Study project report

‘Advanced Signal Processing’

-

Glenmarvin ANTONYDAS  
Théo MARINI  
Guillaume SHERPA

Abstract

This work aims to process biosignal data as part of a research topic on the development of multimodal sensors embedded on a wearable device that will help to monitor patient health. Through a protocol designed to generate stress stimuli, four biosignals have been recorded : electrocardiogram, electrodermal activity, respiratory rate and accelerometer signal. In order to predict stress state, machine learning models have to be designed to interpret the characteristics of these signals. Before the processing of the data, pre-processing steps have to be done to get clean data. Effectively the recording comes with gaussian white noise and motion artefacts from the patient that could distort data metrics. Regarding the electrocardiogram, adaptive filtering and thresholds algorithms have been used to determine metrics such as tachogram or heart rate variability representative of the health state. The electrodermal activity data has been processed using filtering, stationary wavelet transform and resolution of convex optimization problem to extract response-to-stimuli peak metrics. Finally, the respiration rate signal has been filtered to then evaluate its correlation with the acceleration signal using windowed correlation and instantaneous phase synchrony.

*Ce travail a pour but de traiter des biosignaux dans le cadre d'un sujet de recherche sur le développement de capteurs multimodaux embarqués sur un vêtement intelligent qui permettra la surveillance de la santé des patients. Grâce à un protocole conçu pour générer des stimuli de stress, quatre biosignaux ont été enregistrés : l'électrocardiogramme, l'activité électrodermale, la fréquence respiratoire et un signal d'accéléromètre. Afin de prédire l'état de stress d’un patient, des modèles d'apprentissage automatique doivent être conçus pour interpréter les caractéristiques de ces derniers signaux. Avant le traitement des données par ces algorithmes, des étapes de prétraitement doivent être effectuées afin d’obtenir des données exploitables. En effet, l'enregistrement est accompagné de bruit blanc gaussien et d'artefacts de mouvement du patient qui peuvent fausser la mesure des données. En ce qui concerne l'électrocardiogramme, des algorithmes de filtrage adaptatif et de seuils ont été utilisés pour déterminer des métriques telles que le tachogramme ou la variabilité de la fréquence cardiaque, représentatifs de l'état de santé. Les données d'activité électrodermale ont été traitées par filtrage, transformation en ondelettes stationnaires et résolution d'un problème d'optimisation convexe afin d'extraire les pics de réponse aux stimuli. Enfin, le signal de fréquence respiratoire a été filtré pour ensuite évaluer sa corrélation avec le signal d'accélération en utilisant la corrélation par fenêtre et la synchronisation de phase instantanée.*

# Table of contents

[**Table of contents**](#_ofxkn5unc2ol) **2**

[**I. Introduction**](#_6oghqyy8bcpd) **3**

[**III. Experiment**](#_qfco1j8560n0) **4**

[A. Work organization](#_ez9qj3oalxiq) 4

[B. Recording protocol](#_r7yhp1kpxaj6) 4

[C. Data pre-processing](#_3lppdq3t7ecc) 5

[0. ICA attempt](#_3rwptadhhf8j) 5

[1. Electrocardiogram signal](#_ciojiiiq0n9h) 5

[2. Electrodermal activity signals](#_4sxf07ds0o81) 6

[3. Respiratory Rate and Accelerometer](#_6yvtqn94snf2) 7

[D. Data processing](#_3ejthz3mqchy) 7

[1. Respiration rate and acceleration correlation](#_nuifox1ksinz) 7

[**IV. Observations**](#_6oeqioir7yo9) **8**

[1. Electrocardiogram signal](#_pnv2k1any662) 8

[2. Electrodermal activity](#_2pwnlghk0hwx) 8

[3. Respiration rate and acceleration correlation](#_a1q84jy2edvb) 9

[**V. Conclusion**](#_6d7hv314fhk8) **9**

[**References**](#_9i2r3tfeiqx) **10**

# I. Introduction

This study project takes place as part of our Biomedical cursus at the Mines Saint Etienne. In this project we worked with Pr. Esma ISMAILOVA, Dr. Laura FERRARI, and Marina GALLIANI as part of Laura FERRARI works regarding the prediction of stress stage. It consisted of processing four biosignals recorded during a specific stress-generation protocol, to apply supervised machine learning models in order to classify stress stages.

