal combination of features listed above. The number of intervals in

each training phase is not constant because of the fact that each individual crosses the obstacles with a different speed/time. In the test phase, for each individual, we have approximately 187 intervals to consider (which is an average number considering all the subjects), where most of them (two thirds roughly) belong to the non-stress class.

Another aspect about the classification pertains the cross validation methodology. The 10-fold cross validation is used for the SVM, while for the ANN classifier we select the "batch-size" in order to get a matching number of subsets in the cross validation phase.

We can calculate the True Positives (TP), the False Negatives (FN), the True Negatives (TN) and the False Positives (FP), and we can therefore calculate the performance in terms of Accuracy, Sensitivity and Specificity, as defined in Equations 5.29, 5.30 and 5.31.

Since our test data is biased (more non-stress intervals than stress ones), two parameters are also significant, namely, the Balanced Accuracy (BA) and the Geometric Mean (GM), as in

$$\text{BA (\%)} = \frac{1}{2}\left(\frac{\text{TP}}{\text{TP+FN}} + \frac{\text{TN}}{\text{FP+TN}}\right) \cdot 100 \tag{5.34}$$

$$\text{GM (\%)} = \sqrt{\frac{\text{TP}}{\text{TP+FN}} \cdot \frac{\text{TN}}{\text{FP+TN}}} \cdot 100 \tag{5.35}$$

Table 5.9 shows the overall performance of our system (MEAN $\pm$ STD, computed from the results of the 16 subjects). The accuracy and balanced accuracy are similar in the SVM and ANN classifiers, with the ANN slightly better in both of them. Sensitivity of the SVM is greater than that of the ANN, and specificity in SVM is worse than in the ANN. The GM values are comparable in both SVM and ANN.

TABLE 5.9: Performance Comparison Between SVM and ANN

	Acc %	Sens %	Spec %	Balanced Acc %	GM %
SVM	72.86 $\pm$ 8.77	86.69 $\pm$ 7.79	66.46 $\pm$ 13.02	76.57 $\pm$ 6.91	75.40 $\pm$ 7.85
ANN	74.45 $\pm$ 8.65	86.19 $\pm$ 8.13	69.00 $\pm$ 13.38	77.59 $\pm$ 6.50	76.56 $\pm$ 7.30

The hypothesis are the same of Section 5.3.We assume that all the time intervals falling inside an obstacle crossing event should evidence stress and that a person is not stressed outside of an obstacle crossing event. This may not be true, since an individual can be stressed in those time intervals for reasons which we cannot control.

To consider these added occurrences, and recognize stress more reliably, we use the re-label method explained in Section 5.3.5 to the classifiers' output.

Table 5.10 shows the performance using the re-label procedure. The parameters are slightly better than before. As in the previous experiment, Sensitivity is a little lower for both classifiers.

TABLE 5.10: Performance Comparison Between SVM and ANN (the Classifier's Output Is Re-labelled)

	Acc %	Sens %	Specificity %	Balanced Acc %	GM %
SVM	77.61 ± 8.14	85.11 ± 8.82	74.05 ± 11.67	79.58 ± 6.89	79.00 ± 7.30
ANN	78.17 ± 9.13	84.73 ± 9.72	75.16 ± 14.23	79.94 ± 6.88	79.15 ± 7.50

The results of the experiment presented in this section, trace the results obtained Section 5.3. The only difference is that we have not segmented data in "stress" and "non stress" macro-blocks as in the previous test due to the small number of stressing events (4 per subject vs the 12 per subjects of experiment in Section 5.3).

5.5 Real time implementation

Having completed the classifiers training beforehand, classification can be computed with reasonable computational complexity in real time. As described before, we consider 15 seconds long time windows, shifting 5 seconds at a time for the next computation. The time available for each block classification, before we need to consider the next batch of data, is therefore 5 s. We might consider a concrete real time implementation as a future evolution of the system. The normalizing preprocessing step used in Section 4.3 to overcome the problem of possibly different SPR signal amplitudes among different subjects, can also be implemented in real time and can be applied before using

adaptive filters too. Indeed, we need to buffer the past samples in a 5 minute long sliding window and complete the normalization within 1/100 s. Initialization requires to wait 5 minutes to start the procedure.

