POLITECNICO DI MILANO

Facoltà di Ingegneria

Dipartimento di Elettronica, Informazione e Bioingegneria

Master of Science in Biomedical Engineering

MULTIVARIATE ANALYSIS OF PHYSIOLOGICAL VARIABLES FOR CLASSIFICATION OF EMOTIONAL RESPONSES TO VISUAL AND AUDITORY STIMULI

Advisor: Prof. Riccardo BARBIERI

Master Graduation Thesis by:

Matteo CAPONE Student ID. 850384

Francesca GAMBA Student ID. 852870

Anno Accademico: 2016/2017

Abstract

For decades, philosophers and sociologists have been debating about meaning, nature and definition of emotional processes. Actually, there is no total agreement about the discriminating criteria between emotion and what cannot be classified as an emotion. Emotion has been defined as a responsive psyco-physiological process to an emotionally potent antecedent event, causing changes in subjective feeling quality, expressive behavior and physiological activation.

In this study, a passive procedure for emotion induction is used, showing to subjects a presentation of standardized and pre-rated pictures from GAPED database and a videoclip, all with a proper musical background: the presentation includes four arousal sessions (depressed, elated, excited, stressed) spaced by neutral sessions.

Triggered reactions are measured through the acquisition of the three-real-time acquired physiological signals (ECG, Respiration, EDA).

These signals are then processed and appropriately combined in mono-variate and bivariate analysis in order to extract specific features, both in time and frequency domain, for each elicited emotional state.

Classification of detected features is finally performed to evaluate the distribution of the parameters, trying to define a criterion for emotional response identification.

Aims

"*Emotion has no voice*" quoting a very famous Italian song. Emotion has no specific and undeniable definition, it is sensation, it is thrill, it is impulsivity. Studying emotion have been charming researchers and scientists for centuries, without, however providing any certainty. Studying emotions and our physiological reactions is a way to understand more, to interpret, predict and prevent. It is a study about the human, his potential and his consciousness, his stories and his weaknesses.

Our study thus is aimed at recognizing, through the definition of a proper experimental protocol, emotive reactions of subjects, exploiting physiological changes. Our algorithm

seeks to recognize, by means of "low-level" analysis, the presence of variations between a neutral state, during which the subject is stimulated (in this case we deal with audio-visual stimuli) in order to evoke some specific reactions. Considered physiological changes are defined as stressed, elated, excited and depressed reactions.

Such evaluations relevance, nowadays, can match with practical portable applications, like smartphone or other accessories (smartwatch for instance) able to realize real-time records of physiological parameters and sudden fluctuations, sometimes dangerous, and able to provide an instantaneous evaluation of the subject. This could be actually useful in anti-depressive therapy, and to ensure a constant monitoring of the patient remote-controlled.

Furthermore, it would ensure a possible evaluation for overexcitation or undue stress situations, common reasons for heart failure or even stroke in most serious cases.

Prevent is better than cure, as they say, and for this reason we consider emotions as a starting point for the evaluation of exaggerated reactions, sometimes pleasant, sometimes unfortunately harmful, that can characterize daily routine of somebody. Finally, it is of common concern, from a scientific point of view, defining and implementing high level models able to classify human reactions through the evaluation of physiological signals changes, useful in field of applications like the scientific-criminal one, for the polygraph and for interrogations.

State of the art

Emotion studies are based on physiological reactions, described through signals and parameters that can be detected and analyzed in their fluctuations and dynamics. Our nervous system reacts, if stimulated, in different ways: it has been demonstrated that defining a description model for these variations is possible. An emotion has been frequently defined as a psychophysiological reaction to an event characterized by a relevant emotive potential.

The first one who studied emotions was Walter Cannon in 1915[1]. He identified emotion as a multi-component response including a variation in terms of feelings, attitudes, motion and activation. About feelings, in 1884 William James already talked, describing emotive response as slightly subsequent to the physiological change: emotion establishment,

according to the psychologist, is indeed the result of the cerebral reaction to the information received by the nervous system[2].

In 1958 John Lacey started to talk about measurements on physiological signals, also supported by the technology progress, by the emergence of appropriate sensors: he talked about heart rate, blood pressure, blood flow, temperature and skin resistance and other variables considered indicative of a variety of emotional states, since they trigger an autonomic response. Essential, in this framework, the elaboration of the Circumplex Model of Affect, a dimensional approach which classifies emotive states on the basis of two variables: valence (or pleasure dimension) and arousal (or activation dimension). Each emotion can be therefore distinguished and represented as a linear combination of these two dimensions. Thanks to this model, the concept of independent emotion is rejected, and reaction start to be evaluated also from a quantitatively point of view. For instance, an increase of skin conductance or heart rate is related to a precise arousal and valence levels, and thus to a subjective activation of the sympathetic nervous system.

In this framework, an important study of 1994 by the Center for the Study of Emotion and Attention developed a pictures database, called IAPS (International Affective Picture System) including 944 images depicting objects or scenes of different type, evaluated and rated in order to be helpful in eliciting proper reactions. This study is actually derived from the quantitative model mentioned before, so from a 2-D classification system, with the aim to standardize different protocols and categories of experiments through the representation of people, animals, landscapes, objects and other things. This protocol has been used several times with the aim of define a possible classification of the emotion-physiological variation relation.

The most interesting study created a 9 sessions protocol, with alternating neutral and arousal stimulating sessions, while recording ECG, respiratory signal and EDA. From here we started for the definition of our study. Later, another database was created, by the University of Geneva in 2011. This database revised the pictures categories through a macro-division into positive (human, kids, natural landscapes and neutral objects) and negative images, in turn divided into subgroups: snakes, spiders, scenes of violation of human rights and animal mistreatments. The choose of these pictures and of this modification has been moved by the necessity and the desire of a more reliable database, based on a new socio-cultural framework, far from that one at the basis of the IAPS

database. Other stimuli have been used too, to evaluate people emotive responses, from videoclips to short films, obtaining significant results, exploiting facial movements records and video-photopletismography.

Methods

The experimental protocol has been defined starting from the state of the art. The main object of the test is a presentation, projected on a PC screen, lasting 10 minutes, divided into nine sessions, namely five "neutral sessions" (among which one longer lasting has been considered as the baseline condition for reference) and four "emotive sessions", depression, elation, excitation and stress, respectively.

All sessions are composed by pictures chosen from the GAPED databsse: neutral session, lasting 30 seconds each one, show pictures of simple objects or buildings. One of these five neutral sessions is the baseline, lasting 110 seconds, represented by black screen. Emotive sessions last 90 seconds, presenting pictures of different scenes according to the emotion that they should elicit.

In the excitation sessions, pictures are replaced with a video, composed by different clips showing stunts and extreme sports, obtained with GoPro. For depression, elation and stress sessions, appropriate music was chosen as background.

The test involved 105 people, 56 men and 49 women, aging between 20 and 27, students at Politecnico di Milano. Tester were previously asked to fill in a questionnaire, starting point for performing the experiment. The questionnaire is the PHQ-9 (Patient Health Questionnaire – 9) developed and used for diagnosis on the potential presence of a psychological depression and on its seriousness.

Thereafter, subjects were invited to SPiNLabs, at Politecnico di Milano, and were trained about the object of the experiment: a 10 minutes lasting presentation of pictures and video, and the simultaneous acquisition of three physiological signals (ECG, Respiratory signal and electrodermal activity), that would be later analyzed off-line, with the aim of detecting a pattern pointing out the elicited emotive reaction.

Signals were acquired by a real-time analysis encoder, *ProComp Infiniti*. On subjects, analysis evaluating parameters and dynamics in temporal and spectral domain were performed. Starting from the analysis of the cardiac variability analysis, principal indexes found in literature have been computed for each single subject: SDNN, RMSSD,

Triangular Index, SD1, SD1. Then, another analysis was performed in the frequency domain, useful to define fluctuations typical of sympathetic or parasympathetic activation or deactivation.

The analysis was carried out evaluating the power content of the signal in the bands of interest, namely Low Frequencies (LF) and High Frequencies (HF), their ration and normalized values. Respiratory signal was instead analyzed in terms of power contempt at High Frequencies, considering position and amplitude of the power spectral density peak in this frequency band. Last, but not least, skin conductance analysis, in terms of peaks occurrences in the signal for each analyzed section, together with height, width and amplitude values, and width at half maximum.

Results

Signals were acquired by a real-time analysis encoder, *ProComp Infiniti*. Analysis evaluating parameters and dynamics in temporal and spectral domain were performed in all subjects. Starting from ecg-derived heart rate variability analysis, principal indexes found in literature have been computed for each single subject: SDNN, RMSSD, Triangular Index, SD1, SD1. Then, another analysis was performed in the frequency domain, useful to define fluctuations typical of sympathetic or parasympathetic activation or deactivation.

The analysis was carried out evaluating the power content of the signal in the bands of interest, namely Low Frequencies (LF) and High Frequencies (HF), their ration and normalized values. The respiratory signal was also analyzed in the frequency domain, considering position and amplitude of the power spectral density peak in the HF frequency band. Last, but not least, skin conductance was analysed in terms of peak occurrences in the signal for each analyzed section, together with height, width and amplitude values, and width at half maximum. Applying all these evaluation to all subjects, after discarding those with noisy signals, pattern of variation corresponding to expectations, were found. For most of the subjects, a significant activation in the stress and excitation sessions was found, in particular a positive variation with respect to the baseline condition regarding power parameters in the low frequencies range of ECG signal, typical manifestation of sympathetic activation.

For temporal parameters too, results reflecting what is demonstrated and validated in literature were found. Importantly, since our protocol is characterized by short session, HRV indexes in time domain cannot be considered very significant.

Further analysis was performed to compute the coherence resulting from a bivariate analysis between the RR interval series and the Respiratory signal. Results revealed a significant correlation in the high frequencies band, thus indicating how respiratory signal can affect heart rate variability.

Furthermore, classification showed good results in discriminating baseline from the other arousal sessions: in particular, excellent results have been obtained comparing baseline session with excited and stressed ones, pointing out the greatest accuracy level.

Conclusion

In this study, we have developed a novel pilot protocol for emotion elicitation and have tested its efficacy as compared with state-of-the-art analysis of noninvasive physiological variables. Given the challenging nature of the novel settings, combining visual and auditory elicitation, results can be considered as very preliminary.

Both monovariate and bivariate feature analyses yielded results in accordance with the state of the art, regarding obtained physiological fluctuations and changes. In particular, we observed consistent results from stress and exciting sessions, both in terms of classification learner and emotion discrimination. The last stage of this work has been the analysis performed through the classification learner, aimed at discriminating between the four emotions and the related responses, relying on the selected features, i.e. those able to differentiate between populations from a statistical point of view. Indeed, parameter selected for their significance, evaluated by the Wilcoxon test (low frequencies power, normalized low and high frequencies power and simpatho-vagal balance for the ECG, peak frequency for respiration signal and number of peaks for skin conductance) are the relevant features characterizing the classification study. Baseline has been properly discriminated from the four emotions, with accuracy higher than 70%, and acceptable results were obtained by comparing stress and excitation with depression.

Despite the high margin of improvement, these preliminary results establish a good starting point for a more accurate evaluation. Undoubtedly, the greatest limitation we have observed is related to the controlled environment in which the test needs to be performed.

Indeed, we underestimated the importance of performing the test in a dark room, isolated, in a relaxed environment, in order to detect minimal variations in the patient.

We decided to experience a protocol coupling images to music, and we would like to be able to determine, in future, the influence that music can have on our emotive reactions. We believe that music should be considered fundamental and integral part of our everyday life, and thus surely an interesting element for this kind of studies.

We realize our final results have serious limitations. Nevertheless, we would like to highlight the value of the experimental protocol and the consequent time and effort invested by this endeavor. Indeed, this study allowed for important assessments that will be of critical value for future studies of emotion elicitation by visual-auditory stimuli.

Sommario

Per molti anni filosofi e sociologi si sono interrogati sul significato, la natura e la definizione di processi emozionati. Al momento non c'è un accordo totale sui criteri per discriminare le emozioni, né tantomeno su cosa possa essere definito tale.

L'emozione è stata spesso identificata come un processo di risposta psico-fisiologica a un evento antecedente dal forte contenuto emotivo, capace di causa cambiamenti in sentimenti, comportamenti espressimi e attivazioni fisiologiche.

In questo studio è stata utilizzata una procedura passiva per suscitare le emozioni, mostrando ai soggetti una presentazione di immagini standardizzate e precedentemente classificate (dal database GAPED) e un videoclip, tutti opportunamente accompagnati da una musica di sottofondo: la presentazione include quattro sessioni emotive (depressed, elated, excited, stressed) intervallate da sessioni neutrali.

Le reazioni provocate sono misurate attraverso l'acquisizione di tre segnali acquisiti in tempo reale (ECG, respiro, EDA).

I segnali sono stati poi esaminati con analisi mono-variate e bivariate, estraendo le caratteristiche nel dominio temporale e in quello in frequenza, per ogni emozione suscitata.

La classificazione di queste features è stata poi realizzata cercando di identificare una distribuzione dei parametri che permettesse di definire un criterio per l'individuazione della risposta emotiva.

Scopi

"L'emozione non ha voce", citando una famosa canzone italiana. L'emozione non ha definizione precisa e inconfutabile, è sensazione, è brivido, è impulsività. Lo studio delle emozioni affascina ricercatori e scienziati da secoli, senza fornire però certezza alcuna.

Studiare le emozioni e le nostre reazioni fisiologiche è un modo per capirne di più, interpretare, prevedere e prevenire. È uno studio sull'uomo, le sue potenzialità e la sua coscienza, le sue storie e debolezze.

Il nostro lavoro mira pertanto a riconoscere, attraverso la definizione di un opportuno protocollo, le reazioni emotive dei soggetti, sfruttando le variazioni percepite a livello fisiologico. Il nostro algoritmo mira a riconoscere, attraverso un'analisi di "basso-livello",

la presenza di variazioni tra uno stato di baseline /neutralità e fasi in cui al soggetto vengono forniti stimoli, in questo caso audio-visivi, capaci di suscitare reazioni emotive. Le reazioni considerate rientrano nell'ambito di definizione di reazione da stress, da eccitazione, da rilassamento e depressione.

L'importanza di valutazioni di questo tipo, al giorno d'oggi, trova riscontro nelle possibili applicazioni portatili, quali smartphone o accessori (orologi, ad esempio) che possano registrare in tempo reale parametri fisiologici e variazioni repentine, a volte pericolose, e possano fornire e fornirci una valutazione istantanea del soggetto. Questo potrebbe infatti essere utile in ambiti come le terapie anti-depressive, per garantire un monitoraggio costante di un paziente sottoposto a determinati "controlli" a distanza.

Garantirebbe, inoltre, una possibile valutazione di situazioni da sovra-eccitazione o stress eccessivo, cause frequenti di scompensi cardiaci e, in casi gravi, infarti.

Prevenire è meglio che curare, lo si dice spesso, e per questo riteniamo che le emozioni possano essere un punto di partenza importante per valutare le reazioni eccessive, a volte piacevoli, talvolta purtroppo dannose, che possano contraddistinguere la quotidianità di alcune persone.

Infine, è oramai di interesse comune a livello scientifico la definizione e l'implementazione di modelli di alto livello capaci di classificare le reazioni delle persone attraverso la valutazione di segnali fisiologici, utile in ambiti come il campo scientifico-criminale, per test della verità e interrogatori.

Stato dell'arte

Alla base dello studio delle emozioni ci sono le reazioni fisiologiche, descritte attraverso parametri e segnali che possono essere registrati e analizzati, nelle loro variazioni e dinamiche. Il sistema nervoso reagisce, se stimolato, e reagisce in modi diversi: definire un modello di descrizione di queste variazioni è, come tanti studi hanno dimostrato, possibile.

L'emozione è stata più volte definita come un processo di risposta psico-fisiologica a un evento dal potenziale emotivo significativo. Il primo a studiare la fisiologia delle emozioni è stato Walter Cannon, nel 1915[1]. Ha indicato, appunto, l'emozione come una risposta multi-componente composta da una variazione di sentimenti, di atteggiamenti, da movimento, da attivazione. Di sentimenti aveva già parlato, nel 1884, William James, descrivendo la risposta emotiva come una risposta leggermente successiva, temporalmente

parlando, al cambiamento fisiologico: la formazione dell'emozione, secondo lo psicologo, è infatti il risultato della reazione cerebrale all'informazione ricevuta dal sistema nervoso del corpo.

Nel 1958 John Lacey ha iniziato a parlare di misurazioni di segnali fisiologici, anche favorito dall'avanzamento tecnologico, dalla nascita di sensori appropriati: ha parlato di heart rate, pressione sanguigna, flusso, temperatura e resistenza della pelle, e altre variabili che sono considerate indicative di una varietà di stati emotivi, in quanto stimolano una risposta del sistema autonomo.

Fondamentale, in questo contesto, l'elaborazione del circumplex model of affect, approccio dimensionale che classifica gli stati emotivi in base a due variabili: valenza (o dimensione del piacere), e arousal (dimensione di attivazione). Ogni emozione si può quindi distinguere e rappresentare attraverso una combinazione di queste due dimensioni. Attraverso questo studio, si rifiuta il concetto di emozione indipendente, ma si iniziano a valutare le relazioni emotive anche dal punto di vista quantitativo. Ad esempio un aumento nella conduttanza della pelle o nella frequenza cardiaca è correlata con un determinato livello di valenza a arousal, e quindi una soggettiva attivazione del sistema nervoso simpatico.

In questo contesto si sviluppa lo studio eseguito nel 1994 dal Center for the Study of Emotion and Attention, che ha sviluppato un database di immagini, denominato IAPS (International Affective Picture System), che comprende 944 immagini rappresentanti immagini o situazioni di diversi tipi, ma valutate e validate per suscitare determinate reazioni. Lo studio condotto da questo gruppo è partito proprio dal modello prima indicato, quindi da un sistema 2-D di classificazione, per riuscire a standardizzare protocolli e tipologie di esperimenti, attraverso la rappresentazione di persone, animali, paesaggi, oggetti e scene di vario tipo. Questo protocollo è stato utilizzato più volte per determinare una possibile classificazione della relazione emozione-variazione fisiologica. Lo studio più interessante ha visto la determinazione di un protocollo composto di 9 sessioni, alternando sessioni neutrali a sessioni stimolanti, e valutando come segnali fisiologici ECG, RSP, e EDA. È da qui che siamo partiti per la definizione del nostro studio.

Successivamente un altro database è stato creato da un centro di ricerca dell'università di Ginevra, più recentemente (2011). Questo database ha modificato le categorie precedentemente inserite, attraverso una macro-divisione in immagini positive (uomini,

bambini e paesaggi naturali), neutrali (oggetti e arredamento) e negative, a loro volta divise in gruppi: serpenti, ragni, scene di violazione dei diritti umani e maltrattamenti animali. La scelta di queste immagini e di questa modifica è stata mossa dal bisogno e dalla volontà di creare un database più affidabile, in quanto basato su un nuovo contesto socio-culturale, lontano da quello in cui si è definito lo studio del database IAPS.
Anche altri stimoli sono stati utilizzati per valutare le risposte emotive delle persone, partendo da videoclips fino a estratti di film, ottenendo anche risultasti considerevoli, sfruttando la registrazione dei movimenti facciali e della video-foto-pletismografia.

Metodi

Il protocollo sperimentale è stato definito partendo da quanto riscontrato in stato dell'arte. L'oggetto del test è una presentazione, della durata di dieci minuti, divisa in nove sessioni, di cui cinque neutrali (una più lunga che costituisce la baseline da cui ottenere un riferimento) e quattro sessioni emotive, rispettivamente depressione, rilassamento, eccitazione e stress. Tutte le sezioni sono caratterizzate da immagini scelte dal database GAPED: le sezioni neutrali, della durata di 30 secondi l'una, presentano immagini raffiguranti semplici oggetti o paesaggi. Una di queste cinque sezioni neutrali è la baseline, della durata di 110 secondi, che presenta invece lo schermo nero. Le sessioni emotive presentano invece una durata di 90 secondi, con l'alternanza di immagini rappresentanti situazioni diverse a seconda dell'emozione da suscitare. Solo nella sessione relativa all'eccitazione, le immagini sono state sostituite da un video montato inserendo diversi frammenti di video di acrobazie sportive eseguiti con GoPro. Per le sessioni di depressione, rilassamento e stress vi era un sottofondo musicale opportunamente scelto.
Il test ha visto impiegate 105 persone, di cui 56 uomini e 49 donne, di età compresa tra i 20 e i 27 anni, studenti del Politecnico di Milano, contattati direttamente. Ai tester è stato proposto un questionario da compilare, punto di partenza per lo svolgimento di questo esperimento. Il questionario risponde a quello che è il cosiddetto protocollo phq9 (Patient Health Questionnaire-9) sviluppato e utilizzato per diagnosticare e controllare la presenza, ed eventualmente la misura della sua gravità, di uno stato depressivo nel soggetto. Successivamente, i soggetti sono stati invitati allo SPiNLabs, Politecnico di Milano, e ad essi è stato spiegato molto semplicemente quello che è l'oggetto di questo esperimento: una presentazione di dieci minuti, e la contemporanea acquisizione di tre segnali fisiologici

(ECG,Respirazione e conduttanza della pelle), da analizzare successivamente, offline, alla ricerca di pattern che possano effettivamente identificare la reazione emotiva provocata.