Emotions can be analysed by the record of one’s biosignals such as temperature, cardiac or respiratory rhythm. Within the different emotions, the stress could be a crucial indicator to predict disease. Here we analysed electrocardiogram, electrodermal activity, respiration rate and accelerometer signals in order to identify major stress stages.

An electrocardiogram (ECG) represents the cardiac muscle activity of a subject. ECG could be decomposed as three parts: P wave, the QRS complex and the T wave. The QRS complex gives one’s heart rate that can be processed to inform us about one’s emotional state.

Electrodermal activity (EDA) represents the electrodermal activity of the skin. It is measured using conductance, and its signals are a manifestation of the activity of the sweat glands innervated by the autonomic nervous system. When the sudomotor nerves stimulate sweat production, during stress conditions for instance, the conductivity measured at the skin surface changes due to sweat secretion and changes in ionic permeability of the sweat gland membranes.

Respiration rate (RR) is the rate of an inspiration-expiration cycle during a certain amount of time. An increase of this rate could be correlated with an increase of the cardiac activity and indicate if one is in a stressful stage.

Accelerometer indicates one’s motion. This signal does not give explicit information about stress state, however it is used to remove motion artefacts present inside the other signals.

During this project the objectives were to understand and apply the biomedical notions behind the biosignals, to process the gathered data and finally to analyse them with machine learning algorithms for prediction and classification. One will first explain the thinking and algorithms used for the processing steps, then we will discuss the observations of the processing step output and conclude by our results.

# 

# III. Experiment

## A. Work organization

In this study we initially split the work into three parts for the three of us: the first part on ECG signal analysis, the second on EDA signal analysis and the last one on RR signal and accelerometer signals analysis. To organize our work, we applied the same method: we were looking for references on the treated subject, we applied the algorithms used in the previous references, we published our results within our discussion group to have a cross validation of the algorithms and then we presented it to the researchers.

As this project consists of the development of processing algorithms, we worked using Git as a version control software. We used python as a programming language to develop our functions and tried to apply good practices such as working with common libraries for the three of us. This choice has been motivated by the already numerous existing signal processing and machine learning libraries with associated tutorials.

## B. Recording protocol

A protocol for stress state generation has been applied to the subject in order to gather the data. Dr. FERRARI choose the protocol from [Nom des personnes qui ont mis au point le protocol] which is presented as the following:

* The subject is asked to put oneself in a comfortable position.
* The subject is asked to count backwards from 100 subtracting 3 for 30 seconds.
* The subject is asked to count backwards from 100 subtracting 7 for 30 seconds.
* The subject is asked to watch a relaxing video for 5 minutes.

Thus, we expect the subject to show stressful state during the stage of counting. As our main objective is to apply supervised machine learning on our dataset, L. FERRARI took the time labels for each step of the protocol. These labels also helped us to visualize the stress stimuli on our plots. Once the data are retrieved in a digital format, we can analyze them with programming.

## C. Data pre-processing

### 0. ICA attempt

Cleaning signal from noise and artefact has been the first step of each signal. The first idea was to use an Independent Component Analysis to extract the neat signal we are focusing on for both artefacts and noise. Inspired from its application on other biosignals such as ECG [1] or EEG, we performed the algorithms on the four signals of the experiment. A clear decomposition into the three pur signals with the motion artefact component was expected.

However, this idea was cursed from the beginning. Indeed, all signals recorded the same events but with different kinds of metrics (electric voltage, conductance, acceleration). Considering component extraction requires that all sensors are collecting data from the same nature i.e. within the same space of measurement. It applies for a dozen of the same electrodes on the chest for ECG but not in our experiment with four different types of sensors. Thus ICA was dismissed from the pre-processing technique options and we focused on more classic methods of filtering.

### 1. Electrocardiogram signal

Initially, Glen took on the task of processing the ECG data because it had a lot of unwanted elements : noise and artifacts. To do this, he had to rely on the article by Rahul Kher [2] which presents the main nuisances to the signals and the methods to remove them.

In this article, different filters were presented to obtain the best possible signal. The use of a bandpass filter with a bandwidth between 6 and 40 Hz would allow us to obtain a sufficiently filtered signal.