5.6 Conclusions

In this chapter, we proposed a system for drivers' stress detection. In Section 5.1, we tested the system on data taken in laboratory experiments. Since the tests have been carried out in a controlled environment and stress has been induced through sympathetic stimulations we could rely on a strong ground truth. This allowed us to obtain a set of features whose classes ("stress" and "no stress") are well distinct as we can see from the test performances. We obtained a Recall value of 98.18% considering just the sound triggers, and of 91.77% considering both the sound triggers and the metronome changes. Having reliable data simplify the tests to try different kind of settings and find the way to perfect the system choosing the algorithm and the hyperparameters that give the best performances.

In Section 5.3 and Section 5.4 we carried on more realistic experiments with the aid of a professional simulator which runs on a motorized platform. The results show that, using both SPR and ECG signals, we can recognize stress episodes with good reliability. Also, the experiments evidenced some difficulties of realistic simulations due to the difficulty of design a good ground truth. Besides having to set the experiment to have a calm or stressed driver in certain time lapses; the issues with classifying stress in a prolonged period instead of detect a sudden transitory reaction led us to develop a relabeling procedure and to evaluate the stress considering macro intervals. Hence, an idea for improving the system efficacy, in addition to the implementation of more advanced classifiers and the employing of further signals and features, could be the development of a proper post processing method to take the driving context into account.

Chapter 6

Conclusions

The aim of the present work is to bring a contribution to the field of stress detection systems through the processing of physiological measurements; specifically, we focused on the stress that affects a person during driving activity. The physiological measurements recorded in our work are based on Skin Potential Response (SPR) and Electrocardiogram (ECG).

First, we focused on SPR, an endosomatic Electrodermal Activity (EDA) measure, which is less commonly used for the development of stress recognition systems. Our idea is to exploit the properties of SPR which has a faster response to stimuli than the more common Skin Conductance Response (SCR).

Then, we considered the issue of Motion Artifact which is especially strong for physiological measurements recorded from the hands, like EDA, and prevents any reliable evaluation from the signals. We developed a system to remove the Motion Artifact from SPR signals based on adaptive filters. The system removes Motion Artifact disturbances which are related to hand movements, which are represented by the Steering Wheel (SW) angle excursion. We tested the system on measurements taken in experiments carried out in laboratory, in which the subjects move the Steering Wheel while they are under stress stimulation. We applied the Motion Artifact removal system, and then we highlighted the stress response peaks with a Smoothed Nonlinear Energy Operator (SNEO). The results evidenced that we were able to recognize stress reactions.

In additional tests we considered the problem of removing the Motion Artifact not related to SW handling, such as movements

when holding off the SW or during muscle stretching. Thus, we developed a system for Motion Artifact removal which, instead of relying on hand movements, is based on the processing of two SPR signals measured on both the hands with a double channel SPR sensor. The system processes the two SPR through an algorithm that selects the smoother signal and outputs the stress component of SPR, which is the same for both the signals. Some errors could occur when the driver stretches the hands at the same time, but this kind of situations are rare in a driving context. To test the system, we carried out several experiments. The results showed that the system is able to remove the Motion Artifact and outperforms Principal Component Analysis and Independent Component Analysis algorithms.

Once we were able to get a signal cleaned from the Motion Artifact disturbances, the activity continued through the employing of machine learning techniques to recognize stress state in drivers. We carried out realistic experiments on a motorized platform, in collaboration with a company specialized on professional simulators, VI-Grade S.r.l., Tavagnacco (UD) - Italy. In addition to SPR, we measured ECG signal, then, we extracted several features from SPR and ECG. We used Support Vector Machine (SVM) and Neural Networks (NN) classifiers to detect stress. Analyzing the excursion of the features during the experiments and testing several combinations of features which are given as input to the classifiers, we found the set of features which give the best performance and were able to detect mental stress caused by the obstacles that the subjects had to avoid.