Risultati

I segnali sono stati acquisiti grazie ad un encoder per analisi in tempo reale, come il ProComp Infiniti, dal quale sono stati estratti ECG, Respiro e conduttanza della pelle, poi opportunamente valutati. Sui soggetti sono state svolte analisi che valutassero parametri e dinamiche nel dominio temporale e nel dominio in frequenza. Partendo dall'analisi del segnale di variabilità cardiaca, i principali indici ritrovati in stato dell'arte sono stati calcolati, per ogni soggetto: SDNN, RMSSD, Indice triangolare, SD1, SD2. Successivamente è stata svolta un'analisi nel campo delle frequenze, utile per definire quelle variazioni caratteristiche dell'attivazione, o meno, del sistema simpatico o parasimpatico. L'analisi è stata effettuata valutando la potenza del segnale nelle varie bande di interesse, cioè le basse frequenze (LF) e le alte frequenze (HF), il loro rapporto, i loro valori normalizzati. Il segnale respiratorio è stato invece analizzato nei termini della sua potenza alle alte frequenze, considerando la posizione e l'ampiezza del picco in questa banda. Ultima, ma non meno importante, l'analisi della conduttanza della pelle, in termini di occorrenze dei picchi per ogni sezione analizzata, valori di altezza, larghezza e ampiezza a metà altezza del massimo picco.

Estendendo queste valutazioni a tutti i soggetti acquisiti, dopo aver opportunamente scartato quei soggetti i cui segnali risultavano fortemente influenzati da rumori, o artefatti da movimento, si sono notati pattern di variazioni che rispondono a quelle che sono le attese e le aspettative. I soggetti rispondono, per la maggioranza, con un'attivazione considerevole nell'area di stress e eccitazione, in particolare una variazione di intensità superiore alla baseline per quanto riguarda i parametri della potenza nelle basse frequenze del segnale ECG, tipica risposta di attivazione simpatica.

Anche per quanto riguarda i parametri temporali ci sono risultati che rispecchiano quanto dimostrato e validato in passato da stato dell'arte, ma per quanto riguarda il dominio del tempo bisogna considerare il fatto che, essendo il nostro protocollo caratterizzato da sessioni brevi, gli indici di variabilità cardiaca nel dominio temporale sono poco significativi.

Ulteriore analisi che è stata effettuata è quella che riguarda la coerenza risultante da una valutazione bi-variata tra ECG e segnale respiratorio, che ha evidenziato nei soggetti una correlazione significativa nella banda delle alte frequenze, che effettivamente sottolinea come il segnale respiratorio possa influire sulla variabilità cardiaca.

Conclusioni

Il protocollo da noi sviluppato ha mostrato sicuramente dei buoni risultati, ma ancora approssimativi. Costituiscono un ottimo punto di partenza per una valutazione più accurata.

In termini di analisi dei singoli parametri, le valutazioni effettuate hanno riportato risultati perfettamente in linea con lo stato dell'arte, nelle variazioni e nei cambiamenti fisiologici ottenuti.

Le sessioni di stress e eccitazione sono quelle che hanno ottenuto un riscontro migliore, anche in termini di classificazione e distinzione delle emozioni. L'ultimo atto di questo studio è stato analizzare, tramite apposito classification learner, la possibilità di distinguere le emozioni e le risposte in base ai parametri selezionati, cioè quei parametri che hanno presentato le popolazioni analizzate come differenti in modo significativo da un punto di vista statistico. Infatti, i parametri che hanno "superato" il test di Wilcoxon (potenza alle basse frequenze, potenza normalizzata alle basse e alte frequenze, e equilibrio simpato-vagale per quanto riguarda l'ECG, frequenza del picco della respirazione e numero dei picchi per sessione per la skin conductance) sono i features che hanno caratterizzato lo studio di classificazione.

La baseline è stata adeguatamente distinta dalle altre emozioni, superando il 70% di accuratezza in questa valutazione, e buoni risultati sono stati ottenuti anche dal confronto tra le sessioni di stress ed eccitazione e quella di depressione.

Ancora lontani dalla possibilità di distinguere tutti e cinque gli stati emozionali nella stessa analisi, ma restano promettenti i risultati raggiunti, come punto di partenza per un nuovo set-up sperimentale.

Sicuramente riteniamo un grosso limite quello di aver svolto i test in un ambiente "disturbato", da rumori, persone, situazioni. Probabilmente i test andrebbero svolti in una stanza buia, isolata, in un ambiente totalmente rilassato, per poter cogliere le minime variazioni del soggetto.

Abbiamo voluto sperimentare un protocollo che accoppiasse le immagini alla musica, e vorremmo riuscire, in futuro, a determinare l'influenza che questa possa apportare alle nostre reazioni emotive. Riteniamo la musica parte integrante e fondamentale della quotidianità, e quindi elemento di sicuro interesse.

Il percorso intrapreso attraverso questo studio è stato dispendioso, ma fortemente affascinante. Lo studio delle emozioni intriga, colpisce e sorprende. Quello che siamo, proviamo e pensiamo, quanto di più irrazionale e spontaneo ci possa colpire, attraversare e rappresentare. Le nostre emozioni sono il nostro più grande mistero ma anche la nostra più evidente debolezza, per quanto tu possa nasconderle, per quanto tu possa evitarle, si manifestano, reclamano spazio, attraverso ciò che non puoi controllare, né mascherare: il tuo sistema nervoso.

Contents

List of tables

List of figures

Acronyms

ANS	Autonomic Nervous System
CMA	Circumplex Model of Affect(s)
CR	Cardio-Respiratory
ECG	Electrocardiogram
EDA	Electrodermal Activity
HR	Heart Rate
HRV	Heart Rate Variability
HF	High Frequency spectral HRV
LDA	Linear Discriminant Analysis
LF	Low Frequency spectral HRV
LF/HF	Simpatho-vagal balance
PCA	Principal Component Analysis
Pnn50	Percentage of successive normal sinus RR intervals > 50 ms
PPG	Photo-pletysmography
RF	Reticular Formation
RMSSD	Root-mean-square of successive normal sinus R-R interval differences
RR	Respiration Rate
R-R	R-R intervals
SC	Skin Conductance
SCL	Skin Conductance Level
SCR	Skin Conductance Response (amplitude)
SKT	Skin Conductance Temperature
SDNN	Standard deviation of the normal-to-normal intervals
SVM	Support Vector Machine
VLF	Very Low Frequency spectral HRV
SBP	Systolic Blood Pressure
Te	Expiratory time
Ti	Inspiratory time

1. Introduction

1.1 Emotions

"The best and most beautiful things in the world cannot be seen or even touched. They must be felt" (Helen Keller).

It's an unexplainable, fateful attraction towards what we are not able to explain, concretize, interpret.

An emotion arises, without asking for permission, without knocking at the door. An emotion leads our decisions, participates to every event of our daily life making it colorful.

The centrality of an emotion destabilizes reason and rationality, reversing the balance.

Childbirth is an emotion, winning is an emotion, just as acknowledgement and pain.

But this reaction induces action, subverts the order. It prompts us, stimulates us, encourages and boosts us. It affects our actions and efforts, coordinating and steering attitudes and habits. It thus becomes the dominating element in social circumstances, affecting our ability to interface with the environment, to face the problems.

Nevertheless, how can we represent it? How can we sum it up in a single, simple, word? We used to associate it with descriptions that reflect the action, the behavior, the answer. We observe people and we summaries in a single word the description of their individuality, but not its nature.

Everyone knows what an emotion is, until asked to give a definition. Then, it seems, no one knows.

For decades, philosophers and sociologists have been debating about meaning, nature and definition of emotional processes. As much as there can exist mutual agreement about the classification of fear, anger, sadness and joy as emotions, divergent opinions arise when analyzing other states (as startle, interest, guilt and others). Actually, there is no total agreement about the discriminating criteria between emotion and what cannot be classified as an emotion.

The "defining" criteria have been based on specific behaviors believed to be produced by the emotions, linguistic properties of the English words used to denote various states and distinctive patterns of physiological activity, such as characteristic facial expressions.
In 1960 Hilman, suggested, taking a cue from the Drever's (work) *"Dictionary* of Psychology", this general and broad definition:

> *"Emotion: differently described and explained by different psychologists, but all agree that it is a complex state of the organism, involving bodily changes of a widespread character-in breathing, pulse, gland secretion, etc.-and, on the mental side, a state of excitement or perturbation, marked by strong feeling, and usually an impulse toward a definitive form of behavior. If the emotion is intense there is some disturbance of the intellectual functions, a measure of dissociation, and a tendency towards action"[3].*

We're talking, like Darwin did, about emotions as adaptation aimed for the survival, whether human or animals, for some social struggle the sentient being wants to escape from or react to, throughout the emotional means; an unspecified goal to achieve.
Emotions serve an adaptive role in our lives by motivating us to act quickly and take actions that will maximize our chances of survival and success. Charles Darwin was one of the earliest researchers to scientifically study emotions. He suggested that emotional displays could also play an important role in safety and survival. If you encountered a hissing or spitting animal, it would clearly indicate that the creature was angry and defensive, leading to you back off and avoid possible danger. In much the same way, understanding the emotional displays of others gives us clear information about how we might need to respond in a particular situation.
Similar definitions leave room for several doubts and unanswered questions about the distinction between "emotion" and "not-emotion", as is evident in "reactions" like surprise, love, the so-called "aesthetic" emotions, those that don't show in the reaction a particular reminder to a behavioural alteration, at least not declared, not plain.

> *"Does this mean that love, generosity, kindness, compassion, honesty, and other commendable human characteristics are nothing but the result of conscious but selfish, survival-oriented neurobiological regulation? Does this deny the possibility of altruism and negate free will? Does this mean that there is no true love, no sincere*

friendship, no genuine compassion? That is definitely not the case. Love is true, friendship sincere, and compassion genuine, if I do not lie about how I feel, if I really feel loving, friendly, and compassionate". (Descartes, Antonio Damasio)[4].

The pureness of an emotion, its greatness and depth, travels throughout crossing roads of experience, education, society, brain, physiology, culture. Its simplification, its scientific description doesn't make it less fascinating or less complicated, it doesn't devalue it.
Damasio describes the human being as an entity brought to life with its own independent base, aimed at the survival, and with an additional "luggage" that let it handle its decision-making processes, improving the quality of its own existence, building itself. All of that creates, in our own consciousness, a rational and judicious ethic, the one that, following the emotional connection bonding each other, can be considered accountable for our adjusting emotions.

1.1.1 Emotional Intelligence

"I don't want to be at the mercy of my emotions. I want to use them, to enjoy them, and to dominate them." (Oscar Wilde-The Picture of Dorian Gray).

Rational intelligence is just one of the ability that the human being owns to deal with situations he faces every-day, it's just an instrument whereby the human can afford and solve problems.
In the decision-making process a no longer negligible component steps in: it should be evident that our success doesn't just account to a pragmatic cold analysis of choices/options but can rely on shortcuts, faster ways found flanking our rationality to the emotive thinking[5].
The talk is of intelligent emotion, and we have been talking about it since 1990, in a Peter Salovey and John Mayer 's work where it is identified as the "ability to control one's and other's feelings and emotions, to discern between each other and use this info to drive our own thoughts and actions" [6].
As it is clear in Figure 1, emotive intelligence is characterized by a complex process of sensations interiorization.

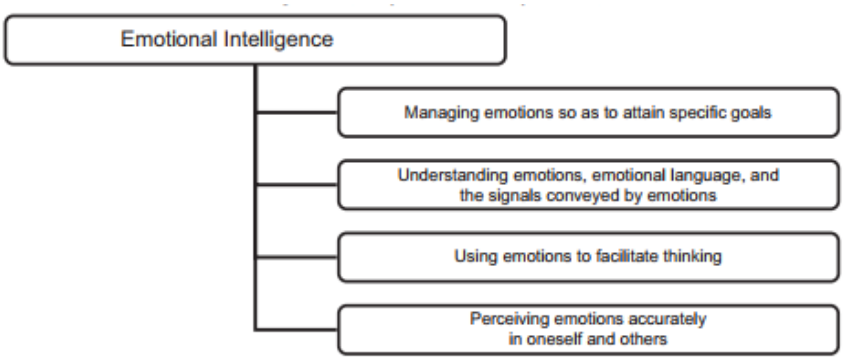


Figure 1 - The Four-Branch Model of Emotional Intelligence (1997)

As an example, imagine a situation in which a young man visits a friend in the hospital who has been in a car accident. The first area of EI involves perceiving emotions. As the young man surveys the hospital room, the visiting relatives, and his unconscious friend, he may wonder, "What is each family member feeling?" Perhaps he perceives the worry and anxiety in their faces. Feelings are complex; also emerging from within him may be fear of his own mortality and a guilty relief—with a surge in energy—in response to being spared the accident himself and remaining unharmed. The anxiety experienced by those around the young man redirects his attention from his own concerns to a focus on the well-being of his friend. Using energy from the fear and relief, he may feel motivated to talk with family members and find out how they are. This is an example of using emotion to facilitate thought. To understand the emotions of the situation involves asking "What sorts of feelings emerge from such a situation?" and "How can these feelings be expected to change over time?" The accident is unexpected and severe, so the family's shock is palpable. The young man may reason that one feature of such shock is its emergence from a rapid combination of surprise, sadness, and other mostly negative emotions (Goodrum, 2005). Knowing this, and understanding these feelings, he may find that one possible course would be to engage in emotion management. After regulating his own emotions, perhaps by observing them, and thereby psychologically distancing himself from them, the young person may inquire of the parents how they came to learn of the accident and how they are holding up, what their days are like, and how he can be of assistance. Listening creates a caring environment while helping to clarify the disturbing, ongoing events [6].

Emotive intelligence involves the ability to feel, evaluate and express an emotion, the ability to get into feelings and to foster the emotional and intellectual growth. The

cognitive sciences allowed to learn that the ability to understand and manage our own emotions gives us the chance to handle our own and other's feelings with the intention to achieve other goals. Thus, where do emotions rise?

1.2 Physiology of emotions

1.2.1 Autonomic nervous system (ANS)

Emotion has been defined as a responsive psico-physiological process to an emotionally potent antecedent event, causing changes in subjective feeling quality, expressive behavior and physiological activation. However, there is no one-to-one relationship between emotion and changes in autonomic activation.

Autonomic responding in emotion has been an active research topic since Walter Cannon, in 1915, first studied the physiology of motion; nowadays in fact, ANS activity is viewed as a major component of the emotion response in many recent theories of emotion.

Emotion and Autonomic Nervous System are intricately intertwined: emotions are a major consumer of autonomic resources, and the ANS provides a key to understanding many of the functions of human emotions. Refining the autonomic architecture of emotion and its integration with other processing and control systems is an important part of human evolution [7].

1.2.2 ANS architecture

The autonomic nervous system (ANS), is a division of the peripheral nervous system. It is a control system that primarily innervates the smooth musculature of all organs, the hearth and the glands, and mediates the neuronal regulation of the internal environment in order to ensure an appropriate balance or adjust it as required. This system, as its name suggests, doesn't act under direct voluntary control.

The ANS cooperate with the somatic nervous system, whose role is mediating afferent and efferent communications with the environment and which is subject to voluntary control and accessible to consciousness.

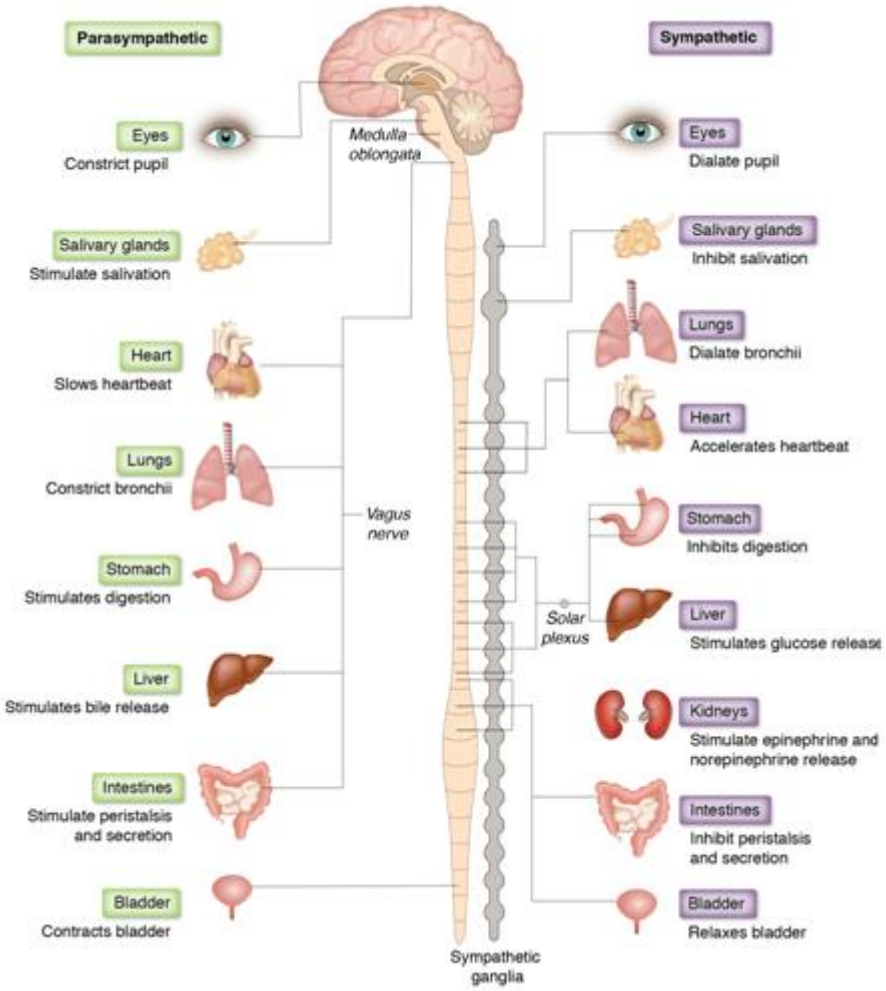


Figure 2 - Autonomic Nervous System: Parasympathetic and Sympathetic branches

The peripheral autonomic nervous system is distinguished into three parts: the sympathetic, the parasympathetic and the enteric nervous systems (Fig. 2). They all innervate the heart.

The basic peripheral element in both the sympathetic and the parasympathetic system is a set of two neuron populations in series. These two autonomic systems originate in different parts of the neuraxis. The sympathetic preganglionic neurons emerge from the thoracic cord and the upper 2-3 segments of the lumbar cord (thoracolumbar system), while those of the parasympathetic system emerge from the brainstem and the sacral cord (S2-S4) (craniosacral system).

The preganglionic neurons are those that send their axons into the ganglia and have their somata in the spinal cord and brainstem. The terminal sympathetic and parasympathetic neurons, named postganglionic neurons, are entirely outside the central nervous system,

with their cell bodies grouped in autonomic ganglia and their axons projecting from the ganglia to the target organs.

The enteric nervous system has an independent function.

Central coordination of autonomic activity represents a cornerstone of current views of integrated nervous system functioning. Unlike the original conceptualization of the ANS as functioning independently of the rest of the nervous system (e.g., involuntary, automatic, and autonomous control), close interactions between the central and autonomic nervous systems exist in various ways.

Thus, like the somatic nervous system, the ANS is integrated at all levels of nervous activity. Whereas segmental autonomic reflexes are coordinated by the spinal cord, suprasegmental integration higher in the brainstem is required for regulation of functions such as respiration, blood pressure, swallowing, and pupillary movement.

More complex integrating systems in the hypothalamus influence the brain stem autonomic subsystems. Many of the activities of the hypothalamus are, in turn, governed by certain cortical areas, particularly the insular, anterior cingulate, and ventromedial prefrontal cortices as well as the central nucleus of the amygdala, that process inputs. from the external environment.

Thus, fundamental adjustments of the organism to its environment can only be attained by the concerted coordination and integration of somatic and autonomic activities from the highest level of neurological activity in the cortex down to the spinal cord and peripheral nervous system. This high degree of specificity in ANS organization is needed for precise neural regulation of homeostatic and protective body functioning during different adaptive challenges in a continuously changing environment. In this context, emotions may provide quick and reliable responses to recurrent life challenges [7].

1.2.3 The Limbic System and the Amygdala

Emotions find place in the brainstem, near the lower part of the limbic system, more precisely in the amygdala, an almond-shaped group of nuclei located deep and medially within the temporal left and right lobes of the brain, through which the emotion is born, directed and guided. The Amygdala is considered part of the limbic system (Fig. 3), a set

of brain structures located on both sides of the thalamus, immediately beneath the cerebrum, whose structures are involved in motivation, emotion, learning, and memory [8].

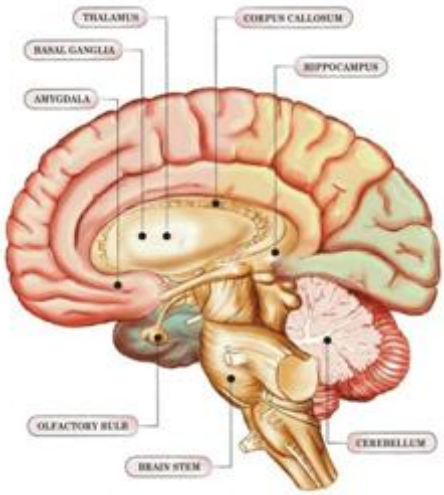


Figure 3 - The Limbic System

Emotional life is largely housed in the limbic system, and it has a great deal to do with the formation of memories. The limbic system is where the subcortical structures meet the cerebral cortex: it operates by influencing the endocrine system and the autonomic nervous system [8].

The central amygdala is at times viewed as a “controller of the brainstem” and uses its widespread projections to the hypothalamus and other brainstem nuclei, including the midbrain reticular formation (RF), to coordinate behavioral, autonomic, and neuroendochrine responses [9].

The assumption that emotional conditions induce long-term neural plasticity in the amygdala suggests that the interrelations between the amygdala and brain regions, such as the hippocampus, may not be static but dynamic. The way the amygdala will affect memory-related processes in the hippocampus may thus largely depend on the previous history of the individual[10].

1.2.4 Physiological changes

As said before, part of the complexity in studying emotion is defining it. Emotion is often considered synonymous with mental feelings in the popular culture, but its etymological roots are consistent with the biological imperative: the word emotion derives from the Latin *movere*, meaning to move. When emotions are intense, people move: they act, they react.