For the implementation of a code in Python, it has been done in several steps:

- The extraction of the data from an .h5 file

- Displaying the raw data in graphs

- Filtering the signals

- Displaying metrics (heart rate, power spectrum)

### 2. Electrodermal activity signals

To process the EDA signals we firstly decided to plot them in order to have a visual appearance of our dataset. After that, we observe the noisy data with unwanted artefacts such as white noise (maybe coming from the recording) and as motion artefacts (coming from a subject's movement). In order to have metrics we need to have clean data to process further with machine learning algorithms. Our next task then has been to search for filtering and artefact removing methods from the literature.

We found numerous filter examples for the EDA. From the paper of Kelsey et al. [] we decided to implement three low-pass filters as indicated, a function designed to remove high frequency components from a signal, at different order from one to eight with a respectively cutoff frequencies of 0.35 Hz, 0.5 Hz and 5Hz. We obtained results that happened to be incorrect as we remove a major part of the signal. To continue with a more scientific approach we decided to visualize the frequency content of the signal using the Fast Fourier Transform well-known algorithm. When we observed the content of the signals, we understood our mistake to choose a too small cutoff frequency.

Thus, with the advice of our tutors, we looked for tutorials on the website of the manufacturer of the probe used for data recording that would help us to go forward in the preprocessing step. According to their reference, we designed two other filters: a second order low-pass filter with a cutoff frequency at 35 Hz and a first order band-pass filter with a frequency band of 0.05-35Hz. By a visual comparison between these two, we decided to work with the first one that also brought convenient results with every EDA signal. Now that we removed the noisy components, we processed the motion artefacts by using an algorithm developed by Chen et al. using Stationary Wavelet Transform. After we reconstructed the signal and smoothed it, we obtained a processable signal for eight of the eleven given signals in the datasets.

Furthermore, once we had clean signals, we were able to extract metrics that would allow us to determine the data characteristics to feed our future machine learning algorithm. We applied an algorithm developed by Greco et al that decomposes the input signal into three components : the phasic, the tonic and a white noise. The tonic component explains the slow drifts of the baseline skin conductance level and spontaneous ﬂuctuations in skin conductance. The phasic component explains the skin conductance response and reﬂects the short-time response to the stimulus. The creators of this algorithm consider four hypotheses that they traduct into a mathematical problem of optimization, whose resolution allows us to extract these three components.

In our case, as we want to analyze the data regarding the short stress stimuli, we were more interested in the phasic component. As input of this algorithm we used the raw signals, the filtered signal (with the filter of reference) and the last preprocessed signal (after applied stationary wavelet) in order to compare the phasic component of these three.

### 3. Respiratory Rate and Accelerometer

The first step was to normalize both signals with a min max scaling rather than a standard scaling due to the presence of important artefacts remaining.

In case of the respiratory rate signal, frequencies of interest were slightly lower than for the ECG and EDA signals. That is why, with the previous conclusion on filtering for the other signals, a low pass filter was chosen for the denoising step. After several tests with a wide range of parameters variation, a cut off frequency of 3 Hz and an order of 4 appears to be quite adapted. Indeed, on the one hand the order is not too high which limits the phase shift feature of the filter and on the other hand, the cut off frequency is sufficient to smooth motion artefacts without affecting the respiration cycle component of the RR signal.

Because acceleration and RR signals focus both on the respiration cycle, they record the same motion source. Thus the previous low pass filter appears to be equally effective on the acceleration recording.

## 

## D. Data processing

### 1. Respiration rate and acceleration correlation

An objective of this project has been to quantify correlation between these two signals and see whether the Breath Rate Variation can be extracted from the acceleration signal only.

The first method applied to measure correlation was the instantaneous phase synchrony due to oscillating properties of the two signals. To calculate phase synchrony, we need to extract the phase of the signal which can be done by using the Hilbert transform which splits the signal into its phase and power. The instantaneous phase synchrony measure is a great way to compute moment-to-moment synchrony between two signals without arbitrarily deciding the window size as done in rolling window correlations.

Since the human normal breath rate is well known in the litterature, this second correlation method is worth applying (typical respiratory rate for a healthy adult at rest is 12–18 breaths per minute). Indeed, windowed correlation can be a good approximation of synchrony between two signals. When filtering is difficult due to uncertainty about which frequencies to analyze due to the stress triggered during the experiment, windowed correlation can be a good approximation of synchrony between two signals. In addition, this method provides a much more stable measure of synchrony robust to the high frequency random noise.