Future work will involve an in-depth analysis during car driving, focusing on SPR feature extraction. While the studies on the efficacy of ECG and SCR features are in an advanced stage, the studies on SPR features are very scarce. Also, the use of other types of classifiers will be carried out. The use of Recurrent Neural Networks, for example, will allow to classify data taking into account the excursion of the features over time. A decisive improvement for the performances would be probably obtained by raising the complexity of the Neural Network. Signal processing by means of Deep Learning systems, which are

not feature-based, will allow to perform the classification based on properties that would be hard to identify manually. Another perspective is represented by the stress recognition in subjects while they are in vehicles which have an autonomous driving system. As mentioned in the introduction, people are less likely to rely on automatic driving systems. Furthermore, autonomous driving systems still cannot handle every kind of situations and human intervention is often necessary. Thus, it is important to analyze the driver mental state in this context too. Experiments in autonomous driving contexts are being carried out along with the research activity.

Bibliography

[1] U Rajendra Acharya et al. "Heart rate variability: a review". In: *Medical and biological engineering and computing* 44.12 (2006), pp. 1031–1051.

[2] A Affanni. "Wearable instrument to measure simultaneously cardiac and electrodermal activities". In: *2016 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*. IEEE. 2016, pp. 1–5.

[3] A Affanni and G Chiorboli. "Design and characterization of a real-time, wearable, endosomatic electrodermal system". In: *Measurement* 75 (2015), pp. 111–121.

[4] A Affanni and G Chiorboli. "Wearable instrument for skin potential response analysis in AAL applications". In: *20th IMEKO TC4 International Symposium and 18th International Workshop on ADC Modelling and Testing Research on Electric and Electronic Measurement for the Economic Upturn Benevento*. 2014, pp. 807–811.

[5] Antonio Affanni et al. "Driver's stress detection using skin potential response signals". In: *Measurement* 122 (2018), pp. 264–274.

[6] Antonio Affanni et al. "Dual channel Electrodermal activity sensor for motion artifact removal in car drivers' stress detection". In: *2019 IEEE Sensors Applications Symposium (SAS)*. IEEE. 2019, pp. 1–6.

[7] Charu C Aggarwal. *Data classification: algorithms and applications*. CRC press, 2014.

[8] Jonathan Aigrain et al. "Multimodal stress detection from multiple assessments". In: *IEEE Transactions on Affective Computing* 9.4 (2016), pp. 491–506.

[9] Yair Amichai-Hamburger. *Technology and psychological well-being*. Cambridge University Press, 2009.

[10] PE An, B Brown, and Chris J Harris. "On the convergence rate performance of the normalized least-mean-square adaptation". In: *IEEE transactions on neural networks* 8.5 (1997), pp. 1211–1214.

[11] Ann-Lis Backman. "Health survey of professional drivers". In: *Scandinavian journal of work, environment & health* (1983), pp. 30–35.

[12] DS Bari et al. "Electrodermal responses to discrete stimuli measured by skin conductance, skin potential, and skin susceptance". In: *Skin Research and Technology* 24.1 (2018), pp. 108–116.

[13] Hanna Becker et al. "Emotion recognition based on high-resolution EEG recordings and reconstructed brain sources". In: *IEEE Transactions on Affective Computing* (2017).

[14] Asa Ben-Hur and Jason Weston. "A user's guide to support vector machines". In: *Data mining techniques for the life sciences*. Springer, 2010, pp. 223–239.

[15] Mathias Benedek and Christian Kaernbach. "Decomposition of skin conductance data by means of nonnegative deconvolution". In: *Psychophysiology* 47.4 (2010), pp. 647–658.

[16] James Bergstra, Dan Yamins, and David D Cox. "Hyperopt: A python library for optimizing the hyperparameters of machine learning algorithms". In: *Proceedings of the 12th Python in science conference*. Citeseer. 2013, pp. 13–20.

[17] N Bershad and O Macchi. "Comparison of RLS and LMS algorithms for tracking a chirped signal". In: *International Conference on Acoustics, Speech, and Signal Processing*, IEEE. 1989, pp. 896–899.

[18] Olga Vl Bitkina et al. "Identifying traffic context using driving stress: a longitudinal preliminary case study". In: *Sensors* 19.9 (2019), p. 2152.