An aspect of emotion upon which most agree is that, in emotional situation, the body acts. The heart pounds, flutters, stops and drops; palms sweat; muscles tense and relax; blood boils; faces blush, flush, frown, and smile. Emotions are known as multi-componential responses that are composed of coordinated changes in subjective feeling, motor expression, and physiological activation. Additionally, they are processes directed towards a specific internal or external event or object, which result in changes in both behavior and bodily state (i.e., physiological change) [11] [12].

These changes are clear and strong in the so-called "coarser emotions", i.e. fear, rage, grief, love, in which everyone recognizes a strong organic reverberation. In a positive state, optimistic feelings dominate and cognitive functions are improved. In negative ones, pessimistic feelings dominate, our capacities are underestimated and analyitical thinking is increased.

> *"If we fancy some strong emotion, and then try to abstract from our consciousness of it all the feelings of its characteristic bodily symptoms, we find we have nothing left behind, no "mind-stuff" out of which the emotion can be constituted, and that a cold and neutral state of intellectual perception is all that remains. ... What kind of an emotion of fear would be left, if the feelings neither of quickened heart-beats nor of shallow breathing, neither of trembling lips nor of weakened limbs, neither of goose-flesh nor of visceral stirrings, were present, it is quite impossible to think. Can one fancy the state of rage and picture no ebullition of it in the chest, no flushing of the face, no dilatation of the nostrils, no clenching of the teeth, no impulse to vigorous action, but in their stead limp muscles, calm breathing, and a placid face? The present writer, for one, certainly cannot. The rage is as completely evaporated as the sensation of its so-called manifestations" [13].*

In 1884, psychologist and philosopher William James proposed that physiological changes actually precede emotions, which are equivalent to our subjective experience of physiological changes, and are experienced as feelings. In his words, "our feeling of the same changes as they occur is the emotion."

The basic premise of the theory is that physiological arousal instigates the experience of emotion. Instead of feeling an emotion and subsequent physiological (bodily) response, the theory proposes that the physiological change is primary, and emotion is then experienced when the brain reacts to the information received via the body's nervous system.

Emotion is a psychophysiological process, produced by the limbic system activity in response to a stimulus, which in turn leads to activation of the somatosensory system. Different peripheral physiological changes lead to different emotions, and a corporal feedback is necessary for the emergence of emotion.

We can represent this theory as follows:

Stimulus → Physiological responses → Peripheral changes sensation → Emotion

With the advent of electronic sensors and amplifiers it has become possible to measure a broad range of physiological reactions to emotional stimuli. In 1958, John Lacey state:

> *"Such measures as skin resistance, heart rate, blood pressure, blood flow, skin temperature, blood-oxygen saturation, gastric motility, pupillary diameter, muscle tension, and other variables have been shown to be remarkably sensitive and responsive measures in a variety of emotional states. Conflict, threat and frustration, anxiety, anger, and fear; startle and pain; embarrassment; pleasant and unpleasant stimuli; - all these produce autonomic changes".*

Physiological signals have some advantages as the following although they may be venerable to artefact caused by motions or other external factors. First, the acquisition of physiological signals by noninvasive sensors is relatively simple and it makes us possible to monitor user's autonomic activity linked to emotional states in real time. Second, physiological responses are spontaneous, less sensitive in social and cultural differences.

1.2.5 Physiological signals

- *Skin Conductance (SKC):* is a measure for determining the electrical level of skin conductance; this response is the phenomenon in which the skin becomes momentarily a better conductor of electricity when external or internal stimuli occur. Typical value: 2 micros. It is also measured as electrodermal activity (EDA) and galvanic skin response (GSR).
- *Blood volume pulse (BVP):* measurement of blood volume may provide information on changes in sympathetic activation. This activity, acting on the diameter of blood vessels, leads to changes in blood volume and blood flow.
- *Respiratory volume (RV):* rest and relaxation lead to slower and more superficial breaths. Emotional excitement and physical activities generate deeper breaths. A stress state will be detectable by frequent breaths. Generally, emotions with negative valence cause irregular breathing. The energy of this signal is in the range of 0.1–10 Hz [6].
- *Electromyogram signal (EMG):* emotional tone is an involuntary, permanent and moderate contraction of muscle fuelled by nervous energy. The slight tension that affects any muscle at rest to exaggerate the effort is an expression of emotion changes as a state of mental stress. It is shown that muscle activity increases during stress and during emotions with negative valence.
- *Skin temperature (SKT):* changes in skin temperature are related to vasodilatation of peripheral blood vessels induced by increased activity of the sympathetic nervous system. If the person is stressed the temperature of the body extremities decreases because blood is redirected to vital organs as a protection measure. The dominant energy of SKT signal is in the band 0–1 Hz.
- *Heart rate (HR):* cardiovascular changes are necessary to prepare for action and likely reflect emotional experiences. In particular, it was found that the emotional valence is predicted by the heart rate. In this study the heart rate is determined through the measurement of the successive peaks in the BVP signal.
- *Photoplethysmography (PPG):* is a simple optical technique used to detect volumetric changes in blood in peripheral circulation. It is a low cost and non-invasive method that makes measurements at the surface of the skin. PPG makes uses of low-intensity

infrared (IR) light. When light travels through biological tissues it is absorbed by bones, skin pigments and both venous and arterial blood.

- *Blood-oxygen-level dependent (BOLD) imaging:* is a method used in functional magnetic resonance imaging (fMRI) to observe different areas of the brain or other organs, which are found to be active at any given time.
- *Electroencephalography (EEG):* is an electrophysiological monitoring method to record electrical activity of the brain. It is typically noninvasive, with the electrodes placed along the scalp, although invasive electrodes are sometimes used such as in electrocorticography. EEG measures voltage fluctuations resulting from ionic current within the neurons of the brain
- *Facial electromyography (fEMG):* refers to an electromyography (EMG) technique that measures muscle activity by detecting and amplifying the tiny electrical impulses that are generated by muscle fibers when they contract. The zygomaticus is predominantly involved in expressing happiness. The corrugator muscle can be used to measure the expression of negative emotions including anger and fear.
- *Pupillary dilation (PD):* is a physiological response that varies the size of the pupil, via the optic and oculomotor cranial nerve. The pupillary response, whose neural pathways are mediated by the ANS, is determined by the activity of two smooth iris muscles [14].

1.3 The circumplex model of affect

The circumplex model of affect is a dimensional approach proposing that all affective states are the results of two neurophysiological systems, one related to valence (or pleasant dimension) and one to arousal (or activation dimension) (Fig. 4) [15] .

Each emotion can be considered as a linear combination of these two dimensions.

Thus, a specific emotion arises out of patterns of activation within these two neurophysiological systems, paired with cognitive interpretations. This model provides an experimental framework to explore the neural basis of affect. Thanks to its dimensional nature, this approach provides powerful statistical and methodological tools for use in genetic, neuroimaging and neurobiological studies of affective disorders.

The theory of basic emotions assumes that humans have a discrete and limited set of basic emotions, each one independent of the others in terms of behavioral, physiological and psychological manifestations, each one arising from the activation within specific neural pathways of the central nervous system.

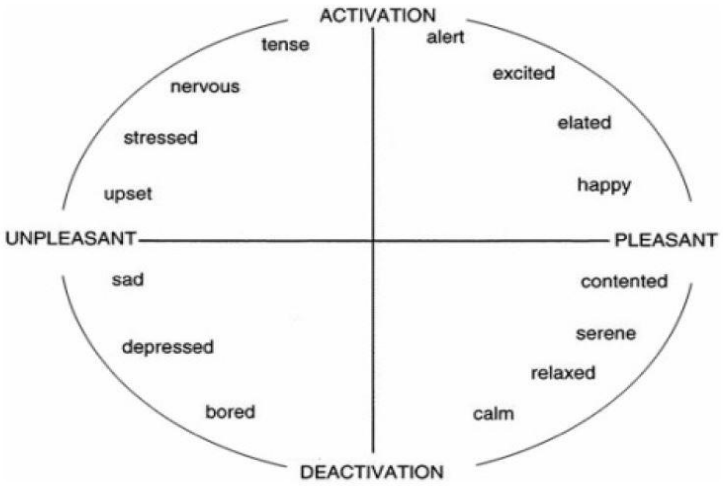


Figure 4 - The Circumplex Model of Affects

1.4 Music & emotion

"For me there is something primitively soothing about this music, and it went straight to my nervous system, making me feel ten feet tall." (Eric Clapton).

Music surrounds us everywhere: at home, at work, in cars, at shopping centers – or when making music ourselves. Juslin, Liljeström, Västfjäll, Barradas, and Silva (2008) confirmed the omnipresence of music in everyday life by using the experience sampling method: In 37% of all examined episodes, participants listened to music[16] [17].

Music is present in every culture, and it plays a prominent role in people's everyday lives.

It has the power to stimulate strong emotions within us: from sad, nostalgic, and tense, to happy, relaxed, calm, and joyous. The attraction of music comes from its "emotional powers": i.e., people tend to value music because it expresses and induces emotions[18].

But why do we perceive emotional information in musical features? Why do we have feel the urge to move when hearing music? Why does sound talk to our emotional brain?

One powerful emotional effect induced by music is a shivery type of skin sensation (commonly called "chills" or "thrills"), which may reflect the brain's ability to extract specific kinds of emotional meaning from music. A thrill is "a subtle nervous tremor caused by intense emotion or excitement (as pleasure, fear,etc.), producing a slight shudder or tingling through the body "[19] .

Emotions are relatively brief, intense, and rapidly changing reactions to potentially important events (subjective challenges or opportunities) in the external or internal environment - often of a social nature - which involve a number of subcomponents (cognitive changes, subjective feelings, expressive behavior, and action tendencies) that are more or less 'synchronized' during an emotional episode. Several experiments have shown that music listening can give rise to physiological responses very similar to those shown to other 'emotional' stimuli, including changes in heart rate, skin temperature, skin conductance, breathing and hormone secretion. Different pieces of music can produce different patterns of physiological response, such that it is possible to discriminate among emotions based on psychophysiological variables by using multivariate techniques[20].

Notably, the experience of pleasure to these abstract stimuli is highly specific to cultural and personal preferences, which can vary tremendously across individuals.

Moreover, the social condition (alone vs. with others) seems to influence which emotions are induced. Some emotions, such as happiness-elation, pleasure-enjoyment, and anger-irritation, often occur in "social" settings (during social interactions). Others, such as calm-contentment, nostalgia, sadness-melancholy, often occur in "solitary" settings (listening alone). Mechanisms of emotion induction are regarded as information-processing devices at various levels of the brain.

"Your Brain is Musical, Even If You Think You Might Not Be"

Through mere exposure, people learn to predict and anticipate the movements of the music; brainstem reflexes are quick, automatic and unlearned. Neuroimaging studies have shown that music can activate the brain areas typically associated with emotions: the deep brain structures that are part of the limbic system like the amygdala and the hippocampus as well as the pathways that transmit dopamine. The specific brain mechanisms that music arouses to yield emotional responses remain to be identified, but related neurodynamic issues are beginning to be understood with the tools of modern brain research.

The emotional power of music is the reason for its application in areas as diverse as the gaming industry, film industry, marketing, and music therapy, yet the scientific insights into this phenomenon are far from complete or revealing. Contemporary research on music and emotion is a popular topic within the fields of music cognition, music psychology, and neuroscience of music, and is, by definition, interdisciplinary. To date, a large number of research approaches have been used to explore why and how music has such a strong grip on listeners.

Emotions are one of the most fascinating features of the human mind. Music is an equally extraordinary characteristic. Understanding the special interaction between the two may take us closer to understanding the fundamental nature of both.

2. State of the art

2.1 The Circumplex Model of Affects

Modeling emotion is not a trivial task, because of the variety and extent of emotions. In literature, several approaches have been used, all developed from the oldest Darwinian hypothesis about universality of basic emotions, which, through time, resulted inappropriate due to neglecting intercultural differences (more crucial than intracultural differences)[21].

Discrete models consider emotions as the result of a selective adaptation for survival, concept which refers to this flow: danger → fear → escape → survival[21].

Ekman was the first at identifying the "basic emotions", choosing 6 emotional states to describe all: anger, disgust, fear, joy, sadness, surprise. Although, for mixed emotions, a proper description cannot be provided by this possible model. [art. 50]Dimensional models then, are based on a multidimensional continuous space in which all emotions can be described by those selected dimensions: Russel proposed those dimensions which nowadays result to be the most appropriate, namely valence and arousal[21].

Appraisal models thus, introduced by Arnold and based on the hypothesis of subjectivity of emotions, evaluate remembered or imagined situations. Dynamical models, finally, describe emotions through a dynamical process, starting from a concept of evolution and depicting each emotion through elicited response tendencies, considering mood and personality of the subject too[21].

In this work we used the Circumplex Model of Affects (CMA), aimed at complementing data from developmental, neuroimaging, and behavioral genetics studies about affective disorders and affective neuroscience in general. The circumplex model of affect is a dimensional approach, stating that valence and arousal, the two independent neurophysiological pathways, process all affective sensations, and that various other cognitive processes are aimed at interpreting and refining emotional experience[15]. (Fig. CMA)

This model rejects the theories of basic emotions, according to which each emotion is independent, in terms of its behavioral, psychological and physiological manifestations and is related to the activation of an independent neural pathway belonging to the CNS. Intercorrelations within and between emotions and people reporting them were in previous studies found by researchers and described by this quantitative model[15].

So, according to this CMA, each emotion, each affective state, comes out to be a linear combination of arousal and valence values which can then be interpreted as representing a particular emotional state. These complex interactions between cognitions occur primarily in neocortical structures. Peripheral physiological responses to affective stimuli vary incrementally with subjective rating of valence and arousal[15].

An increase in skin conductance and heart rate is correlated with subjective ratings of valence and arousal, and thus to the sympathetic nervous system activity. Also in fMRI an increase in signal intensity in the visual cortex and ratings of arousal can be seen. EEG shows an increase in cerebral activity in relation to subjective ratings of arousal [15].

Valence rating turned out to be correlated also to facial electromyographic measurements: an increase in corrugators activity is related to negative valence while the increase in the zygomatic muscle activity is related to positive valence ratings. This 2D approach is thus inconsistent with theories stating that specific responses to affective stimuli are associated with discrete affective states[15].

2.1.1 Previous studies

In literature, the CMA plane has often been considered an orthonormal space where each point is the linear combination of an arousal and a valence value. A previous study divided the CMA plane into 25 regions using classification (not necessarily representing 25 different emotions) in order to be later able to recognize emotions[22].

2.1.2 The Valence neural circuitry

The mesolimbic dopamine system has long been associated with pleasure and plays a significant role in the experience of negative emotions. It represents a neural substrate for the valence dimension of the circumplex model. Positive and negative emotions can furthermore be associated to an asymmetry of activity in the frontal lobe of the prefrontal

cortex: greater activation on the right or on the left accompanies the experience of more negatively and positive "valenced" emotions respectively[15].

2.1.3 The Arousal neural circuitry

The reticular formation (RF) regulates arousal levels of the CNS through its connections with the limbic system and thalamus. Sensory stimuli are sent from the thalamus to the amygdala, responding to both appetitive and aversive stimuli, but with a greater activation in response to more arousing stimuli. Determinations of emotional arousal are then sent from the amygdala to the RF through the association cortices of the parietal lobe. Descending tracts of the RF constitute the spino-reticular pathways that modulate muscle tone and sweat glands activity. The structure of this described Arousal Network is shown in Figure 5[15].

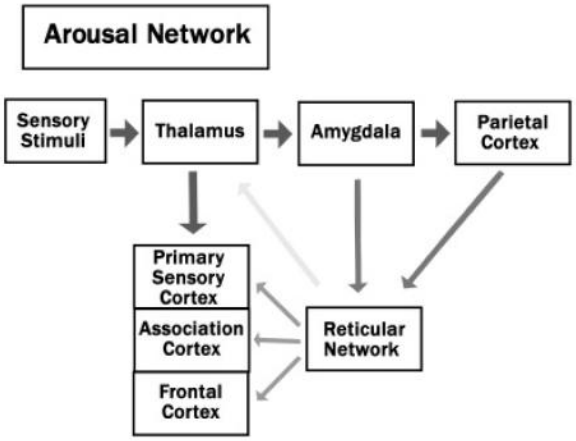


Figure 3. The pathways of the arousal network.

Figure 5 - The pathway of the arousal network

Mesolimbic and reticular networks support the phylogentically primitive sensations of pleasure and arousal.

Within the context of the CMA, prefrontal cortex integrates, organizes and structures the primitive feeling of pleasure and arousal with knowledge of the temporal events that relate prior experiences of stimuli within different life contexts with expectations for the future[15].

The most widely used procedure for emotion induction is the passive presentation, to the patient, of emotion eliciting materials of different type, such as images, film clips,

personalized recalls and others. Triggered reactions have then to be measured acquiring physiological signals, feelings, expressivity and behavior[23].

An emotion recognition algorithm requires: emotion elicitation, training, and final classification of detected features[21].

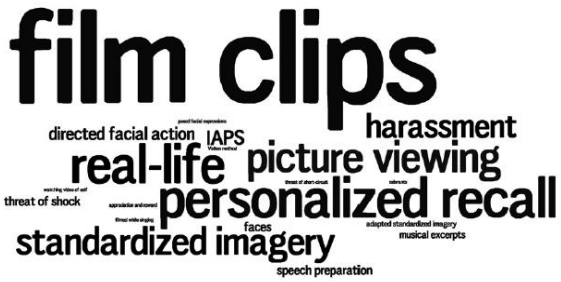


Figure 6 - Emotion induction methods

2.2 Static visual stimuli: pictures

2.2.1 IAPS Database

Emotions are systemic responses to delayed or inhibited motivated actions: in this sense they are attitudes and not the acts themselves, they reflect central activation and preparation for action: so, by processing ANS response to a suitable elicitation process, emotions can be discriminated [24].

Emotions have many different shapes, and each theory must face this diversity: they are the result of a Darwinian development and can be defined as motivationally tuned states of readiness. In this readiness, different responses are included: subjective experience, peripheral/autonomic nervous system and central nervous system activation and consequent physiological changes, and, lastly, behavioral consequences (Fig. 7) [24].

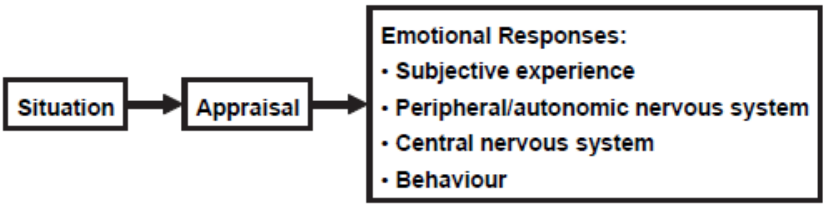


Figure 7 - A consensual component model of emotional responding [17]

Another work, proposed the existence of two different motive systems in the brain, an appetitive and an aversive one, according to the valence dimension. On the other hand, arousal doesn't have a separated substrate, but emotional responses are reflected in the level of both neural and metabolic activation[24].

Although the variety of emotion manifestations, many researchers choose for a 2-D organization.

On the circumplex model of affects (CMAs) the majority of emotion recognition system in literature bases its researches: in this model, each emotional affective state is assigned to a specific region of the defined 2-D arousal-valence space. Valence represents the pleasant or unpleasant character of an experienced emotion, with a positive or a negative value respectively. The other dimension, Arousal, measures the intensity of the stimulus categorizing it as activating or deactivating (positive and negative arousal value respectively). Both valence and arousal values are defined on a 0-1 scale. Thus, modeling elicitation stimuli, as well as modeling emotional response, is such a complex task.

In literature, many elicitation procedures were used, like introspection, movements, lights, set of actions, images, sounds, movies, speech, games, virtual reality interaction, reliving emotions and real experiences [25][26].

Although, since the most widespread and easy passive elicitation of emotions is through pictures, an international database was created in 1994 by the Center for the Study of Emotion and Attention, called *International Affective Picture System* (IAPS). It includes 944 pictures, somehow sampling the visual representation of the world, properly pre-rated in terms of valence and arousal level: rating of these pictures has been carried out through the Self-Assessment Manikin (SAM), a test used to rank images, submitted by many subjects in different studies[24] [22].

This picture database permits a controlled exposure, in terms of time, intensity and repeatability. The induced stimuli should be calibrated to become a standard measurement useful for scientific replication in emotion research[24].

The majority of previous works realizes static emotional stimulation with these IAPS pictures, which include photos representing people, animals, natural landscapes, objects and other scenes[24].

Distribution on the valence-arousal plane of a representative set of IAPS pictures is shown in Figure 8.

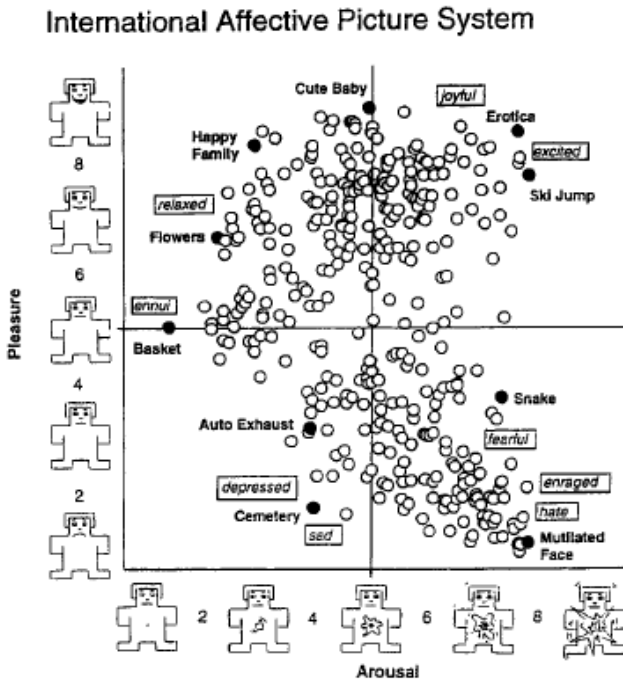


Figure 8 – Distribution of 360 Photographic Images from the IAPS database

2.2.1.1 Previous Studies

Many different elicitation protocols, using IAPS images can be found in literature.