# 

# IV. Observations

## 1. Electrocardiogram signal

When displaying the raw data in the form of a graph, we can see the artifacts and noise recorded in the signal. It was therefore necessary to remove them to move on to the data processing.

First, with the help of the "biosignalsnotebooks" library, we tried to filter the signal by low-pass and band-pass filters that we can see on the following figures. Although the signal has been smoothed while removing noise, the artifacts are still present. The work done has not been sufficient to avoid these errors. It is possible that the "stationary wavelet transform" method presented in the paper [3] would free us from the artifacts

Since the artifact motions prevented us from having suitable metrics, it was necessary to find another way to obtain the heart rate variations so that it could be correlated with the other data.

The method used is a simple one, the purpose of which is to find the number of (R-peaks) of the QRS complex (which represents the contractions of the heart muscle) during a period and then find the BPM by the following formula,

bpm=

- nbpeaks : number of peaks during the interval

- tn : last peak

- t0 : first peak

- 60 : To obtain minutes

## 2. Electrodermal activity

Considering EDA signals we finally have numerous annotated plots like the figure X where we can observe the noise and unwanted artefacts. Our focus of interest remains in figure X and X where we observe clean correlation of the electrodermal activity (sweat generation) and the stress stimulus. We can also observe on figure X that each subject has a different response to the protocol.

## 

## 3. Respiration rate and acceleration correlation

The correlation methods performed allow us to assess if two signals are in phase (moving up and down together) or out of phase.

*to be continued*

# V. Conclusion

To conclude, we could only implement a code for signal pre-processing, the first step of this project. We had to face some difficulties that we had to overcome to advance on the project, with more or less ease depending on one's abilities.

To reach our results, we did bibliographic studies that we were able to implement and apply to our data. And this by using available libraries specially designed for the processing and analysis of biosignals. Thus, we were able to put in relation the data coming from the EDA and the breathing to better understand the functioning of the stress on the human body.

For the organization method, we noticed that even if this project could be realized in telecommuting, we preferred to work directly in the same room to advance on the project. It was easier to help each other when we encountered difficulties or to clarify the guidelines to follow.

Due to delay, we did not have the time to practice Machine Learning. We regret it because we think that it is the one which seemed the most interesting.

Although we were not able to complete the project, there were positives. Working on this kind of project was really rewarding because we were able to combine our own skills to work on this project. Also, we are glad to learn a little more about what is currently done in the biomedical field

# 

# References

[1] : *Application of independent component analysis in removing artefacts from the electrocardiogram*, He Taigang, Gari Clifford and Lionel Tarassenko, “Neural Computing & Applications ” 15.2 (2006)

[2] : *Signal Processing Techniques for Removing Noise from ECG Signals,* Rahul Kher,March 2016

[3] : *Motion artefact removals for wearable ECG using stationary wavelet transform,*  [Shuto Nagai](https://www.ncbi.nlm.nih.gov/pubmed/?term=Nagai%20S%5BAuthor%5D&cauthor=true&cauthor_uid=28868151), [Daisuke Anzai](https://www.ncbi.nlm.nih.gov/pubmed/?term=Anzai%20D%5BAuthor%5D&cauthor=true&cauthor_uid=28868151), and [Jianqing Wang](https://www.ncbi.nlm.nih.gov/pubmed/?term=Wang%20J%5BAuthor%5D&cauthor=true&cauthor_uid=28868151), June 2017

[4] : *Kelsey, M., “Artifact detection in electrodermal activity using sparse recovery”, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 2017, vol. 10211. doi:10.1117/12.2264027.*

[5] : *Jason J Braithwaite, Derrick G Watson,* *A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments, Technical Report, 2nd version: Selective Attention & Awareness Laboratory (SAAL) Behavioural Brain Sciences Centre, University of Birmingham, UK}, 2015*

[6] : *Chen W, Jaques N, Taylor S, Sano A, Fedor S, Picard RW. Wavelet-based motion artifact removal for electrodermal activity. Annu Int Conf IEEE Eng Med Biol Soc. 2015;2015:6223-6226. doi:10.1109/EMBC.2015.7319814*

[7] : *Greco, Alberto & Valenza, Gaetano & lanatà, Antonio & Scilingo, Enzo & Citi, Luca. (2016). cvxEDA: A Convex Optimization Approach to Electrodermal Activity Processing. IEEE Transactions on Biomedical Engineering. 2016. 797-804. 10.1109/TBME.2015.2474131.*