[19] Esther Bosch et al. "Emotional GaRage: A Workshop on In-Car Emotion Recognition and Regulation". In: *Adjunct Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM. 2018, pp. 44–49.

[20] Bernhard E Boser, Isabelle M Guyon, and Vladimir N Vapnik. "A training algorithm for optimal margin classifiers". In: *Proceedings of the fifth annual workshop on Computational learning theory*. ACM. 1992, pp. 144–152.

[21] Wolfram Boucsein. *Electrodermal activity*. Springer Science & Business Media, 2012.

[22] Adrian Burns et al. "SHIMMER^{TM}–A wireless sensor platform for noninvasive biomedical research". In: *IEEE Sensors Journal* 10.9 (2010), pp. 1527–1534.

[23] Giancarlo Calvagno et al. "A multiresolution approach to spike detection in EEG". In: *2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No. 00CH37100)*. Vol. 6. IEEE. 2000, pp. 3582–3585.

[24] A John Camm et al. "Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology". In: (1996).

[25] Henry Candra et al. "Classification of facial-emotion expression in the application of psychotherapy using Viola-Jones and Edge-Histogram of Oriented Gradient". In: *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE. 2016, pp. 423–426.

[26] Theodora Chaspari et al. "Sparse representation of electrodermal activity with knowledge-driven dictionaries". In: *IEEE Transactions on Biomedical Engineering* 62.3 (2014), pp. 960–971.

[27] Tong Chen et al. "Detection of psychological stress using a hyperspectral imaging technique". In: *IEEE Transactions on Affective Computing* 5.4 (2014), pp. 391–405.

[28] Weixuan Chen et al. "Wavelet-based motion artifact removal for electrodermal activity". In: *2015 37th Annual International Conference of the*

IEEE Engineering in Medicine and Biology Society (EMBC). IEEE. 2015, pp. 6223–6226.

[29] Minho Choi et al. "Wearable device-based system to monitor a driverís stress, fatigue, and drowsiness". In: *IEEE Transactions on Instrumentation and Measurement* 67.3 (2017), pp. 634–645.

[30] Anuva Chowdhury et al. "Sensor applications and physiological features in driversí drowsiness detection: A review". In: *IEEE Sensors Journal* 18.8 (2018), pp. 3055–3067.

[31] Elisabeth Chroni et al. "The effect of stimulation technique on sympathetic skin responses in healthy subjects". In: *Clinical Autonomic Research* 16.6 (2006), pp. 396–400.

[32] Wan-Young Chung, Teak-Wei Chong, and Boon-Giin Lee. "Methods to detect and reduce driver stress: a review". In: *International Journal of Automotive Technology* 20.5 (2019), pp. 1051–1063.

[33] Jeffrey F Cohn, Zara Ambadar, and Paul Ekman. "Observer-based measurement of facial expression with the Facial Action Coding System". In: *The handbook of emotion elicitation and assessment* (2007), pp. 203–221.

[34] Giulia Da Poian, Riccardo Bernardini, and Roberto Rinaldo. "Gaussian dictionary for compressive sensing of the ECG signal". In: *2014 IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications (BIOMS) Proceedings*. IEEE. 2014, pp. 80–85.

[35] Ivan Nunes Da Silva et al. "Artificial neural networks". In: *Cham: Springer International Publishing* (2017).

[36] Paul Ekman et al. "Invited article: Face, voice, and body in detecting deceit". In: *Journal of nonverbal behavior* 15.2 (1991), pp. 125–135.

[37] PH Ellaway et al. "Sweat production and the sympathetic skin response: Improving the clinical assessment of autonomic function". In: *Autonomic Neuroscience* 155.1-2 (2010), pp. 109–114.

[38] Anna Feldhütter, Christoph Segler, and Klaus Bengler. "Does shifting between conditionally and partially automated driving lead to a loss of mode awareness?" In: *International conference on applied human factors and ergonomics*. Springer. 2017, pp. 730–741.

[39] Anna Feldhütter et al. "How the duration of automated driving influences take-over performance and gaze behavior". In: *Advances in ergonomic design of systems, products and processes*. Springer, 2017, pp. 309–318.