Usually, a set of IAPS pictures with coherent arousal and valence levels is shown, as a passive emotional elicitation, alternating each arousing session to a neutral one.

In a previous study, arousing sessions composed by 20 pictures, each one lasting 10 s, are used, ensuring an elicitation of 40 seconds which allows the evaluation of the low frequencies indicating sympathetic activity[22].

Another experimental setup, similar to the one of the present work, was used by Valenza et al. The protocol consisted in an affective elicitation of the patient performed by 9 sessions of 20 IAPS images each one, alternating neutral and arousal sessions, the latter inducing increasing levels of valence[27].

Valenza et al. proposed a novel probabilistic framework for the characterization of emotions through the analysis of heartbeat dynamics exclusively. IAPS images were used as visual stimuli, in a sequence that alternated arousal and valence levels[28].

In a previous study, which used IAPS database as emotional stimuli, the criterion for the pictures selection was the following.

Two ranges of arousal and valence level were defined: L-M and M-H. by deciding a selective combination of these two regions of the valence-arousal space, four emotional states were distinguished: sadness (L-M valence and L-M arousal), anger (L-M valence and M-H arousal), happiness (M-H valence and M-H arousal), and relaxation (M-H valence and L-M arousal). For each of the mentioned classification problems, a leave-one-out procedure was at the end performed on the N available features using Support Vector Machine (SVM) algorithm[28].

IAPS elicitation consisted in a presentation of pictures composed by 9 sessions. Neutral sessions, placed between each two arousal sessions, includes 6 neutral images with valence rating in the range 5.52-7.08 and arousal rating in the range 2.42-3.22. Arousal sessions are composed by 20 images with increasing valence level (from unpleasant to pleasant). Signals recorded during emotional elicitation were ECG, RSP, EDA[22].

2.2.2 GAPED Database

IAPS database was created in order to reach a sufficient control in the induced stimuli and to allow a comparison and replication of protocols using the same material. This desired successful emotion induction, until 2011, has been reached with the use of pictures from the only reliable database, the IAPS. Nevertheless, the impact of these images becomes poor when the awareness of participants increases. Another limit concerns the little number of images.

IAPS database has been the only one about emotional responses induction for a long time. But its limits let to the creation of a new database, the GAPED one, whose main goal is to provide additional rated pictures to stimulate and elicit responses in a subject. It consists in positive and negative images: Positive (P) ones represents human and babies and natural landscapes, and include a Neutral (N) category containing pictures of objects, buildings or furniture. Negative images, because of their frequent use in psychology neuroscience, were instead divided into groups: spiders(Sp), snakes (Sn), scenes of violation of human rights (H) and scenes depicting animal mistreatment (A).

All pictures were rated in terms of valence and arousal values just like IAPS pictures. But another evaluation, taken into account in the rating process, was that concerning the congruence of the picture with moral and legal norms. Negative pictures have in general greater arousal values than neutral and positive ones.

Concerning arousal, a high level was matched to images including human concerns and animal mistreatment. Low arousal was instead elicited by positive and neutral pictures, while a medium arousal level was elicited by other categories.

Regarding valence, a low level was paired to negative pictures significantly differing from values of positive and neutral categories, while medium valence to preliminarily positively rated pictures concerning arousal.

Both valence and arousal values are defined on a 0-100 scale.

Aim of the creation of this new database was to provide a more reliable and reproducible base for an experimental elicitation protocol.

Representation of the ratings distribution in the valence-arousal space for each category of the database is shown in Figure 9: each polygon represents the area of the plane occupied by all the images belonging to a category [23].

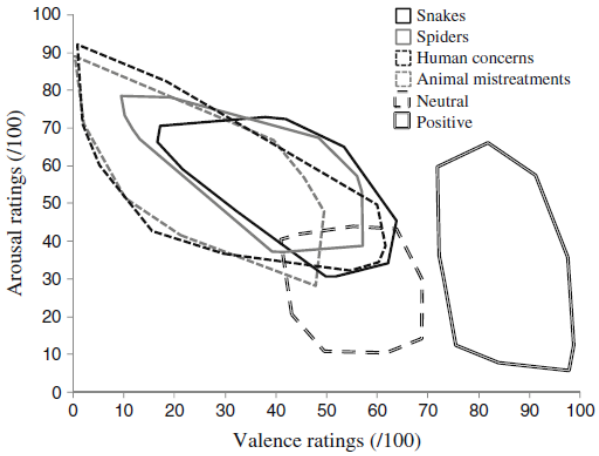


Figure 9 - Representation of the outcome ratings in the valence-arousal space for each GAPED database category. Polygons represent the surface occupied by all the images in a given category

Valence and arousal levels for each group of GAPED images are shown in Figure 10.

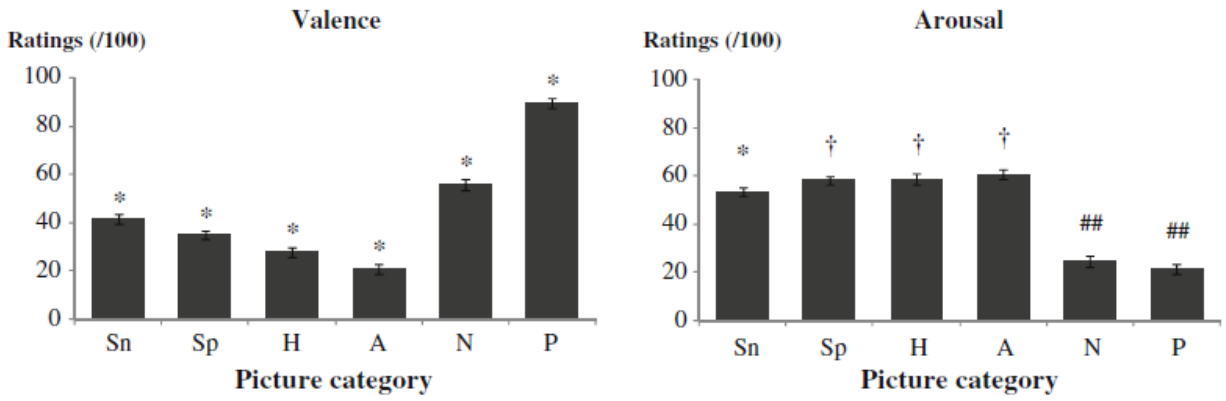


Figure 10 - Valence and Arousal for each group of images in GAPED database

2.3 Dynamic visual stimuli: videoclips

To arouse target emotions in the subject, a previous work used film clips as stimuli lasting from 3 to 8 minutes, which turned out to be more reliable and naturalistic in inducing emotions with respect to other emotion elicitation techniques. After stimuli proposal, a questionnaire was submitted to the subject who had to report his emotional state and to rate the intensity of the felt sensation on a five-point scale. Different videoclips were successfully chosen to elicit fear, neutral feelings, joy. The classification rates for fear, neutral, joy have been 76%, 94%, 84% respectively[29].

Another study used well-validated films to induce amusement and sadness, considered as positive and negative emotions respectively. The 5 minutes lasting videos projected on the screen was composed by three scenes extracted from the same film: two amusing scenes, spaced by a sad one. Each participant watched the film three times in a row while behavioral and physiological responses were recorded in order to be later analyzed[30].

In another study positive emotion films were shown to female subjects and the smoothness of their smiles was measured: enjoyment smiles were longer lasting and smoother than smiles for non-enjoyment. Facial behavior was recorded[31].

Another study was performed showing to subjects some brief films eliciting disgust and fear and subsequently asking them to report the emotion they felt during this affective episode: meanwhile facial expression were recorded. An evident coherence between facial-feedback and self-report was found. Disgust turned out to be associated with an activation in the right hemisphere in both frontal and temporal regions, while happiness was associated with a left side activation in the anterior temporal region[32].

2.4 Other emotion elicitation methods

Some researchers considered visual stimulation (performed through IAPS images) not sufficient to obtain strong emotional responses, so they created a new multimodal protocol, involving audio, visual and cognitive stimuli to evoke target emotions. Sadness, for example, is elicited through: a visual stimulation consisted in blue illumination, an audio stimulation in background music and cognitive stimulation in voice narrating a sad story[33].

Another way to elicit target emotions in the subject is asking him to report a biographical recall, recording his reaction while reliving target emotions[34].

2.5 Music

"Without music, life would be a mistake" (Nietzsche)

Life goes through labyrinthic ways, sometimes tough, sometimes easier ones, but it travels moved by music. Each aspect of our daily life beats following some music.

Even though music is universally appreciated, relation between our emotional reactions and listening to music is a fairly recent research field. The starting point is Sloboda's research (last century ending) in the book "Current trends in the study of music and emotion" , where some questions about music power in emotions elicitation are proposed, trying to reach the complexity and diversity of this link[20]. Does music evoke emotions in

a listener? Obviously, this question led us to a previously described discussion about definition of emotions. It has been found a common agreement about physiological changes happening due to music, and the same for cerebral reactions or behaviours[35]. Several studies stated that listeners feel “something” listening to music during their experiments: Orne, in 1962, underlined that verbal report of emotional status is not so reliable, since the tester could replace the real emotional significance of the listened music with his own emotion; sometimes it happens that the tester reports a reaction just because of the study expectations [36].

Phyiological response is surely more reliable, and several researches have been showed that: in 1997 Nyklicek and Thayer tried to find evoked reactions (happiness, serene, sad, agitation) from previously rated songs, through respiratory signal. In 1999 Davis and Thaut realized a study to measure physiological response to preferred music by listeners. Testers choose songs to listen and physiological signals were: vascular constriction, heart rate, muscle tension and skin temperature. Results showed an increase in relax condition and decrease in anxiety state during the test with reference to pre and post experiment[37].

In 2001 Knight and Rickard focused again on anxiety evaluating heart rate, blood pressure, cortisol, salivary IgA, in a rest condition and after stressor presentation, causing increase in anxiety, heart rate and systolic blood pressure in control subjects [38]. Other tester were evaluated in the same stressed situation, but soon after listening to music: effects have been shown as significantly inferior, and this suggests and stresses how music can be considered as a good treatment for anxiety reduction.

In 2004 Juslin and Laukka strongly criticized previous works, since they did not consider social context during listening[18]. There are different configurations of emotional responses to music, to evaluate and interpet them in a proper way several features have to be combined and considered, starting from musical aspects, as showed in the Table 1.

Emotion	Musical features
Happiness	Fast tempo, small tempo variability, major mode, simple and consonant harmony, medium-high sound level, small sound level variability, high pitch, much pitch variability, wide pitch range, ascending pitch, perfect 4th and 5th intervals, rising micro intonation, raised singer's formant, staccato articulation, large articulation variability, smooth and fluent rhythm, bright timbre, fast tone attacks, small timing varibility, sharp contrasts between "long" and "short" notes, medium-fast vibrato rate, medium vibrato extent, micro-structural regularity
Sadness	Slow tempo, minor mode, dissonance, low sound level, moderate sound level variability, low pitch, narrow pitch range, descending pitch, "flat" (or falling) intonation, small intervals (e.g., minor 2nd), lowered singer's formant, legato articulation, small articulation variability, dull timbre, slow tone attacks, large timing variability (e.g., rubato), soft contrasts between "long" and "short" notes, pauses, slow vibrato, small vibrato extent, ritardando, micro-structural irregularity
Anger	Fast tempo, small tempo variability, minor mode, atonality, dissonance, high sound level, small loudness variability, high pitch, small pitch variability, ascending pitch, major 7th and augmented 4th intervals, raised singer's formant, staccato articulation, moderate articulation variability, complex rhythm, sudden rhythmic changes (e.g., syncopations), sharp timbre, spectral noise, fast tone attacks/decays, small timing variability, accents on tonally unstable notes, sharp contrasts between "long" and "short" notes, accelerando, medium-fast vibrato rate, large vibrato extent, micro-structural irregularity
Fear	Fast tempo, large tempo variability, minor mode, dissonance, low sound level, large sound level variability, rapid changes in sound level, high pitch, ascending pitch, wide pitch range, large pitch contrasts, staccato articulation, large articulation variability, jerky rhythms, soft timbre, very large timing variability, pauses, soft tone attacks, fast vibrato rate, small vibrato extent, micro-structural irregularity
Tenderness	Slow tempo, major mode, consonance, medium-low sound level, small sound level variability, low pitch, fairly narrow pitch range, lowered singer's formant, legato articulation, small articulation variability, slow tone attacks, soft timbre, moderate timing variability, soft contrasts between long and short notes, accents on tonally stable notes, medium fast vibrato, small vibrato extent, micro-structural regularity

Table 1 - Musical features

In this work too, music ability to elicit physiological changes has been highlighted.

In 2006 Koelsch and Fritz showed different cerebral reactions to unpleasant and pleasant music, with amygdala, hippocampus, parahippocampal gyrus and temporal poles activations, as showed in other works with negative valence stimuli [39].

Several studies underlined how it is essential to consider that emotion related to music is strictly subjective, depending on age, gender, personality, training, preferences, situation and mood[40]. There are many factors to be considered, as Gabriellson said in 2001: physical factors, social ones and occasional circumstances.

It is clear that music and emotions are strongly related, but it is still open the discussion about how they interact and how much strong this link is.

"Music is probably the only real magic I have encountered in my life. There's not some trick involved with it. It's pure and it's real. It moves, it heals, it communicates and does all these incredible things" (Tom Petty)

2.6 Physiological signals for emotion interpretation

An emotion recognition algorithm requires the records of physiological signals.
Nowadays almost all wearable devices can measure four major physiological signals: skin temperature (SKT), skin conductance (SC), heart rate (HR), and electrocardiogram (ECG). Other signals, like electroencephalogram (EEG), facial electromyogram (f-EMG), and blood pressure may be helpful, the attachment of electrodes to the scalp or face is not practical[29][41][42].

Figure 11 - All measures of autonomic functioning

Monovariate analysis of signals is always conducted in order to extract both standard and nonlinear features[22].

2.6.1 Cardiac Signal (ECG)

Figure 12 - Cardiovascular measurements

From ECG signal both time-domain and frequency-domain features can be extracted, since heart electrical activity is affected by ANS changes[43].

Peak detection on the QRS complex with Pan-Tompkins algorithm is firstly performed on ECG signal to extract intervals between successive R-waves and to determine the heart rate (HR): heart rate variability (HRV) is then obtained[22].

2.6.1.1 Heart Rate (HR)

Heart rate (HR) is dually controlled by the sympathetic (HR increase) and parasympathetic (HR decrease) branch of the ANS that may act independently, as a result of the intrinsic automaticity of the sinoatrial (SA) node and the modulating influence of the ANS. Increasing in sympathetic activity becomes progressively more important in accelerating the cardiac rate.

Sympathetic activation is also called the "fight-or-flight response", referring to the preparation to actions, while parasympathetic activation is called "rest and digest response", referring to relaxation as signals for resting and recovery.

All emotions are characterized by an increased HR (decreased RR intervals) with respect to neutral condition. [art. 5]. For HR an higher increase results in anger and fear. HR also increases in surprise, paired to stronger sweat activity, vasoconstriction, activated by sympathetic nervous system [34].

Possible standard features computable from HRV are the following:

- Time domain statistical indexes: SDNN, RMSSD, pNN50
- Spectral (frequency) domain: LF/HF ratio
- Geometric methods: HRV triangular index (TI), TINN, SD1/SD2 ratio

2.6.1.2 Time domain indexes

Time domain features and frequency domain features are extractable from HRV.

Statistical parameters and morphological indexes are included in the time domain analysis: mean value and standard deviation of the NN interval and RMS of successive differences of intervals; pNN50, representing the number of successive differences of intervals which differ by more than 50 ms, expressed as a percentage of the total number of heartbeats; triangular index, reflecting morphological pattern of HRV.

HRV doesn't change in anger, while has significant decrease in fear and happiness[44].

In addition, time-domain features, such as mean and standard deviation (SD) of the HRV

(heart rate variability) time series have also been considered to be significant for the exploration of autonomic nervous system in many previous studies for cardiacfunction assessment and psychophysiological investigation[29].

For total HRV, stress sessions present the highest rates[45].

Time indexes result to be not relevant if investigated in short-time periods. Frequency-domain indexes, on the other hand, turn out to be more meaningful in emotion elicitation/recognition algorithms: HRV and HR indexes can be derived from ECG signal[29].

2.6.1.3 Frequency domain indexes (spectral indexes)

Time domain analysis is moreover unable to discriminate between sympathetic and parasympathetic activations, which is instead achievable by a frequency domain analysis.

Frequency domain features are based on the PSD of the signal HRV, which can be estimated by an autoregressive (AR) model. Once obtained the PSD, 3 main spectral components can be identified: Very Low Frequencies (VLF, below 0.04 Hz), Low Frequencies (LF, from 0.04 Hz up to 0.15 Hz) and High Frequencies (HF, from 0.15Hz up to 0.4 Hz). Thus power, in correspondence of these 3 frequency ranges can be calculated together with LF/HF ratio, which gives useful information about sympatho-vagal balance, by measuring the shift towards a vagal or a sympathetic predominant control.

From baseline to stress, high frequency (HF) absolute values decrease significantly, whereas differences between engagement and relaxation are not significant. The LF/HF graphs are the most interesting, as the highest values are for engagement. The stress epoch values are between engagement and (lowest) relaxation[45].

2.6.1.4 Previous studies

HR increases more in anger and fear, while skin temperature increases in anger. Differences both between positive (happiness) and negative (anger and fear) emotion and among negative emotions were found. Emotion were grouped into 3 sets on the basis of HR and skin temperature as shown in the Figure 13[46].

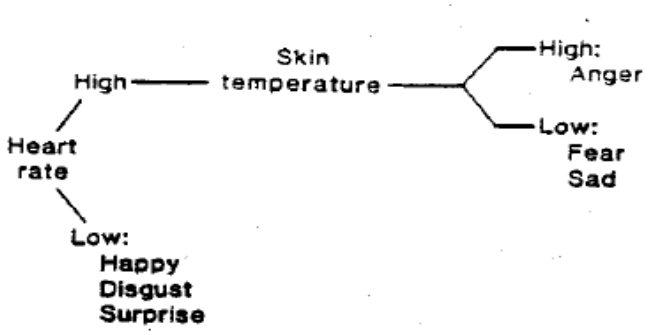


Fig. 2. Decision tree for discriminating emotions in direction facial action task.

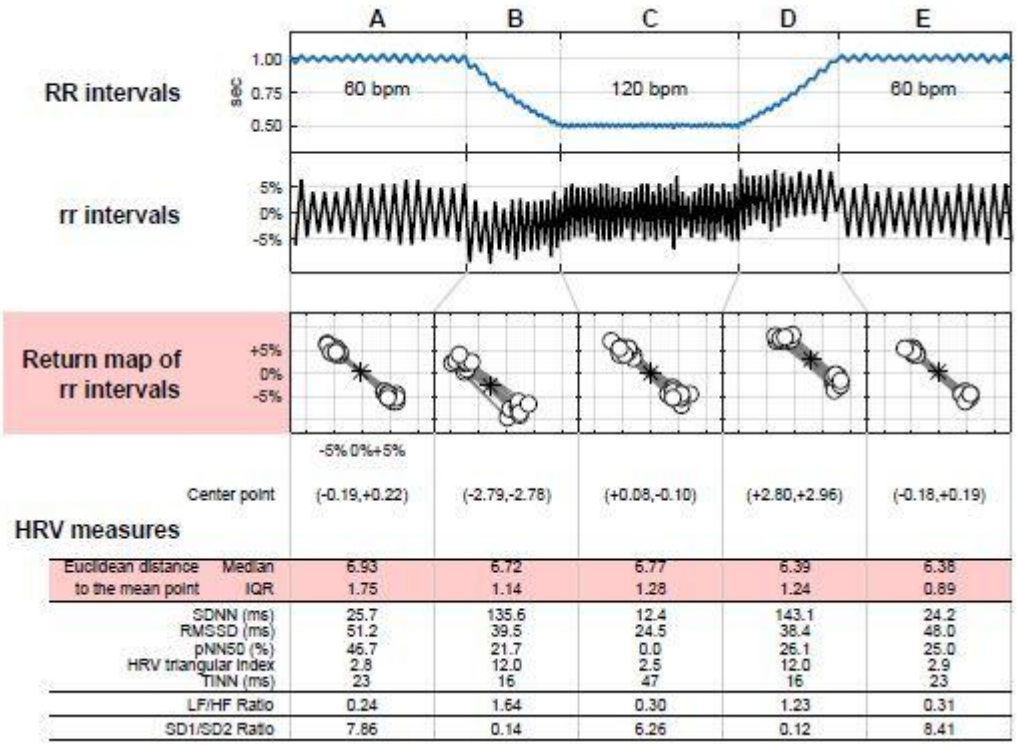
Figure 13- Decision tree for discriminating emotions in direction facial action task

HRV (heart rate variability) is a marker of autonomic system activity. Parameters measures of HRV are very useful to identify diseases or heart failures, allowing to evaluate heart rate in any time or for successive intervals. The starting point is NN intervals series, i.e. intervals between adjacent normal QRS, the ones resulting from sinus node depolarizations (non-ectopic beats).

These measures are numerical and comparable indexes about cardiac rhythm. Time domain analysis is detailed explained in Materials and Methods section (with a better explanation of each index).

The first study about these indexes describes from a general point of view their utility and values, underlining that it could be recommendable to do this kind of analysis on short-term 5 min recordings or long-term ones (24h, as Holter)[47] .