[40] Marshall L Fisher. "The Lagrangian relaxation method for solving integer programming problems". In: *Management science* 27.1 (1981), pp. 1–18.

[41] James S Fleming. *Interpreting the electrocardiogram*. Springer Science & Business Media, 2012.

[42] Don C Fowles et al. "Publication recommendations for electrodermal measurements". In: *Psychophysiology* 18.3 (1981), pp. 232–239.

[43] Stuart I Fox. "Human Physiology 8th edition Chapter 3 Multiple Choice 2". In: (2004).

[44] Christian Gold et al. "Taking over control from highly automated vehicles in complex traffic situations: the role of traffic density". In: *Human factors* 58.4 (2016), pp. 642–652.

[45] Gisele C Gotardi et al. "Adverse effects of anxiety on attentional control differ as a function of experience: a simulated driving study". In: *Applied ergonomics* 74 (2019), pp. 41–47.

[46] Alberto Greco et al. "Arousal and valence recognition of affective sounds based on electrodermal activity". In: *IEEE Sensors Journal* 17.3 (2016), pp. 716–725.

[47] Shalom Greene, Himanshu Thapliyal, and Allison Caban-Holt. "A Survey of Affective Computing for Stress Detection: Evaluating technologies in stress detection for better health". In: *IEEE Consumer Electronics Magazine* 5.4 (2016), pp. 44–56.

[48] Chang Hoon Ha et al. "Investigation on relationship between information flow rate and mental workload of accident diagnosis tasks in NPPs". In: *IEEE Transactions on Nuclear Science* 53.3 (2006), pp. 1450–1459.

[49] Zahid Halim and Mahma Rehan. "On identification of driving-induced stress using electroencephalogram signals: A framework based on wearable safety-critical scheme and machine learning". In: *Information Fusion* 53 (2020), pp. 66–79.

[50] Mustapha Harb et al. "Projecting travelers into a world of self-driving vehicles: estimating travel behavior implications via a naturalistic experiment". In: *Transportation* 45.6 (2018), pp. 1671–1685.

[51] Mohammad Mehedi Hassan et al. "Human emotion recognition using deep belief network architecture". In: *Information Fusion* 51 (2019), pp. 10–18.

[52] Corneliu-Eugen Havarneanu, Cornelia Măirean, and Simona-Andreea Popuşoi. "Workplace stress as predictor of risky driving behavior among taxi drivers. The role of job-related affective state and taxi driving experience". In: *Safety science* 111 (2019), pp. 264–270.

[53] Simon S Haykin. *Modern filters*. Macmillan Coll Division, 1989.

[54] Jennifer Healey and Rosalind Picard. "SmartCar: detecting driver stress". In: *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*. Vol. 4. IEEE. 2000, pp. 218–221.

[55] Alberto Hernando et al. "Inclusion of respiratory frequency information in heart rate variability analysis for stress assessment". In: *IEEE journal of biomedical and health informatics* 20.4 (2016), pp. 1016–1025.

[56] Eun-Hye Jang et al. "A study on analysis of bio-signals for basic emotions classification: recognition using machine learning algorithms". In: *2014 International Conference on Information Science & Applications (ICISA)*. IEEE. 2014, pp. 1–4.

[57] Eun-Hye Jang et al. "Analysis of physiological signals for recognition of boredom, pain, and surprise emotions". In: *Journal of physiological anthropology* 34.1 (2015), p. 25.

[58] Oliver Jarosch, Hanna Bellem, and Klaus Bengler. "Effects of Task-Induced Fatigue in Prolonged Conditional Automated Driving". In: *Human factors* (2019), p. 0018720818816226.

[59] J. D. Johnston. "Transform coding of audio signals using perceptual noise criteria". In: *IEEE Journal on Selected Areas in Communications* 6.2 (1988), pp. 314–323. ISSN: 1558-0008. DOI: 10.1109/49.608.

[60] Christos D Katsis et al. "Toward emotion recognition in car-racing drivers: A biosignal processing approach". In: *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* 38.3 (2008), pp. 502–512.

[61] Yılmaz Kaya and Ömer Faruk Ertuğrul. "Estimation of neurological status from non-electroencephalography bio-signals by motif patterns". In: *Applied Soft Computing* 83 (2019), p. 105609.

[62] Malia Kelsey et al. "Artifact detection in electrodermal activity using sparse recovery". In: *Compressive Sensing VI: From Diverse Modalities to Big Data Analytics*. Vol. 10211. International Society for Optics and Photonics. 2017, p. 102110D.

[63] Mohit Kumar et al. "Fuzzy evaluation of heart rate signals for mental stress assessment". In: *IEEE Transactions on fuzzy systems* 15.5 (2007), pp. 791–808.

[64] Antonio Lanatà et al. "How the autonomic nervous system and driving style change with incremental stressing conditions during simulated driving". In: *IEEE Transactions on Intelligent Transportation Systems* 16.3 (2014), pp. 1505–1517.

[65] Boon Giin Lee et al. "Wearable mobile-based emotional response-monitoring system for drivers". In: *IEEE Transactions on Human-Machine Systems* 47.5 (2017), pp. 636–649.

[66] Yuan-Pin Lin et al. "EEG-based emotion recognition in music listening". In: *IEEE Transactions on Biomedical Engineering* 57.7 (2010), pp. 1798–1806.

[67] Yuan-Pin Lin et al. "EEG-based emotion recognition in music listening: A comparison of schemes for multiclass support vector machine". In: *2009 IEEE international conference on acoustics, speech and signal processing*. IEEE. 2009, pp. 489–492.

[68] Monika Lohani, Brennan R Payne, and David L Strayer. "A review of psychophysiological measures to assess cognitive states in real-world driving". In: *Frontiers in human neuroscience* 13 (2019).

[69] Daniel Lopez-Martinez, Neska El-Haouij, and Rosalind Picard. "Detection of Real-world Driving-induced Affective State Using Physiological Signals and Multi-view Multi-task Machine Learning". In: *arXiv preprint arXiv:1907.09929* (2019).

[70] Choubeila Maaoui, Alain Pruski, and Faiza Abdat. "Emotion Recognition for hHman-Machine Communication". In: *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2008, pp. 1210–1215.

[71] Omae Manabu et al. "The application of RTK-GPS and steer-by-wire technology to the automatic driving of vehicles and an evaluation of driver behavior". In: *IATSS research* 30.2 (2006), pp. 29–38.

[72] Olvi L Mangasarian. "Arbitrary-norm separating plane". In: *Operations Research Letters* 24.1-2 (1999), pp. 15–23.

[73] Anthony D McDonald, Thomas K Ferris, and Tyler A Wiener. "Classification of Driver Distraction: A Comprehensive Analysis of Feature Generation, Machine Learning, and Input Measures". In: *Human factors* (2019), p. 0018720819856454.

[74] Daniel McDuff, Sarah Gontarek, and Rosalind W Picard. "Improvements in remote cardiopulmonary measurement using a five band digital camera". In: *IEEE Transactions on Biomedical Engineering* 61.10 (2014), pp. 2593–2601.

[75] Nadine Marie Moacdieh and Nadine Sarter. "The effects of data density, display organization, and stress on search performance: An eye tracking study of clutter". In: *IEEE Transactions on Human-Machine Systems* 47.6 (2017), pp. 886–895.

[76] Mohammadhossein Moghimi, Robert Stone, and Pia Rotshtein. "Affective Recognition in Dynamic and Interactive Virtual Environments". In: *IEEE Transactions on Affective Computing* (2017).

[77] JD Montagu and Edward Michael Coles. "Mechanism and measurement of the galvanic skin response." In: *Psychological Bulletin* 65.5 (1966), p. 261.

[78] Sudipta Mukhopadhyay and GC Ray. "A new interpretation of nonlinear energy operator and its efficacy in spike detection". In: *IEEE Transactions on biomedical engineering* 45.2 (1998), pp. 180–187.

[79] Andrew Ng. "Machine learning yearning". In: *URL: http://www. mlyearning. org/(96)* (2017).