Second report is a brief analysis of these parameters on a one-hour ECG signal, with five sections (as shown below): first part is 60bpm frequency (A), in the second area frequency gradually increases (B), until it reaches 120 bpm (C) in the central section[34]. Then, it slowly decreases again (D), reaching the starting point of 60bpm (E). Signal variability is limited between 5 and -5 %, and it is shown below (Fig. 14) the representation of RR intervals, rr sequences, return map and time indexes values for each section[21].



		A	B	C	D	E
Center point		(-0.19,+0.22)	(-2.79,-2.78)	(+0.08,-0.10)	(+2.80,+2.96)	(-0.18,+0.19)
HRV measures						
Euclidean distance to the mean point	Median	6.93	6.72	6.77	6.39	6.36
	IQR	1.75	1.14	1.28	1.24	0.89
SDNN (ms)		25.7	135.6	12.4	143.1	24.2
RMSSD (ms)		51.2	39.5	24.5	38.4	48.0
pNN50 (%)		46.7	21.7	0.0	26.1	25.0
HRV triangular index		2.8	12.0	2.5	12.0	2.9
TINN (ms)		23	16	47	16	23
LF/HF Ratio		0.24	1.64	0.30	1.23	0.31
SD1/SD2 Ratio		7.86	0.14	6.26	0.12	8.41

Figure 14 - RR intervals, rr intervals and return maps for a task; temporal indexes

2.6.2 Respiration

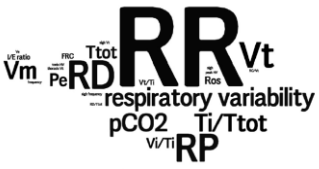


Figure 15 - Respiratory measures

In addition, respiration is very important in the emotion research. From respiratory signals, mean and standard deviation of the first and second derivative are usually computed: these measurements characterize variation of the respiration signal, standard deviation of the breathing amplitude. Respiration rate (RR) and statistical parameters are also calculated, like the maximum and minimum values of breathing amplitude and their difference, skewness, kurtosis and standard error of the mean. Analysis in the frequency domain concern the computation of spectral power in the bandwidths 0–0.1 Hz, 0.1–0.2 Hz, 0.2–0.3 Hz and 0.3–0.4 Hz.

Thus, also respiratory parameters, such as those concerning volume quantification, timing, and shape, are able to map into the affective space dimensions of valence and arousal[29].

So, in this research, four physical signals--SKT, SC, ECG and Respiration are selected to extract features for emotion recognition[29].

Some emotional responses, such as crying, laughing, or shouting, have unique respiratory characteristics. Respiratory period decreases in fear and happiness while respiratory period variability increases in sadness[34].

2.6.2.1 Previous studies

Decreased respiratory period was in previous studies observed in happiness and fear, and less in anger. No significant respiratory amplitude changes were observed but respiratory variability (in terms of respiratory amplitude standard deviation) was observed in anger and sadness and less in fear[44].

2.6.3 Electrodermal Activity

Skin conductance (SC), also called Electrodermal activity (EDA) or Galvanic skin response (GSR), is another important signal representing autonomic nervous system activity, often used in literature[48].

It characterizes changes in the electrical properties of the skin due to the activity of sweat glands and is physically interpreted as conductance. Sweat glands distributed on the skin only receive output from the sympathetic nervous system, and thus SC is a good indicator of arousal level due to

cognitive stimulus[29][49].

About changes in Skin Conductance, two components observation can be done: a tonic component, SCL (skin conductance level), with intrapersonal differences, and a phasic component, SCR (skin conductance response) superimposed on the baseline (represented by the tonic component) changing with external stimulation. Many studies stated that changes in electrodermal activity are linearly correlated to the intensity of the felt emotion with the arousal level, reason why EDA is considered an index of sympathetic activity[21].

Figure 16 - Electrodermal measures

An increased Skin Conductance signal (namely a drop-in skin resistance) is detected in correspondence with a sympathetic activation and reflects cognitively or emotionally mediated motor preparation[43][11]. An example of GSR signal is shown in Figure 17.

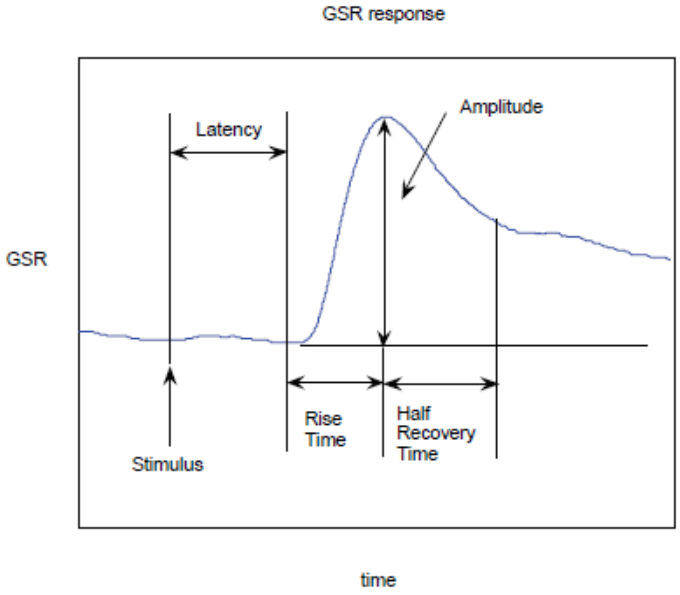


Figure 17 - Example of recorded GSR signal, showing a single Skin Conductance Response (SCR)

2.6.4 Other signals

Below, we report other signals often used in literature to help modeling emotion recognition algorithms.

- *SKT*: it is an important and effective indicator of emotion states, and it is mainly influenced by sympathetic adrenergic vasoconstrictor nerves, reflecting autonomic nervous system activity. [art.8] SKT extremely decreases in stress and fear, increases in relaxation, boredom and sleep and mildly increases in pain and surprise
- *EEG*: electrical activity of the brain produces the EEG signal, measured by applying electrodes on the scalp of the patient. [art.44]
- *Pneumograph*: the breath of a patient reflects the ANS shifts: collection of respiration pattern data requires a pneumograph, which is a respiration transducer placed on a belt around the chest. [art.44]
- *BVP*: changes in the BVP signal could indicate relative changes in the vascular bed due to vasodilatation or vasoconstriction (increase or decrease in blood perfusion) as well as changes in the elasticity of the vascular walls, reflecting changes in blood pressure [35]. BVP decreases in surprise and greatly decreases in pain: decrease in BVP

amplitude during pain compared to the baseline state might be implying peripheral vasoconstriction in the finger associated with arousal [art. 1]

- *PPG*: Variations in Blood Volume Pulse (BVP) can be measured through the photo-plethysmograph (PPG), an infrared functioning device to place on the finger of the patient: these signals provide information about ANS functioning too. [art.44] . PPG is a non-invasive technique and measures the blood volume in skin capillary. This signal is usually used to estimate heart rate variability (HRV), since a decrease in skin blood flow is an indicator of sympathetic reflex response to stimuli, caused by vasoconstriction in hairless area of the hand, so thus fingers [art.44]. An example of PPG signal is shown in Figure 18.

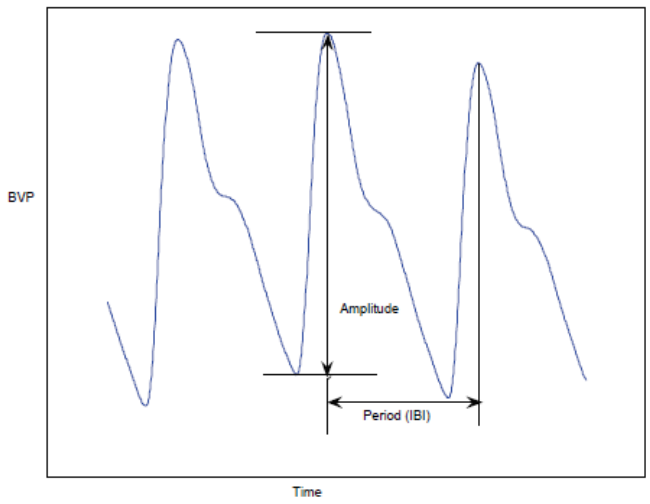


Figure 18 - Example of recorded BVP signal, showing three cardiac cycles (beats)

- *PUPIL_DIAMETER (PD):* Pupil diameter variation pattern can be measured through a webcam.[43]. This measure is determined by the contractile activity of the sphincter and the dilator papillae, two opposing groups of muscles in the iris, governed by the ANS in humans. This measure is basically determined by the amount of light and by the accommodation reflexes. Furthermore, researchers found a correlation between pupil diameter variation and emotional states, seen as a cognitive information process. Pupil diameter and point of gaze can be detected with infrared aye-tracking systems[43].

2.6.5 Bivariate analysis: CR coupling

Bivariate analysis is often performed between cardiovascular and respiratory signals in order to extract other significant information.

Another work attempt to extend a previous study developed by Valenza et al. on an emotion recognition system based only on monovariate analysis, by embedding bivariate analysis.

Here, CR coupling is introduced in an innovative way, by providing it from features extracted with a mono-variate analysis: this CR coupling thus refer to a phase synchronization between cardiac and respiratory systems[22].

Including CR coupling the method actually became able to identify the correct affective states referring to the CMA, thus improving the previous achieved recognition of 5 levels of both arousal and valence. More accurate allocation of emotions into the 25 different affective regions of the valence-arousal space was therefore obtained by using both mono-variate and bivariate features[22].

Physiologically, the crucial point of the CR coupling comes from the overlap of the power spectra of the two interacting systems: spectral peaks of both signals may become almost equal and peak frequencies become closer due to the synchronization of the systems[22].

2.6.5.1 Previous studies

In a study, the contribution of cardiorespiratory activity to the production of different somatic states linked to the feeling of different target emotions in different subjects was tested[44].

Subjects were asked to choose and recall an autobiographical episode eliciting him a strong specific emotion: cardiac and respiratory activity were recorded and analyzed starting from the moment the subject was feeling that specific emotion[44].

Several measures of both HR and respiration were extracted. Through PCA the independent factors were then extracted to be used in discriminant analysis to verify their contribution in predicting motions[44].

Specific basic emotions can be predicted from the observed pattern of cardiorespiratory activity: Cardiorespiratory activity associated with emotions can be described through two

influences, sympathetic and parasympathetic, and the parasympathetic influence can be further distinguished into respiration-coupled and respiration-uncoupled components[44].
In the same study, an effort to discriminate the four investigated emotions was done, primarily through a pair of emotions comparison: anger was clearly separated from both fear and happiness thanks to the combination of mean respiration period and RR-High Frequency, similarly, anger and fear were distinguished with an univariate approach for both dependent variables (Fig. 19).
However, it is noteworthy that happiness did not differ significantly from anger on either Respiration Period-mean or RR-High Frequency: separation between these two emotions was achievable only when considering the bi-dimensional space, thus overcoming the limits of univariate analysis[44].

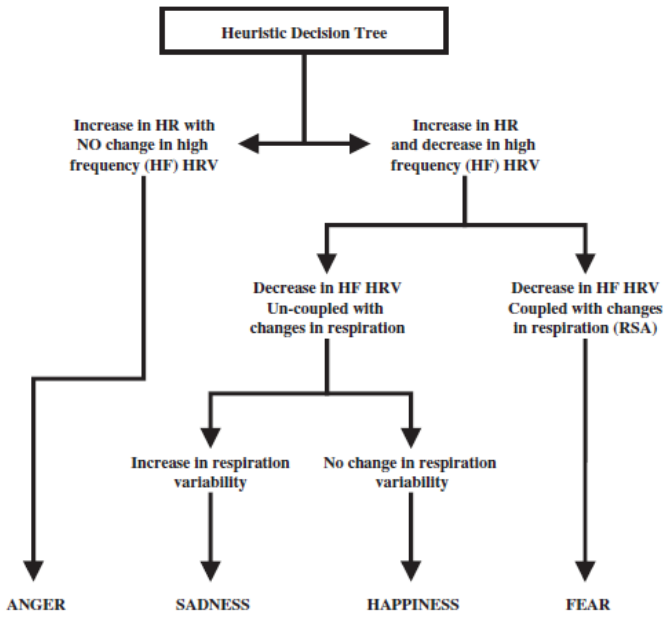


Figure 19 - Heuristic decision tree derived from the observations of distincts patterns of cardiorespiratory activity during basic emotions[44]

For each bio-signal, mono-variate analyses were applied in order to extract significant features using both standard and nonlinear techniques. Moreover, coupling measures by means of a bivariate analysis have been extracted from the RR interval series along with the RSP. Then, the obtained feature space dimension was reduced using the principal component (PCA) analysis method. Finally, features were classified using various machine

learning methods. QDC showed the highest recognition accuracy and consistency in both arousal and valence multi-classes (Fig. 20)[22].

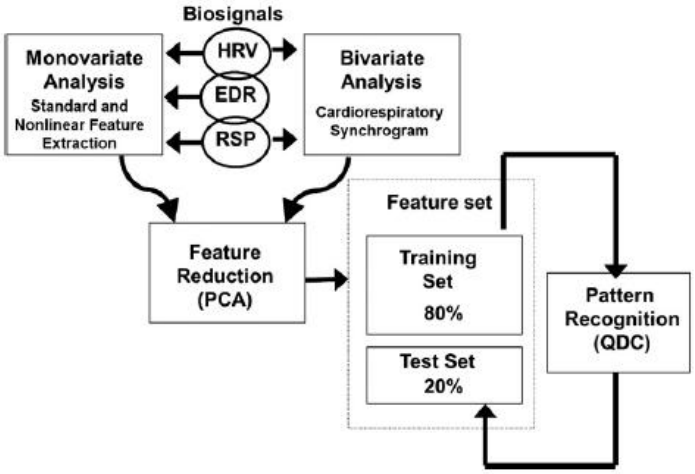


Figure 20 - Block diagram representing a possible acquisition and processing chain

2.7 Emotion discrimination

Reminding that increased sympathetic and parasympathetic activations is caused by different stimuli and cause different physiological and behavioral reactions, a correlation between the kind of autonomic nervous system activation and the changing physiological signals can be established, in order to achieve the aim of each emotion recognition algorithm: discrimination between different emotional condition[43][50][51].

In literature, a complete work achieve a detailed distinction between many different emotions in terms of autonomic nervous system physiological responses, analyzing the features extracted from signals[7].

Figure 21 - Investigated emotions

2.7.1 Negative emotions

- *Anger: m*odal response pattern results in a reciprocal sympathetic activation and increased respiratory activity (faster breathing, namely shorter Ti and Te). Increased HR and SBP. Decreased HRV indicates a cardiac parasympathetic inhibition, and increased EDA is the result of the sympathetic activation
- *Anxiety:* characterized by a sympathetic activation and vagal deactivations, which form a pattern of reciprocal inhibition, combined with faster and shallower respiratory activity. HR increases, HRV decreases, LF and LF/HF ratio increases, and SBP and EDA (SCR) too. Decreased Ti and Te in respiration signal also indicate increased Respiration rate (RR). Some IAPS standard pictures make an exception and report decreased HR and increased HRV: pictures showing snakes, sharks, knife or attack.
- *Disgust*: a distinction between disgust elicited with images of contamination and pollution and disgust elicited by mutilation, injury and blood must be made. For the first kind of disgust, a sympathetic-parasympathetic co-activation, HR acceleration, increased HRV and faster breathing are present, while the second kind of elicitation is characterized by a pattern of sympathetic deactivation, HR deceleration, no changes in HRV and faster breathing.
- *Embarrassment:* induced by experimenter humiliation. Studies revealed a broad sympathetic activation and vagal withdrawal composing a pattern of reciprocal inhibition.
- *Fear:* elicited by showing threatening pictures, film clips, standardized images, personal recalls or real-life manipulation. An increase in HR can be observed, due to vasoconstriction.
- *Sadness:*its elicitation evokes heterogeneous pattern of sympathetic-parasympathetic co-activation. Responses to sadness elicitation must be distinguished in two classes: an activating response and a deactivating response. Activating sadness (also called crying sadness) implies increased cardiovascular sympathetic control (increased HR) and changes in respiratory activity, while deactivating response (or non-crying sadness) is characterized by a sympathetic deactivation, decreased HR, increased or unchanged HRV, and can be distinguished by all other negative emotions by its decrease in EDA. Sadness elicited by means of standardized images depicting scenes of catastrophe,

soldiers in action, dead animals usually implies increased HR and unchanged HRV and ventilation. Decreased HR and Ti, increased Te and HRV, unchanged SCR for depressive pictures such as cemetery, plane crash, war victims.

2.7.2 Positive emotions

- *Affection*:it is identified by a decreased HR, unspecific increase in SCL and increased variation in respiratory activity.
- *Amusement*:it is usually elicited by film clips and the overall response shows an increased vagal control, (and alfa adrenergic) respiratory and EDA coupled with sympathetic deactivation (and beta adrenergic). HR has a much more variable response and in different studies result increased, or decreased or unchanged. Increased HRV can be seen and unchanged LF/HF
- *Contentment*:results in decreased cardiovascular and respiratory activity, low EDA, indicating mild cardiac vagal activation like in amusement, and sympathetic deactivation, stronger than in amusement. So decreased HR and RR, decreased or unchanged HRV and SCL were reported when contentment was elicited. Furthermore, the use of music for emotion induction brings to increased RR and decreased HRV.
- *Happiness*:it is characterized by an increase in cardiac activity coupled with a vagal withdraw, vasodilatation, increasing EDA and respiratory activity. Increased or unchanged HR, decreased or unchanged HRV and unchanged LF, increased EDA and RR can be observed. Nevertheless, decreased HR and slightly increased HRV were observed in some studies in which emotions were elicited by using IAPS images. The variance presented in physiological response when eliciting happiness is due to the fact that with the term “happiness” also admiration, contentment, excitement, joy and pride are sometimes included.
- *Joy*: increased vagal activity and increased cholinergically mediated sympathetic activity and respiratory activity characterize joy manifestation. Joy is the only one, among positive emotions, showing an increase in beta adrenergic sympathetic activation, associated with a major motivational engagement. An increase in HR, SCL and HRV and RR can be seen.

- *Pleasure, anticipatory*: under these terms, appetite and sexual arousal are included and physiological responses concern physiological deactivation when standardized images are shown to the subject, resulting in increased vagal activity, EDA and in decreased RR. Increased HR and SCR are present when erotic pictures are shown.
- *Pride*: it results in decreased HR and unchanged HRV, and in general in increased cholinergic sympathetic activity and unchanged vagal control.

Other emotions, such as surprise and suspense doesn't show a clear valence characteristic pattern: just an HR increase and decrease were observed in some studies for surprise and suspense respectively.

Summing up all the described evidences, HR show an increase in negative emotions (anger, anxiety, contamination-related disgust, embarrassment, fear, crying sadness) and In some positive emotions (happiness, joy and surprise), as well a decrease in mutilation-related disgust, imminent-threat fear, non-crying sadness, affection, contentment, anticipatory pleasure, suspense, all of those involving a passivity characteristic typical of vagal control. Contamination related disgust was the only negative emotion eliciting an increase in vagal control, visible thanks to the increased HRV. HRV also increases in amusement and joy while decreases in happiness and anticipatory pleasure: however, all show increased EDA and faster breathing. Decreased HR and LF/HF ratio were detected in mutilation related disgust.

For a long time, convenience measures were exclusively used to study ANS response to emotion inducing stimuli: these convenience measures were HR and electrodermal activity, used alone to indicate the activation state of the subject. Later, some other quantifiable indices were included for this kind of analysis, to quantitatively evaluate both cardiovascular, eccrine and respiratory response.

Decreased EDA was pointed out only in non-crying sadness and contentment. All the other emotions reported an increase in EDA, which reflects the emotionally mediated motor preparation, which indeed is not present when sadness is experienced. In Table 2, physiological changes overview is shown.

		anxiety	disgust (contamination)	disgust (mutilation)	sadness (non-crying)	amusement	contentment	happiness	joy
cardiovascular	HR	↑	↑ –	↓	↓	↓↑	↓	↑	↑
	HRV	↓	↑	–	↓	↑	↓↑	↓	(↑)
	LF	(↑)							
	LF/HF	(↑)		↓		–			
electrodermal	SCR	↑	↑	↑		↑	–		(↑)
	SCL	↑	↑	↑	↓	↑	↓	↑–	–
respiratory	RR	↑	↑	↑	↑	↑	↓↑	↑	(↑)
	Ti	↓	↓	–		(↓)	(↑)	↓	
	Te	↓	↑	–			(↑)	↓	
ANS resp. Comp.	α-adrenergic	(↑)	↑	–	(↑)	↑–	(↓)	(↑)	–
	ß-adrenergic	(↑)	(↑)	(↑)		(↓)	(↓)	↓	↓↑
	cholinergic	↑	↑	↑	↓	↑	↓	↑–	↑–
	vagal	↓	↑	–	↓	↑	↓↑	↓	(↑)
	Respiratory activity	↑	↑	–	↑	↑	↓	↑	(↑)

Table 2 - Physiological changes in different emotions

2.8 Signal analysis: point process

In .2003 a new model of human heartbeat intervals was defined: the point-process model. It gives a new definition to both heart rate (HR) and heart rate variability (HRV). This probability model takes into account the stochastic structure/nature of human heartbeats considering it an history-dependent process, and deriving from this, a probability density function [52].

More detailed explanation about this algorithm can be found in chapter "Materials and Methods".

2.9 Statistical analysis and Classification

In 2004 Kim and colleagues [Kim et al., 2004] developed an emotion recognition system which used short-term monitoring of 4 physiological signals: ECG, BVP (measured through PPG), SKT (skin temperature or EDA). [art. 44]

Support Vector Machine classifier, and this group reported a

correct classification ratio of 78.4% in identifying instances of "sadness", "anger" and "stress", and 61.8% in identifying instances of "sadness", "anger", "stress" and "surprise" [43].

In 2003, Barreto and Zhai, using the Stroop test to elicit stress and monitoring 4 physiological signals, derived 11 different features which were used to differentiate between non-stress and stress sessions in the test. Then, Naïve Bayes classifier; a Decision Tree classifier, and a Support Vector

Machine classifier were used: the best classification was obtained with SVM classifier (90.10%)[43].

Arousal and valence multi-class recognition can be performed by applying a Bayesian classifier (that uses quadratic discriminant classifier QDC) on extracted features[22].

3. Materials and methods

3.1 Experimental protocol

3.1.1 Subjects

Participants were 105 healthy volunteers, recruited from the Politecnico di Milano bachelor and master course in Biomedical Engineering.

Within these 105 subjects, 56 were men and 49 women, all aged between 20 and 27. Subjects were divided in two groups according to the presentation shown them. Table 3 shows the division of the subjects, number of males and females in each group and related age (mean ± standard deviation).

Subjects	Presentation protocol	Gender		
		male	female	Tot.
	n. 1	30	21	51
	n. 2	26	28	54
	Tot.	56	49	105

Age (years)	Presentation protocol	Gender		
		male	female	Tot.
	n. 1	24,2 ± 1,38	23,73 ± 1,14	24 ± 1,30
	n. 2	23,8 ± 1,28	23,2 ± 1,00	23,52 ± 1,18
	Tot.	24 ± 1,33	23,43 ± 1,08	23,75 ±1,25

Table 3 - Subjects for the present study

The majority of the subjects doesn't show pathological inheritance concerning cardiovascular diseases, respiratory disorders or neural system disorders.

Except for two subjects, a man and a woman, showing bradycardic (low heart rate) and tachycardic (rapid heartbeat) phenomena, respectively.

The experiment was carried out in the SPiNLabS Laboratory in Politecnico di Milano, Italy. Signal acquisition took place between September and October 2017.

Subjects were first contacted and were scheduled, according to their availability, to participate in the experiment after filling out a short questionnaire (just one minute required to fill it out). The Patient Health Questionnaire-9 (PHQ-9) is a short tool self-administered, specifically developed for its usage in Primary Health Care (Spitzer et al.,1999).

This tool can be used for the screening, for diagnosis, to monitor and measure the seriousness of the eventual depressive state of the subject. A detailed explanation is present in the dedicated section.

Before performing the experiment, the subject is told about experimental setting, in terms of duration, acquired signals and presentation structure. Then, the sensors were applied to the patient: three electrodes for the ECG signal acquisition, a respiratory abdominal belt and two Velcro bands secured around two fingers embedding two sensors for the Skin Conductance signal acquisition.

Before starting the test, the patient was asked to calm down, in order not to strain muscles for the following 10 minutes (full test duration), choosing a comfortable and proper position on the chair, located in front of the PC screen on which the presentation was projected.

Then, a proper verification of the correct positioning of electrodes and cleanness of signals was performed, by a visual check on the screen showing the output signals in the acquisition software interface.

Then lights were turned off, the subject was asked to remain as still as possible, and the test started.

3.1.2 PHQ-9 Questionnaire

The PHQ-9 Questionnaire is a self-administered version, used as a diagnostic instrument for common mental disorders. This module uses scoring from 0 (“not at all”) to 3 (“nearly every day”).

Validity has been assessed on a 6000 patients sample in eight different clinics for primary health-care and in seven obstetrics-gynecology clinics.

It is a new instrument for the recognition of depression and other mental disorders, based on daily life activities and sensations. Even if shorter than other questionnaires and instruments for measuring depression, it has comparable specificity and sensitivity values.

Built on the nine criteria used in the type DSM-IV depressive disorders diagnosis, the relevance of this instrument lies in its dual ability in establishing a diagnosis and the depression degree too.

The questionnaire is three pages long and can be simply self-ministered to the patient. Once filled out, the clinician checks the answers and evaluates the total score, applying the diagnostic algorithm later described (Fig. 22).

Eight different diagnosis can be defined according to different thresholds: major depressive disorder, panic disorder, other anxiety disorder, bulimia nervosa, and others. The worst-case is that one related to major depression, recognized when five or more symptoms have occurred at least “more than half the days” in the past two weeks and one of the symptoms is the depression mood or the anhedonia.

Other depressive disorders are diagnosed if two, three or four symptoms have occurred at least “more than half the days” in the past two weeks and one of the symptoms is the depression mood or the anhedonia.

One among the nine criteria (“thoughts that you would be better off dead or of hurting yourself in some way”) has its own validity, regardless its duration.

Moreover, before actually making a diagnosis, we must ask the patient about any other possible physiological symptom or bipolar overreactions.

The total score of the test can range from 0 to 27, since each answer can be scored from 0 to 3.

A final question is then added, to understand how much these symptoms are actually affecting the patient’s daylife; the question is the following: “How difficult have these problems made it for you to do your work, take care of things at home, or get along with other people?”.

The Patient Health Questionnaire (PHQ-9)

Patient Name _____ Date of Visit _____

Over the past 2 weeks, how often have you been bothered by any of the following problems?	Not At all	Several Days	More Than Half the Days	Nearly Every Day
1. Little interest or pleasure in doing things	0	1	2	3
2. Feeling down, depressed or hopeless	0	1	2	3
3. Trouble falling asleep, staying asleep, or sleeping too much	0	1	2	3
4. Feeling tired or having little energy	0	1	2	3
5. Poor appetite or overeating	0	1	2	3
6. Feeling bad about yourself - or that you're a failure or have let yourself or your family down	0	1	2	3
7. Trouble concentrating on things, such as reading the newspaper or watching television	0	1	2	3
8. Moving or speaking so slowly that other people could have noticed. Or, the opposite - being so fidgety or restless that you have been moving around a lot more than usual	0	1	2	3
9. Thoughts that you would be better off dead or of hurting yourself in some way	0	1	2	3
Column Totals	_____ +	_____ +	_____	
Add Totals Together	_____			

10. If you checked off any problems, how difficult have those problems made it for you to Do your work, take care of things at home, or get along with other people?

☐ Not difficult at all ☐ Somewhat difficult ☐ Very difficult ☐ Extremely difficult

Figure 22 - The PHQ-9 questionnaire

Validity and reliability of this test have been studied for years in the late years of 20th century, between 1997 and 1999, by comparing the outcomes of the questionnaire to evaluations obtained through other approaches or with direct interview with psychologists and doctors. After actually evaluating the significance of this method, it was decided to apply a partitioning into categories by relying on the scores. Growing levels of "severity" of the disorders were established: 0-4, 5-9, 10-14, 15-19, 20 or more. The categories have been chosen for several reasons, the first is a pragmatic one, because this method ensures the ease in the classification with scales the cut points of those are 5,10,15,20, easy to memorize and apply for clinicians. Moreover, from an empirical point of view, this kind of

association between score and disturb degree, doesn't affect the meaning of the evaluation. Anyway, as told before, all of these decisions and evaluations are conducted through psychologist and doctors' consultation, applying different considerations and statistical studies too.

The "Depression severity" level, linked to different scores, turns out to be: Minimal, 0-4; Mild, 5-9; Moderate, 10-14; Moderately severe, 15-19; Severe, 20-27.

In conclusion, the PHQ-9 questionnaire embeds the criteria for DSM-IV depression with other depression symptoms in a short self-report tool, with a scoring for the evaluation of both the occurrence of symptoms and the presence and duration of self-injuring or suicidal ideas. It offers substantial temporal advantages, as well as ease in the actual compilation and evaluation.

The PHQ-9 with a score equal or greater than 10, has a sensitivity of 88% and a specificity of 88% for the Major Depression.

The total score is obtained by the sum of the scores of the single answers (column 0 = 0 points, column 1 = 1 point ("several days"), column 2 = 2 points ("more than half the days"), column 3 = 3 points ("nearly every day")).

Figure 23 shows the relative diagnosis and the recommended treatment by relying on the result of the test [53].

PHQ-9 Score	Provisional Diagnosis	Treatment Recommendation *Patient Preferences should be considered*
5-9	Minimal Symptoms*	Support, educate to call if worse, return in one month
10-14	Minor depression ++ Dysthymia* Major Depression, mild	Support, watchful waiting Antidepressant or psychotherapy Antidepressant or psychotherapy
15-19	Major depression, moderately severe	Antidepressant or psychotherapy
>20	Major Depression, severe	Antidepressant and psychotherapy (especially if not improved on monotherapy)

** If symptoms present ≥ two years, then probable chronic depression which warrants antidepressants or psychotherapy (ask "In the past 2 years have you felt depressed or sad most days, even if you felt okay sometimes?")*

++ If symptoms present ≥ one month or severe functional impairment, consider active treatment

Figure 23 - The PHQ-9 scores

3.1.3 Visual stimulation

The three physiological signals mentioned above are acquired while presenting different stimuli to the patients.

The stimuli are both visual and auditory and are provided by a presentation on a PC screen that the patient was asked to watch, remaining as still as possible.

Two different presentations were created and tested on an equal number of subjects.

Each presentation lasts 10 minutes and consists of different sessions: the differences are the duration of the first "relax" session as well in the order of the presentation of the four stimuli.

The first session, the one we call "relax" session, was established to let the patient calm down before starting with the arousal sessions. This first session lasts 40 seconds in the first presentation while 110 seconds in the second one.

Then, the sequence of the four arousal sessions follow, alternated with neutral sessions.

Each neutral session lasts 30 seconds while the arousal sessions lasts 90 seconds.

In the first presentation (Fig. 24) the order of the evoked emotions is the following:

1) Stressed
2) Elated
3) Excited
4) Depressed

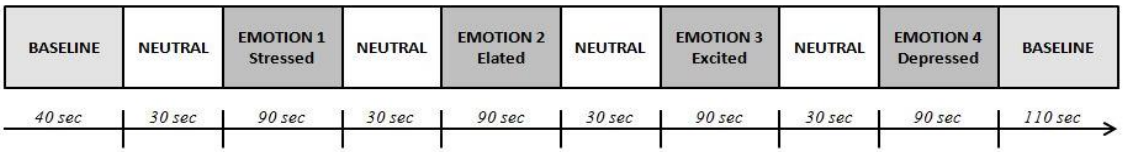


Figure 24 - Presentation 1

For the second presentation (Fig. 25) the first and the last emotion were though switched, so the order becomes:

1) Depressed
2) Elated
3) Excited
4) Stressed

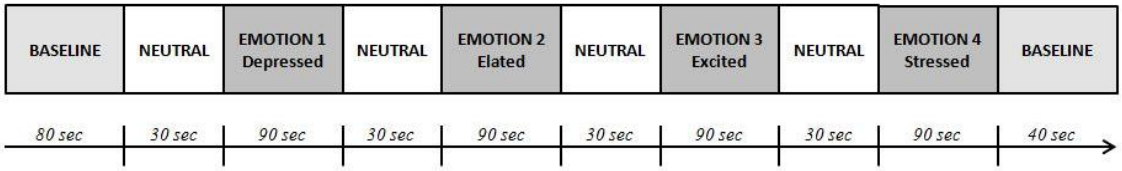


Figure 25 - Presentation 2

Then the presentation lasts with another final "relax" phase, in order to obtain a baseline reference signal; this last session is 110 seconds long in the first presentation and 40 seconds long in the second one.

The stressed, elated and depressed sessions consist each one in a sequence of 10 pictures (9 seconds each one) paired to a background music, properly chosen in order to increase the stimuli provided by the pictures.

In the excited session the pictures have been replaced by an enthralling video. Thus, the neutral phase was a sequence of three (10 seconds each one) meaningless pictures and the relax sessions consists of a simple black slide: both the neutral and the relax sessions are not paired to any musical background but silence.

The whole signal acquisition procedure is performed with the lights out and music is listened by a pair of headphones: this setting allows to improve the isolation and the concentration of the patient. These tests took place in one month and a half of constant and accurate work, through an accurate planification which allow us to optimize times, so that we were able to record 105 acquisitions in less more than a month, considering a 30 minutes lasting time-slot for each subject, to ensure a comfortable and paceful position before projecting the presentation. Sensors and electrodes must be properly positioned and functioning must be checked, then the patient must remain as still as possible for the whole duration of the presentation.

3.1.4 Visual Emotional Stimulation through images: IAPS vs GAPED

It is known that using the visual channel to convey emotional stimulation is the most common and prevalent in research thanks to its convenience and ease and to its relation to

the nowadays well-known visual path physiology and processes. Visual path stimulation can be both static or dynamic according to what is used between pictures and videoclips.

The great part of the works in the Affective science research area uses the *International Affective Picture System (IAPS)*, a picture database created to permit better control in the selection of the induced stimuli and to gain a reliability and repeatability when comparing and replicating different experiments performed with the same material. Each single IAPS image is associated to a standardized rating of valence and arousal.

Although, we decided to use another picture database, the *Geneva Affective Picture Database (GAPED)* to obtain a more reliable induction of expressive and physiological emotion response with these stimuli, taking advantage of the main feature distinguishing the two databases.

The GAPED database is thought to be able to induce emotion according to social relevance and norm significance. Both two main emotional categories in which the GAPED database is divided, negative and non-negative, have been used in our work.

Within the negative category, pictures representing human concerns (depicting scenes violating human rights) are selected to elicit "depression", while pictures representing animal and human mistreatment are used for the "stress" session.

Within the non-negative category, a distinction between positive and neutral is present.

Among positive pictures we chose those representing human babies, young animals and landscapes to elicit "elation". Neutral images, such as those representing objects, buildings and furniture are though used for the neutral sessions, alternating to the "arousal" ones.

GAPED pictures are as well rated in terms of valence and arousal.

Reminding that the arousal value rates the images in term of activation and deactivation (high and low arousal values respectively) while the valence is linked to the pleasant or unpleasant influence of an image (high and low valence values respectively), the graphical representation of the circumplex model of affect is used.

A scatter plot can be used to show the distribution of the images, used in the present work, in a valence-arousal coordinate plane. In this chart, the horizontal axis, representing the valence value, and the vertical axis, representing the arousal value, both range from 0 to 100.

Pictures belonging to different emotional sessions are concentrated in specific parts of this plane:

- *"Depression" session:* upper -left quadrant of the chart (low valence, high arousal)
- *"Elation" session:* lower right part of the 4th quadrant (high valence, low arousal)
- *"Neutral" session:* lower central-right part of the chart, mainly within the 4th quadrant (medium valence, low arousal)
- *"Stress" session:* left upper part of the 2nd quadrant (very low valence, high arousal)

A distribution in the Valence-Arousal plane of the GAPED images that we used in the present study, is shown in Figure 26.

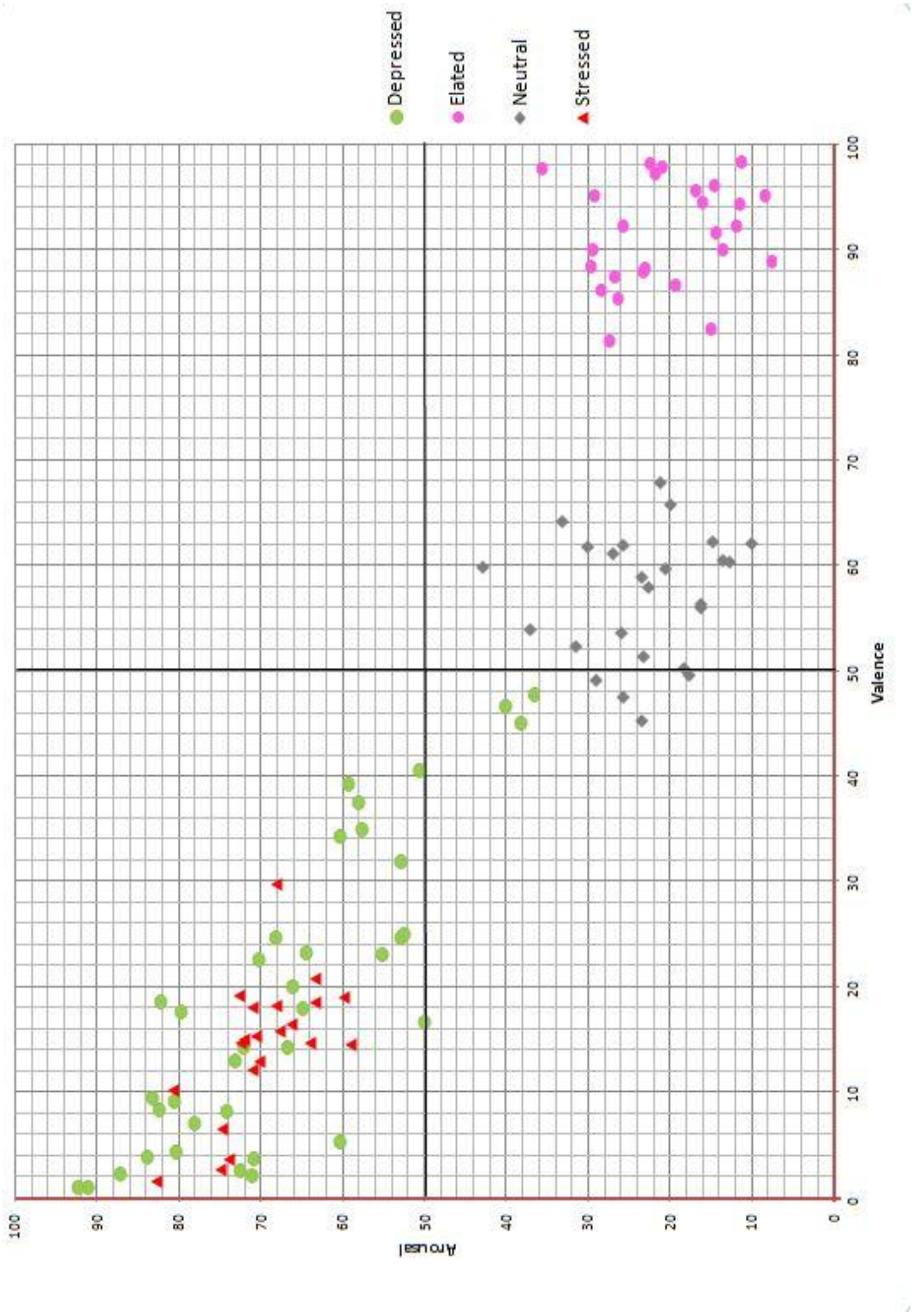


Figure 26 - GAPED distribution of a selection of images on the valence-arousal plane

3.1.5 Benefits of GAPED Database

Within the "State of the Art" chapter and in the previous section, we extensively described the study which lead to the extensive usage of the IAPS images, as well as the basis which drove the Geneva research center to the formulation and to the subsequent creation of a new pictures database.

In a previous study, conducted by a student team from the Bachelor in Biomedical Engineering, and supervised by us, (bibliografia TESI WEMOS), an eight minutes lasting protocol has been used, composed by 8 alternating emotional and neutral sessions, consisting in the presentation of an images sequence chosen from the IAPS database. The obtained results, evaluated through the accuracy of the classifier, raised to 53%, giving rise to different considerations on the validity of the protocol.

The work performed by the WeMos project group was accurate and careful both in signals acquisition and analysis, and in software and hardware programmation, but did not provide proper results. Together with them, we found the possible changes applicable to their protocol and to their processing in order to update and strengthen the results. The starting point has been for sure the setting of a longer lasting protocol, so as to include a signal acquisition session without any visual stimuli, in order to provide a reliable baseline for use as reference; moreover, we decided, after a comparison with the mentioned before students, to change the picture database, choosing for the latest one, realized by the center of Geneva, using images aimed at eliciting more precise negative reactions in the tester.

Within our presentation, therefore, we inserted the GAPED database images, to structure the sessions of the stress and depression elicitation on which the Swiss researchers focused on to find powerful images, but especially inherent to the socio-cultural environment in which nowadays we are developing ourselves.

Images depicting mistreatment, sadness, misery, desolation, mutilations and death are included. Elise S. Dan-Glauser and Klaus R. Scherer, with this work in the 2012, tried to modernize and contextualize human reactions, using more recent images with better definition and quality, a relevant and significant innovative issue, that should not be underestimated. In that direction, other research studies are moving too, as well as the recently released one from the OASIS in 2016, about which we talked in the previous chapter [54].

Moreover, as we recently explained, we decided to add a musical background to the emotional sessions. Last but not the least, we decided to completely change the "exciting" session. The IAPS database would provide, for this session, erotic pornographic pictures, depicting naked men and women, as well as sexual acts, many years ago considered "forbidden/banned".

Although, the IAPS database study is almost twenty years old, twenty years of technology development, Internet, reality Shows. Society is changing, our habits and routine too, the background is developing. Nowadays, sex is not a taboo any more, especially for university students. Images, as well as videos, movies, sex dedicated programs, are almost everywhere in the net; availability and affordability of pornographic material no longer represents an issue, and besides it, talking about it, doesn't worry or embarrass like it used some years ago[55] [56].

Our sex-related reactions quickly changed, as well as the context within which we grow up. Hence, a naked man, a sexual act, a "fake" provocative 2D vintage image does not evoke a surprising reaction as it could be 20 years ago.

That is why we decided to replace the images planned for this protocol "exciting" session, with a videoclip, whose utility and description, as well as significance of evoked emotive involvement, are going to be explained in the following part.

The choice of the images to be included in the protocol hasn't been a trivial task: before starting our study, we co-operated with another students group (WeMos) in order to define a protocol with similar characteristics but based on IAPS database. We established our study looking for an innovation meaningful and successful. GAPED database has been properly evaluated and analyzed, trying to find those images that would have been able to elicit a greater stimulation in terms of both valence-arousal scale classification and our interpretation. Thus images have been chosen on the basis of their quantitative value, as described by the Geneva center of studies, and for their qualitative value too evaluated considering the elicited impression. Furthermore, the video, replaced the exciting session, was edited with a meticulous procedure, trying to use those meaningful frames able to elicit a more powerful and enthralling sensation with respect to IAPS images.