[80] Jiapu Pan and Willis J Tompkins. "A real-time QRS detection algorithm". In: *IEEE Trans. Biomed. Eng* 32.3 (1985), pp. 230–236.

[81] Suja Sreeith Panicker and Prakasam Gayathri. "A survey of machine learning techniques in physiology based mental stress detection systems". In: *Biocybernetics and Biomedical Engineering* (2019).

[82] Corey Park, Shervin Shahrdar, and Mehrdad Nojoumian. "EEG-based classification of emotional state using an autonomous vehicle simulator". In: *2018 IEEE 10th Sensor Array and Multichannel Signal Processing Workshop (SAM)*. IEEE. 2018, pp. 297–300.

[83] Bo Peng and Gang Qian. "Online gesture spotting from visual hull data". In: *IEEE transactions on pattern analysis and machine intelligence* 33.6 (2010), pp. 1175–1188.

[84] Laura Pomarjanschi et al. "Safer driving with gaze guidance". In: *International Conference on Bio-Inspired Models of Network, Information, and Computing Systems*. Springer. 2010, pp. 581–586.

[85] William H Press and Saul A Teukolsky. "Savitzky-Golay smoothing filters". In: *Computers in Physics* 4.6 (1990), pp. 669–672.

[86] Emily B Prince, Katherine B Martin, and Daniel S Messinger. *Facial action coding system*.

[87] Bhanu Priya and S Dandapat. "A subspace projection based approach to improve the recognition of stressed speech". In: *2016 IEEE Annual India Conference (INDICON)*. IEEE. 2016, pp. 1–7.

[88] Zhaoyi Qin et al. "Stress level evaluation using BP Neural network based on time-frequency analysis of HRV". In: *2017 IEEE International Conference on Mechatronics and Automation (ICMA)*. IEEE. 2017, pp. 1798–1803.

[89] S Ramamohan and S Dandapat. "Sinusoidal model-based analysis and classification of stressed speech". In: *IEEE transactions on audio, speech, and language processing* 14.3 (2006), pp. 737–746.

[90] Mohammad Naim Rastgoo et al. "A critical review of proactive detection of driver stress levels based on multimodal measurements". In: *ACM Computing Surveys (CSUR)* 51.5 (2019), p. 88.

[91] Conny MA van Ravenswaaij-Arts et al. "Heart rate variability". In: *Annals of internal medicine* 118.6 (1993), pp. 436–447.

[92] George Rigas et al. "Towards driver's state recognition on real driving conditions". In: *International Journal of Vehicular Technology* 2011 (2011).

[93] Stephen Roberts and Richard Everson. *Independent component analysis: principles and practice*. Cambridge University Press, 2001.

[94] Jean-Claude Roy et al. *Progress in electrodermal research*. Vol. 249. Springer Science & Business Media, 2012.

[95] DE Rumelhart. "GE Hinton, and RJ Williams. 1986". In: *Learning Representations by Back-Propagating errors. Nature* 323 (), pp. 533–536.

[96] Ioana Safta, Ovidiu Grigore, and Constantin Căruntu. "Emotion detection using psycho-physiological signal processing". In: *2011 7th International Symposium on Advanced Topics in Electrical Engineering (ATEE)*. IEEE. 2011, pp. 1–4.

[97] Abraham Savitzky and Marcel JE Golay. "Smoothing and differentiation of data by simplified least squares procedures." In: *Analytical chemistry* 36.8 (1964), pp. 1627–1639.

[98] Ali H Sayed and Thomas Kailath. "Recursive least-squares adaptive filters". In: *The Digital Signal Processing Handbook* 21.1 (1998).

[99] Pedro Schestatsky et al. "Sympathetic skin responses evoked by muscle contraction". In: *Neuroscience letters* 463.2 (2009), pp. 140–144.

[100] Justus Schwan et al. "High-performance and lightweight real-time deep face emotion recognition". In: *2017 12th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP)*. IEEE. 2017, pp. 76–79.

[101] Jerritta Selvaraj et al. "Classification of emotional states from electrocardiogram signals: a non-linear approach based on hurst". In: *Biomedical engineering online* 12.1 (2013), p. 44.