3.1.6 Video Exciting session

The decision to replace the "exciting" images with a videoclip, arises therefore from the need and the intention to evoke a powerful reaction in the patient. The idea was then to rely on a sequence of clips caught from videos published online on GoPro website (https://it.gopro.com/channel/), concerning different extreme sports, acrobatic activities, selected and later assembled using iMovie, an Apple released software (Fig. 27) . By now, we all already know about GoPro and all the video characteristics recordable by means of these cameras. Firsthand experiences, lived from a gripping, fascinating, original perspective, which ensure more steeply results. But everything, obviously, passes through the nature of these sports, the idea of "extreme", and the physiological related reactions.

Everything start from adrenaline.

A paper released in 2006 by the University of Maryland Medical Center, titled "Endocrinology Health Guide", by Justin Anderson, PsyD, sport consultant for the Center for Sport Psychology, in Denton (Texas), explains in an easy and proper way, the subtle but strong line connecting adrenaline to extreme sports.

"The emotion that the adrenaline feeds into is a heightened sense of being alive," Anderson tells. "All your senses are in an acute level of awareness, and it's that fight or flight response. They either do it and live or they die. That is what they are playing into, and that is a very primitive thing that is going on." When it comes to extreme sports, the adrenaline factor likely plays a role in explaining why athletes reach for the outer limits as well.

Adrenaline, or epinephrine, is a stress hormone secreted from the adrenal glands on the kidneys. It plays a major role in preparing the body for a fight-or-flight reaction in threatening environments. An adrenaline rush is a sudden increase in the secretion of adrenaline from the adrenal glands. This happens when the brain communicates to the glands that there will be a need for a fight-or-flight response. The cause of an adrenaline rush need not be an actual physical threat but can also be an imagined threat, strenuous exercise, heart failure, chronic stress, anxiety or a disorder of the brain or adrenal glands.

When you perceive something as threatening or exciting, the hypothalamus in the brain signals to the adrenal glands that it's time to produce adrenaline and other stress hormones. The adrenal glands produce adrenaline by transforming the amino acid tyrosine into dopamine. Oxygenation of dopamine yields noradrenaline, which is then converted into

adrenaline. Adrenaline binds to receptors on the heart, arteries, pancreas, liver, muscles and fatty tissue. By binding to receptors on the heart and arteries, adrenaline increases heart rate and respiration, and by binding to receptors on the pancreas, liver, muscles and fatty tissue, it inhibits the production of insulin and stimulates the synthesis of sugar and fat, which the body can use as a fuel in fight-or-flight situations.

An "adrenaline rush" occurs when the adrenal gland is stimulated through an activity that causes stress on the body, and certainly extreme sports, such as backcountry snowboarding and bungee jumping, fall into the category of causing stress.

According to the University of Maryland Endocrinology Health Guide, the stimulation of the adrenal gland releases a number of hormones, including epinephrine, or adrenaline. This increases the heart rate and the force of heart contractions, facilitates blood flow to the muscles and brain, causes relaxation of smooth muscles, and helps with the conversion of glycogen to glucose in the liver. For extreme athletes, this adrenaline rush is a feeling that can't come often enough.

But this is what we do, what we practice, what we live. But when we play the role of the audience, what does happen? What does change?

Most of us are content to play the role of screaming fan at the end of an Ironman, standing by while elite athletes like Hall cross the finish line. Why is it that we enjoy watching others endure the blood, sweat, and tears of extreme competition? "It's human nature to have curiosity about the outcome of such extreme sports and how people can defy death." In the same paper mentioned before, Berman, a PhD from Beverly Hills, spells as follows. He introduces the TV-reality notion, the meaning behind shared life.

Nowadays, what we see, moments and situations we empathize with, seek to become directly belonging to our emotion. Whatever broadcasted on a screen is not a simple message anymore, not even an experience, a story for the big audience; it is something we perceive as specifically targeted to us.

It is through this process that we develop the more intensive and involved reactions.

"It's different for everyone, but it's exciting to watch these people compete," says Anderson. "They are testing themselves to an extreme measure, and watching them push themselves while you say, 'I could never do that,' is fascinating."

Someone is doing something that you would not be able to do, you could not do, you would not do. But he is doing that in the best way, and you can judge him, you can watch him, you can incite him. We are the audience, we are interested viewers.

The balance and control of regular instincts application is tried in the highest-pressure scenario. The response, nerves and the grip over the uncontrollable is the main zeal why people who play or watch, enjoy it. These games are loved because the natural elements react to human extreme or instincts. The passion of man to overcome the natural barriers is an addiction for these athletes and spectators. Many time these overtures have dire impacts, but the fear of danger is the prime motivator to excel in its skill.

For all these reasons, the choice of this videoclip seemed to appropriate and suitable to our intents.

Figure 27 - An image from the videoclip in the exciting session

3.1.7 Music & songs

The choice of background music comes from an evaluation based on our experience and on an online research, looking for the proper song for each session, through playlists, databases, music critics websites.

Stressed session: for this part our choice is "Tubular Bells" by Mike Oldfield, soundtrack of the famous film "The exorcist", from 1973. Everyone knows this movie for the powerful images and for the soundtrack too, but this song has not been the first choice. Originally, Lalo Schifrin was the designated composer. William Friedkin (the director) entrusted this job to him, but the original song was heavily rejected. Why? Due to the way the public has reacted during the teaser exposition. People involvement was so strong that, as someone said, lot of viewers ran out of the room and others threw up on the chairs. The director himself threw the audio recorder out of the window, deciding to change composer, replacing him with Mike Oldfield. The new version is just slightly "softer" than the first one, and in association with one of the scariest horror movies ever produced, it easily evokes a stress condition, especially if used as background for images representing dead bodies or mutilations.

Elated session: our choice for this part is "All together now", a song written by Paul McCartney in the 1967, published with the legendary group "The Beatles" two years later. This song has elementary lyrics, with a melody that easily reminds songs that children use to sing playing together. McCartney described the song as a "children's sing-along with the title phrase inspired by the music hall tradition of asking the audience to join in". (Re. Lewisohn, Mark).

He also described a "subcurrent" in the song, a dual-meaning where "we are all together now". According to music critic Tom Maginnis of AllMusic (Maginnis, Tom 2009) McCartney created the song "to match the same light-hearted spirit" of Yellow Submarine".

This is the reason why matching this song to smiling and happy babies is a good choice.

Excited session: in this part there is a video from GoPro and a background song by Michal Calfan. This track is called "Resurrection", from 2011, chosen by GoPro as a perfect matching with extreme sports video, with a high adrenaline content.

Depressed session: the song for this session is "Hallelujah", the Jeff Buckley one, from 1992. As Alan Light (author of "The Holy or the Broken: Leonard Cohen, Jeff Buckley, and the Unlikely Ascent of Hallelujah) said, "there is no right way to sing it, it seems to belong to no one, yet remains one of the most ubiquitous songs in popular music". As well as emotions, everyone knows what they are, no one knows how to define them. This is a universal song, but with an ambiguous and unknown meaning. Of course, the song belongs to Leonard Cohen, but in the public eye, Jeff Buckley changed the song's fate. His own fate was far more tragic than the famous song he covered.

"Joyous and despondent, a celebration and a lament, a juxtaposition of dark Old Testament imagery with an irresistibly uplifting chorus, it is an open-ended meditation on love and faith – and certainly not a song that would easily be pegged as an international anthem.", so described by Salman Rushdie, music critic.

According to Bono, who has performed "Hallelujah" with U2, "it might be the most perfect song in the world".

Martin Bandier, chairman and CEO of Sony/ATV Music Publishing (which owns the rights to both Cohen and Buckley's versions of the song), phrased it a little differently: "'Hallelujah' is a brand." This song acquired an "official" deep meaning soon after September 11th, since VH1 (tv channel) put it in a on-air rotation music video, as a background for Twin Towers burning images.

3.2 Experimental setup: signal acquisition

Once the setup was ready, the signals were acquired by the sensors placed on the patients who were asked to sit back and remain as much steady as possible, while watching the presentation on a PC screen.

Three different proper sensors were used to acquire cardiac, respiratory and skin conductance signals. A real-time encoder was then used to convert physiological signals, connected to the computer. Figure 28 shows a schematic representation of the experimental setup.

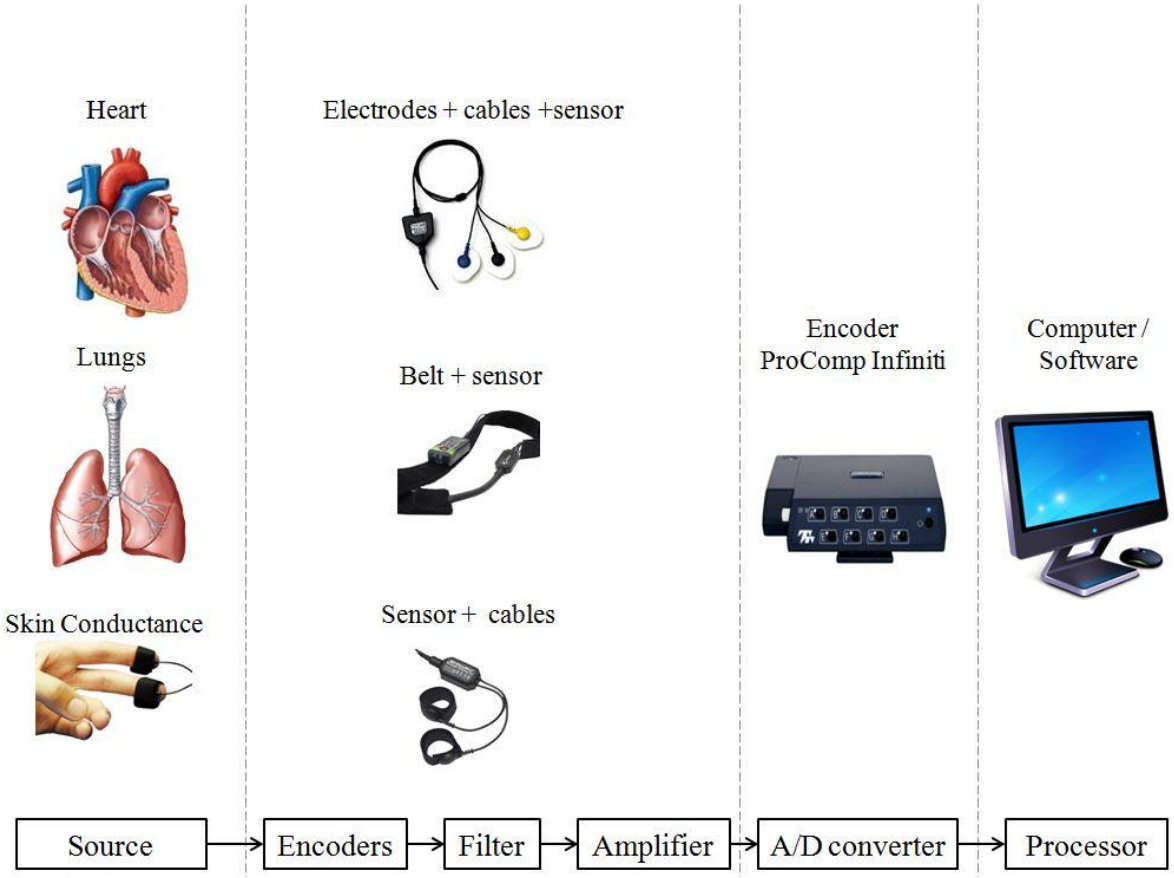


Figure 28 - Experimental setup of the present study

3.2.1 ProComp Infiniti: the device

Figure 29 - ProComp Infiniti

ProComp Infiniti (T7500M) is an encoder for real-time, computerized biofeedback and physiological data acquisition in any clinical setting (Fig. 29).

The device is housed in an ergonomically designed case, it requires only a USB port, and can be used with any IBM-compatible PC.

ProComp Infiniti can capture data in real time by connecting to the PC via fiber-optic cable: an USB adapter was used (TT-USB SA7700) connected to the fiber optical cable to optically isolate client from the computer.

The device has eight protected pin sensor inputs with two channels (A-B) sampling at 2048 samples per second and six channels (C-H) sampling at 256 samples per second.

The sampling rate is the number of measures (samples) per second taken from the continuous signal (analog signal). In our case, the analyzed analog signals are ECG, Respiration signal and skin conductance, and they are captured by non-invasive sensors. ECG was sampled at 2048 (channel A) while Respiration and skin conductance signals were sampled at 256 (channels G and E). The series of samples constitute the digital signal. A higher sampling rate (2048) for ECG signal allows higher precision in HRV calculations.

Furthermore, this device offers internal, user-activated calibration to ensure the highest quality signal, without the costly downtime associated with factory re-calibration.

ProComp Infiniti captures real-time data subsequently analyzed by the *BioGraph Infiniti Software* (*SA7900, Version 3.1.6, © 2003-2007*). The software has a Windows user interface, captures raw data and realizes realtime frequency processing and analysis.

3.2.2 Electrodermal Activity (EDA or SCR)

Figure 30 - EDA sensor (model: M5821) – GA 8056

The SCR assembly is made up of two pieces. The (SA9309P) cable is terminated with two snap receptacles (Fig. 30). The button sensor, is embedded into both Velcro bands which is secured around the fingers. A small voltage is applied to the skin between two sensors and the skin's current conduction is measured.

Skin Conductance is considered to be a function of sweat gland activity and skin pore size, both controlled by the sympathetic nervous system.

In response to stress or anxiety, sympathetic nervous system is aroused, thus activating the sweat glands. As sweat is produced, an increased conductance is measured.

The measurement unit is normally set as micro ohms.

The two sensors are attached to the first or third phalanx of index and middle fingers of the dominant hand.

3.2.3 Respiration Signal

Figure 31 - Respiration signal sensor (Model H4624)

To quantitatively measure the respiration signal, ProComp Infiniti uses the *Resp-Flex/Pro (SA9311M)* sensor/amplifier, which detects breathing by measuring the chest's expansion and contraction (Fig.31). The used unit of measure is relative (%).

This Respiration sensor is composed by a long strap and a sensor, consisting in a rubber tube sensitive to stretch. The belt is strapped around the abdomen of the patient with the sensor placed in the front (Fig. 32), and it converts the expansion and contraction of the abdominal area, to a sinusoidal signal, increasing when breathing in (which reflects in stretching and relaxing of the sensor) and decreasing when exhaling. From this raw signal it will then be possible to calculate the respiration rate and the relative breath amplitude.

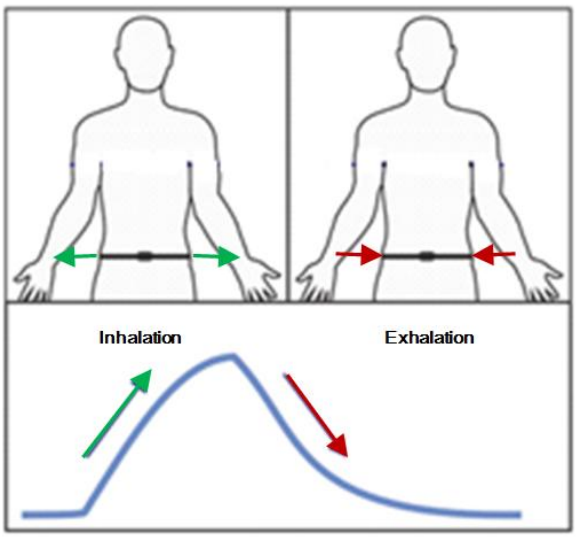


Figure 32 - Functioning of Respiration sensor

(As breathing in takes place the rib cage expands which stretches the device. When exhaling, the stretch relaxes and the sensor returns to its neutral position. The resulting waveform is displayed on the screen.)

3.2.4 Electrocardiogram (ECG/EKG)

To acquire the ECG signal, ProComp Infiniti uses the *EKG-Flex/Pro (SA9306M)* amplifier/Sensor (Fig. 33), which detects heart beats by measuring the physiological electrical activity of the heart. The unit of measure is millivolts (mV).

Figure 33 - ECG sensor (Model R1884)

This electrical signal generated by the heart when contracting, and propagating throughout the body, is captured by the electrodes, then amplified and filtered by the sensor and converted to a digital signal by the encoder. The sensor's output shows the usual beat pattern, then used to calculate the Heart Rate (HR) in beats per minute evaluating the time between consecutive beats.

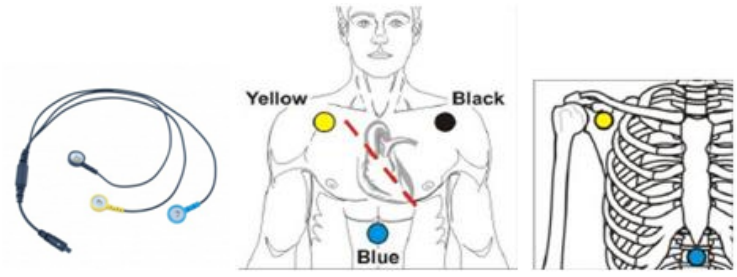


Figure 34 - ECG electrodes placement (triangular configuration

The chest placement for the pre-gelled electrodes was chosen and they were applied on the patient's chest, using a triangular configuration with two electrodes (the yellow and blue

ones) parallel with the heart's main axis. The yellow and black electrodes (positive and negative) should be placed over the right and left coracoid processes respectively (3/4 inch below right and left collarbones), and the blue electrode (reference) over the xiphoid process (just below the last left rib)(Fig.34). This configuration is chosen in order to minimize the risk of interference artifacts from chest and arm muscle activity.

3.2.5 Recording Session

First of all, the channel configuration of the ProComp Infiniti encoder is set:

- Channel A: ECG
- Channel E: SC
- Channel G: Respiration

Secondly, the Script session is created: in this step the duration of the recording is set to10 minutes to let the recording stop itself at the end of the presentation.
Then a new client is added: once name, surname and gender of the patient are inserted in the database, the recording session is ready to start.
The user interface of the Biograph Infiniti software is shown in the picture (Fig. 35), with the three dynamic charts representing the three-channel acquiring ECG, Skin Conductance and Respiration signals through time, respectively.

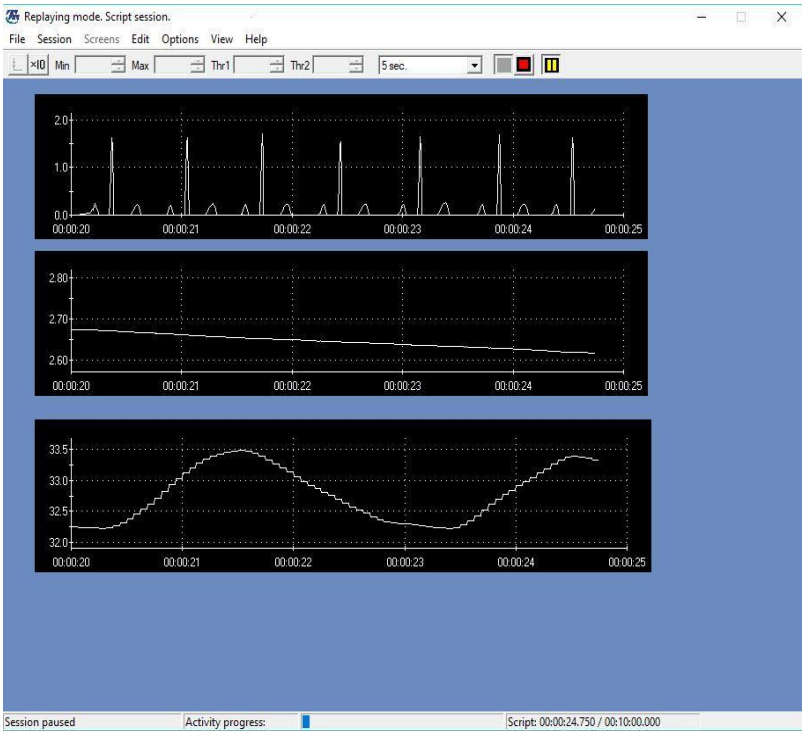


Figure 35 – Interface of the Biograph Infiniti Software

3.3 Signal analysis: ECG

All the signal captured by the sensors are sampled by the encoder ProComp Infiniti.

The sampling rate for the ECG signal is 2048 samples/sec, while for the GSR signal and the respiration signal a sample frequency of 256 is used. Digital signals are analyzed by the BioGraph Infiniti Software in order to extract meaningful features. Then the ECG signal is downsampled to 256 for computational reasons.

Signal analysis are then carried on taking into account the meaning of each different arousal phase related to each corresponding provided stimulus.

For each patient the four sessions related to the four evoked emotions are thus individually analyzed, in order to evaluate the variance and the meaning embedded in each one of them.

3.3.1 Features extraction

From the ECG processing, we obtained signals like the one in Figure 36.

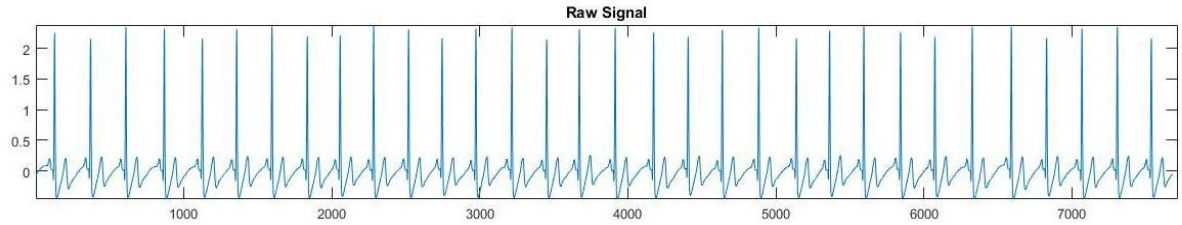


Figure 36 - Raw ECG signal

After developing a first filtering session, with a notch filter (to cut off the network interference at 50 Hz), the Pan-Tompkins algorithm is used to detect the QRS complexes, finding R peaks, allowing us to build the RR signal. This is the one we use to develop the time and frequency analysis.