[102] Cornelia Setz et al. "Discriminating stress from cognitive load using a wearable EDA device". In: *IEEE Transactions on information technology in biomedicine* 14.2 (2009), pp. 410–417.

[103] Jainendra Shukla et al. "Feature Extraction and Selection for Emotion Recognition from Electrodermal Activity". In: *IEEE Transactions on Affective Computing* (2019).

[104] Ramanathan Subramanian et al. "ASCERTAIN: Emotion and personality recognition using commercial sensors". In: *IEEE Transactions on Affective Computing* 9.2 (2016), pp. 147–160.

[105] Yu Sun and Nitish Thakor. "Photoplethysmography revisited: from contact to noncontact, from point to imaging". In: *IEEE Transactions on Biomedical Engineering* 63.3 (2015), pp. 463–477.

[106] Daisuke Tomoi et al. "Estimation of stress during car race with factor analysis". In: *2015 International Symposium on Micro-NanoMechatronics and Human Science (MHS)*. IEEE. 2015, pp. 1–5.

[107] Minoru Toyokura. "Sympathetic skin responses: the influence of electrical stimulus intensity and habituation on the waveform". In: *Clinical Autonomic Research* 16.2 (2006), pp. 130–135.

[108] Minoru Toyokura. "Waveform and habituation of sympathetic skin response". In: *Electroencephalography and Clinical Neurophysiology/Electromyography and Motor Control* 109.2 (1998), pp. 178–183.

[109] Minoru Toyokura. "Waveform variation and size of sympathetic skin response: regional difference between the sole and palm recordings". In: *Clinical neurophysiology* 110.4 (1999), pp. 765–771.

[110] Minoru Toyokura. "Within-subject consistency of sympathetic-skin-response waveform across different modalities of stimulation". In: *Autonomic Neuroscience* 169.2 (2012), pp. 135–138.

[111] Su-Jing Wang et al. "Micro-expression recognition using color spaces". In: *IEEE Transactions on Image Processing* 24.12 (2015), pp. 6034–6047.

[112] Ye Xia et al. "Predicting driver attention in critical situations". In: *Asian Conference on Computer Vision*. Springer. 2018, pp. 658–674.

[113] Qianli Xu, Tin Lay Nwe, and Cuntai Guan. "Cluster-based analysis for personalized stress evaluation using physiological signals". In: *IEEE journal of biomedical and health informatics* 19.1 (2014), pp. 275–281.

[114] Yimin Yang et al. "EEG-based emotion recognition using hierarchical network with subnetwork nodes". In: *IEEE Transactions on Cognitive and Developmental Systems* 10.2 (2017), pp. 408–419.

[115] Ya Jun Yu et al. "Investigation on driver stress utilizing ECG signals with on-board navigation systems in use". In: *2016 14th International Conference on Control, Automation, Robotics and Vision (ICARCV)*. IEEE. 2016, pp. 1–6.

[116] Shizhe Zang et al. "The Impact of Adverse Weather Conditions on Autonomous Vehicles: How Rain, Snow, Fog, and Hail Affect the Performance of a Self-Driving Car". In: *IEEE Vehicular Technology Magazine* 14.2 (2019), pp. 103–111.

[117] Jie Zhang et al. "Recognition of Real-Scene Stress in Examination with Heart Rate Features". In: *2017 9th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*. Vol. 1. IEEE. 2017, pp. 26–29.

[118] Yuning Zhang, Maysam Haghdan, and Kevin S Xu. "Unsupervised motion artifact detection in wrist-measured electrodermal activity data". In: *Proceedings of the 2017 ACM International Symposium on Wearable Computers*. ACM. 2017, pp. 54–57.

[119] Rencheng Zheng et al. "Biosignal analysis to assess mental stress in automatic driving of trucks: Palmar perspiration and masseter electromyography". In: *Sensors* 15.3 (2015), pp. 5136–5150.

[120] Yali Zheng et al. "Unobtrusive and multimodal wearable sensing to quantify anxiety". In: *IEEE Sensors Journal* 16.10 (2016), pp. 3689–3696.