3.3.1.1 Pan-Tompkins Algorithm

This algorithm has been developed in order to detect QRS complexes run time, very useful for Holter applications (24 hours patient monitoring) or to analyze arrhythmias. Its application allows to lower artefacts and noises influence on the signal, thanks to a very precise analysis of signals heights, slopes and amplitudes.

Noises can come from different sources, as movement artefacts, measuring instruments interferences, wires and electrodes movements. Another source of noise, very tough to identify, is that one related to the presence of T waves with high frequency components, similar to QRS complexes, that can lead to misclassifications.

The Pan-Tompkins algorithm is designed to filter signals using variable parameters and thresholds, adapting them to the signal in a run time way, to correctly detect interesting shapes[57].

Validated for 24 hours signals, with an accuracy of 99,3% of QRS complexes, it is now used to detect arrhythmias and heart failures, broadcasting in real time an alarm signal to a physician ready to evaluate the situation.

The structure is made of three parts: linear digital filtering, non-linear transformations and a decision rule algorithm.

First step is composed by a bandpass filter, a derivative one and a moving window integrator.

The bandpass filter is used to detect the muscular artifact, the 60Hz interference and the T waves one. An IIR ButterWorth filter (5-15 Hz) is used to delete the frequency not belonging to the QRS complex; the derivative filter detects the ECG slope and, in particular, the QRS one. This filter has a quite linear response with 30 Hz as the cut off frequency. The moving window filter is used to delete the fake peaks.

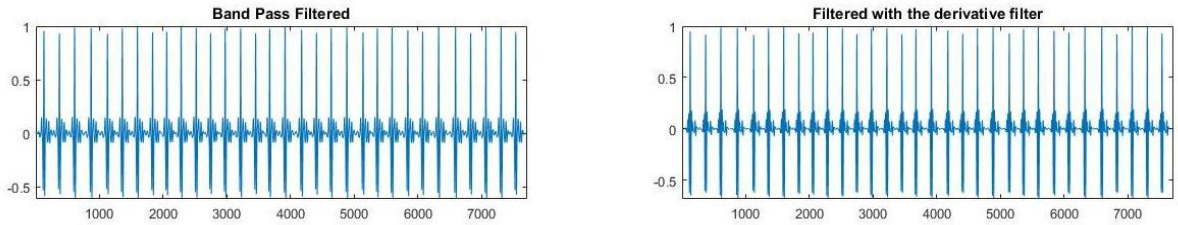


Figure 37 - Signal filtering

The second step has a signal amplitude squaring as a non-linear transformation. The signal is squared point by point to have only positive values and to highlight high frequencies. Decision rule is based on adapting and variable thresholds, set in an automatic way. Outputs are shown in Figure 37, 38 and 39.

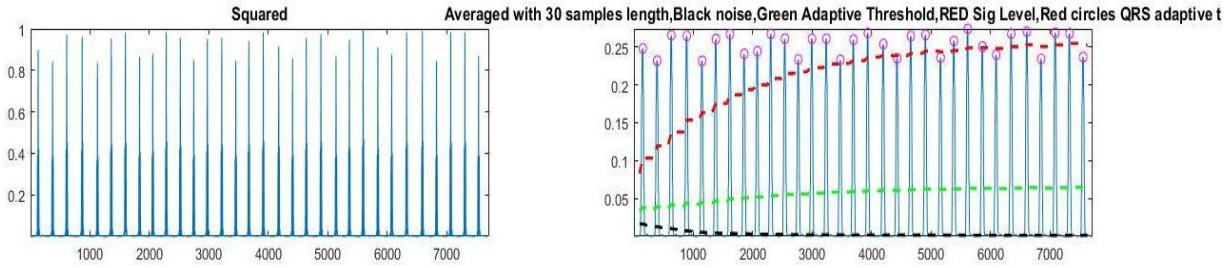


Figure 38 - Squared signal

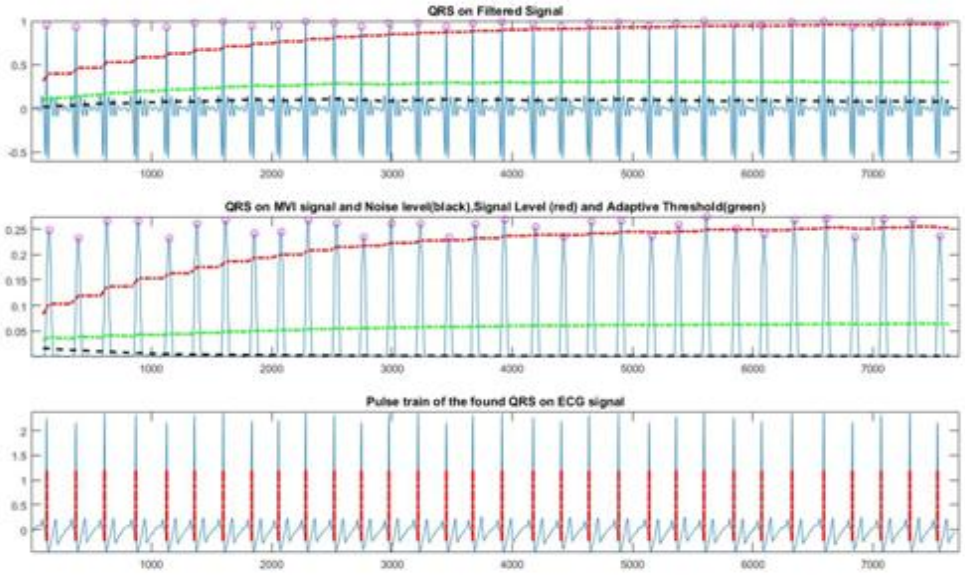


Figure 39 - Pan-Tompkins algorithm output : peak detection

After Pan-Tompkins application, the result is the one shown in Figure 40, with R peaks detection. Distances between R peaks are used to build the tachogram signal, plotting these values in relation to heart beat indexes.

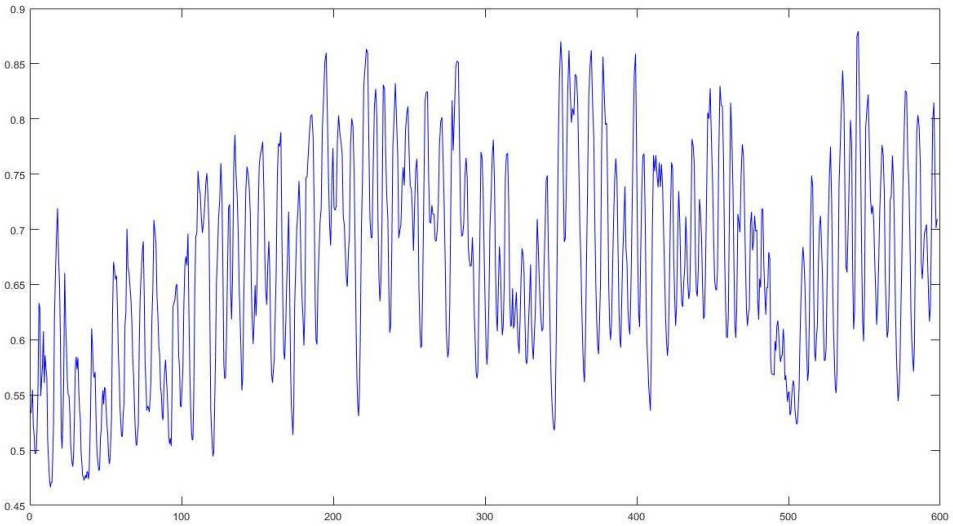


Figure 40 - Tachogram (R-R intervals)

3.3.2 Heart Rate Variability (HRV)

HRV is an important quantitative marker of cardiovascular regulation performed by the ANS. It is considered a reliable predictor of mortality after the occurrence of an acute myocardial infarction and it is also indicative of ventricular dysfunction in patients with congestive heart failure. It is clinically useful to evaluate HRV to find signs of pathological processes, functional disorders, but also to evaluate physical condition of a subject.

HRV is an important quantitative marker of cardiovascular regulation performed by the ANS.

HRV is considered a reliable predictor of mortality after the occurrence of an acute myocardial infarction and it is also indicative of ventricular dysfunction in patients with congestive heart failure.

HR is estimated by the computation of the number of R-wave events (heartbeats) per unit time on the electrocardiogram (ECG).

The model we are going to explain, unlike common approaches, is able to model the R-R intervals (or heart rate series) in order to reflect the point process structure of the R-wave events. The R-wave events are a sequence of discrete occurrence in continuous time, thus forming a point-process structure. The R-wave represents the electrical impulse from the heart's conduction system provided by ventricular contractions.

Heartbeat intervals are here modeled as an history-dependent inverse Gaussian (HDIG) point process which allows to describe the probability density and the subsequent definition of HR and HRV.

By estimating the time-varying parameters of the inverse Gaussian model, maximizing the local likelihood.

3.3.3 Physiology of R-wave events

R-R intervals are the time lapses between two successive R-wave events. Each R-wave event is triggered by the synchronous depolarization of the sinoatrial (SA) node cells, which then propagates through the proper conduction system towards the left atrium and the ventricles. After depolarization, transmembrane potentials of those cells return to their resting potentials and then begin anew their rise toward threshold, rise that can be modeled

as a Gaussian random walk. The elapsing time between two consecutives thresholds crossing are the R-R intervals, and they can be modeled as independent inverse Gaussian random variables[58].

Even if SA inputs from the ANS occur in ms, their effects on the output can persist for many seconds. For example, an increase in sympathetic activity, received in input from the SA node, produces, as output, shorter R-R intervals for a while. R-R intervals lengths cannot thus be considered independent but, rather, are a function of the recent history of these inputs from the ANS to the SA node.

The point process model takes into account the dynamic nature of the inputs, related to the fact that they are part of the continuously changing cardiovascular control circuitry. Given the R-wave events detected from the ECG signal, each RR interval is extracted by this point-process model of the ventricular contraction events.

3.3.4 Heart rate probability model (Point Process)

Given any R-wave event u_k, the length of the next R-R interval obeys an HDIG probability density function [58].

$$f\left(t\middle|H_{u_k},\theta\right) = \left[\frac{\theta_{p+1}}{2\pi(t-u_k)^3}\right]^{\frac{1}{2}} exp\left\{-\frac{1}{2}\frac{\theta_{p+1}\left[t-u_k-\mu\left(H_{u_k},\theta\right)\right]^2}{\mu\left(H_{u_k},\theta\right)^2(t-u_k)}\right\}$$

Equation 1 is a probability model for the R-R intervals representing the effects of the last sympathetic and parasympathetic input to the SA node (history dependence) through its mean parameter $\mu(Hu_k,\theta)$.

Where t is any time satisfying $t> u_k$ and Hu_k is the history of the R-R intervals up to u_k, θ is the vector of model parameters.

This R-R probability model represents the dependence of the R-R interval length on the recent history of parasympathetic and sympathetic inputs to the SA node.

Assuming R-R intervals as independent, the equation (1) simplifies to a renewal inverse Gaussian model.

The mean and standard deviation of this R-R probability model are the following:

$$\mu_{RR} = \mu\left(H_{u_k}, \theta\right)$$

$$\sigma_{RR} = \left[\mu\left(H_{u_k}, \theta\right)^3 \theta_{p+1}^{-1}\right]^{\frac{1}{2}}$$

Then, once characterized the stochastic properties of the R-R intervals, a precise definition of HR and HRV can be obtained.

The heart rate probability density function can be defined from the RR-interval probability function, through a standard change of variables: this function defines the stochastic properties of heart rate in terms of probability density.

$$\mathrm{f}\left(\mathrm{r}|\mathrm{H}_{\mathrm{u_k}}, \theta\right) = \left|\frac{\mathrm{dt}}{\mathrm{dr}}\right| \mathrm{f}\left(\mathrm{t}|\mathrm{H}_{\mathrm{u_k}}, \theta\right)$$

$$= \left(\frac{\theta_{\mathrm{p+1}}^{*}}{2\pi \mathrm{r}}\right)^{\frac{1}{2}} \exp\left\{-\frac{10_{\mathrm{p+1}}^{*}[1 - \mu^{*}\left(\mathrm{H}_{\mathrm{u_k}}, \theta\right)\mathrm{r}]^2}{\mu * (\mathrm{H}_{\mathrm{u_k}}, \theta)^2 \mathrm{r}}\right\}$$

Where r is the heart rate random variable, containing the conversion of R-R interval measured in ms to heart rate measurements in beats per minute.

The mean and standard deviation of the heart rate probability density (previous equation) are defined as follows:

$$\mu_{\mathrm{HR}} = \mu * \left(\mathrm{H}_{\mathrm{u_k}}, \theta\right)^{-1} + \theta_{\mathrm{p+1}}^{*(-1)}$$

$$\sigma_{\mathrm{HR}} = \left[\frac{2\mu * \left(\mathrm{H}_{\mathrm{u_k}}, \theta\right) + \theta_{\mathrm{p+1}}^{*}}{\mu * \left(\mathrm{H}_{\mathrm{u_k}}, \theta\right) {\theta_{\mathrm{p+1}}^{*}}^2}\right]^{\frac{1}{2}}$$

Heart rate standard deviation and RR standard deviation are then used to construct the indexes of heart rate variability in the Local maximum-likelihood estimation of instantaneous HR and HRV. Instantaneous HR and HRV can be tracked by following the time evolving variant theta and the continuous evaluation of the evolution of the last two equations, heart rate mean and standard deviation.

3.3.4.1 Local maximum-likelihood estimation of instantaneous HR and HRV

A local maximum likelihood procedure is then used to estimate the time-varying parameter theta. The ECG recording interval is (0,T], the length of the local likelihood observation interval is l for t Є [l,T]. Defining the local likelihood observation interval as t^l = (t-l,t], and assuming that within t^l we observe n_t R-wave event times($u_{t-l:t}=(u_1,\ldots,u_{nt})$), then at time t, the local maximum likelihood estimate of theta (θ_t estimate) is the estimation of theta on t^l. The local log likelihood is the following:

$$\log f(u_{t-1:t}|\theta_t) = \sum_{i=2}^{n_t} w(t-u_i)\log f\left(u_i - u_{i-1}\middle|H_{u_{i-1}},\theta_t\right) + w\left(t-u_{n_t}\right)\log\int_{t-u_{n_t}}^{\infty} f\left(v\middle|H_{u_{n_t}},\theta_t\right)dv$$

Where w(t) is the weighting factor θ.

Once θ_t estimate is computed in the t^l interval, the t^l interval is shifted to (t-l+Δ, t+ Δ] and the local maximum likelihood estimation is repeated. The procedure stops when t=T.

Given the local maximum-likelihood estimation of theta at time t (θ_t estimate), the instantaneous estimates of mean R-R, R-R standard deviation, HR mean and HR standard deviation can be computed as follows:

$$\hat{\mu}_{RR_t} = \mu(H_t, \hat{\sigma}_t)$$

$$\hat{\sigma}_{RR_t} = \left[\mu(H_t, \hat{\sigma})^3 \hat{\theta}^{-1}{}_{p+1,t}\right]^{\frac{1}{2}}$$

$$\hat{\mu}_{HR_t} = \mu_t * \left(H_t, \hat{\theta}\right)^{-1} + \hat{\theta}^{*-1}{}_{p+1,t}$$

$$\hat{\sigma}_{HR_t} = \left[\frac{2\mu_t^*(H_t, \hat{\theta}^*{}_{p+1,t})}{\mu_t^*(H_t, \hat{\theta}_t)\hat{\theta}^{*2}{}_{p+1,t}}\right]^{\frac{1}{2}}$$

The key feature of this analysis is the estimation of these indices in continuous time.
The local likelihood function is defined at t, the right end point of the observation interval and the point where the local maximum likelihood estimate is computed. t Є (l,T]
Thus, because theta is estimated in continuous time, all the indexes can be computed in continuous time as well, being continuous functions of θ_t.
Equations of RR standard deviation and HR standard deviation are designed as indexes of HRV.

3.3.5 Time domain analysis

The easiest way to get information from Heart Rate Variability is a time domain analysis, allowing us to evaluate point by point and interval by interval variations in time. The starting point is QRS complexes detection on a continuous ECG, finding and selecting the so-called normal-to-normal intervals, i.e. an interval between two adjacent QRS complexes that are not ectopic beats [47] [34] [59].

3.3.5.1 Statistical measures

The standard HRV measurements are related to mean and variance of these intervals. Parameters are presented below with formulas and descriptions (Table 4).
Entropy index (E) is evaluated, using Shannon definition of entropy.

$$\mathrm{RMSSD} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n-1}(\mathrm{RR_{i+1}} - \mathrm{RR_i})^2}$$

$$\mathrm{SDNN} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n-1}(\mathrm{RR_i} - \overline{\mathrm{RR}})^2}$$

$$\mathrm{pNN50} = \mathrm{P}(|\mathrm{RR_{i+1}} - \mathrm{RR_i}| > 50ms)$$

SDNN represents the square root of the variance of the NN intervals: since the variance is a way to define the total power of the spectral analysis, in this index variability is described in its cyclic components. It is strictly related to signal length (in a proportional way), and for this reason it has to be compared only if signals have the same duration. This measure is also useful for the Holter recordings (24 hours), since it shows the heart's intrinsic ability of reaction to hormonal influences. RMSSD is used to give an estimation of the high-frequency variations in heart rate, but in recordings characterized from short duration, because it gives info about the parasympathetic regulation of the heart. Both measures are expressed in milliseconds [47].

Variable	Units	Description
		Statistical measures
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.
		Geometric measures
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 7·8125 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)[21].
Logarithmic index		Coefficient φ of the negative exponential curve $k \cdot e^{-\varphi t}$ which is the best approximation of the histogram of absolute differences between adjacent NN intervals[22].

Table 4- Temporal indexes for ECG [47]

3.3.5.2 Geometric measures

From the NN series geometric patterns can be also described.

For instance, sample density distribution of NN interval durations, and sample density distribution of differences between adjacent NN intervals (described below as RR sequence) or Lorentz plot, etc. Moreover, HRV triangular index is used, i.e. density distribution integral (number of all NN intervals) divided by the most frequent value (mode) with its absolute frequency k. HRV can be measured as the ratio of the standard deviation SD2 along the identity line (RRi+1=RRi) and the standard deviation SD1 along the perpendicular axis (RRi+1= - RRi). SD1 is based on successive differences, therefore named “short term HRV”, whereas the SD2 is based on the summation

of successive RR intervals like a moving average. Its deviation represents” long term HRV”.

$$rr_i = \frac{2(RR_i - RR_{i-1})}{RR_i + RR_{i-1}}$$

$$\text{HRV triangular index} = \text{n/k}$$

$$\frac{SD1}{SD2} = \frac{\sqrt{\frac{1}{2}}\sigma(RR_{i+1} - RR_i)}{\sqrt{\frac{1}{2}}\sigma(RR_{i+1} + RR_i)}$$

3.3.6 Frequency domain analysis

PSD (power spectral density) provides a general outline of the power distribution in the different frequencies. To evaluate that there are parametric and non-parametric methods. Power spectral density has an informative meaning if evaluated on 4-5 minutes periods, but in this work sessions last about 90 s. So, as described in a detailed way in Point Process section, it is useful to develop a punctual analysis of the power spectrum of a signal.

Parameters evaluated are the ones that link power and frequency, as VLF (very low frequency), LF (low frequency) and HF (high frequency) power components. These values are normally described as absolute power (ms^2), but LF and HF components can be also measured in normalized units, representing the relative value of each power component proportionally to the total power minus VLF component. This is done to underline the behavior of the SNA [47].

Another important parameter is the sympatho-vagal balance, whose behavior is described as LF on HF, useful to define sympathetic system activation.

To process this analysis Point Process is used.

Spectral features are summed up in Table 5.

Variable	Units	Description	Frequency range
		Analysis of short-term recordings (5 min)	
5 min total power	ms^2	The variance of NN intervals over the temporal segment	approximately ≤0·4 Hz
VLF	ms^2	Power in very low frequency range	≤0·04 Hz
LF	ms^2	Power in low frequency range	0·04–0·15 Hz
LF norm	n.u.	LF power in normalised units LF/(Total Power–VLF) × 100	
HF	ms^2	Power in high frequency range	0·15–0·4 Hz
HF norm	n.u.	HF power in normalised units HF/(Total Power–VLF) × 100	
LF/HF		Ratio LF [ms^2]/HF [ms^2]	
		Analysis of entire 24 h	
Total power	ms^2	Variance of all NN intervals	approximately ≤0·4 Hz
ULF	ms^2	Power in the ultra low frequency range	≤0·003 Hz
VLF	ms^2	Power in the very low frequency range	0·003–0·04 Hz
LF	ms^2	Power in the low frequency range	0·04–0·15 Hz
HF	ms^2	Power in the high frequency range	0·15–0·4 Hz
α		Slope of the linear interpolation of the spectrum in a log-log scale	approximately ≤0·04 Hz

Table 5 - Spectral indexes from ECG [47]

Total power is an approximate way of estimation of the overall autonomic activity. Power in LF, as indicated in the table above, is a marker of activity of the sympathetic and para-sympathetic activity, but gives stronger info about the first one; HF power, instead, reflects vagal activity, and it is frequently called “respiratory” band, related to the NN variations caused by the respiration signal. The ratio between these two power indexes is a reflection of the balance between sympathetic and parasympathetic systems. Normalized powers are useful to underline the changes in sympathetic and parasympathetic regulation, avoiding the influence of VLF effects.

3.4 Signal analysis: respiratory signal

The respiration sensor detects abdominal expansion/contraction and outputs a respiration waveform with a trend like the one in Figure 41:

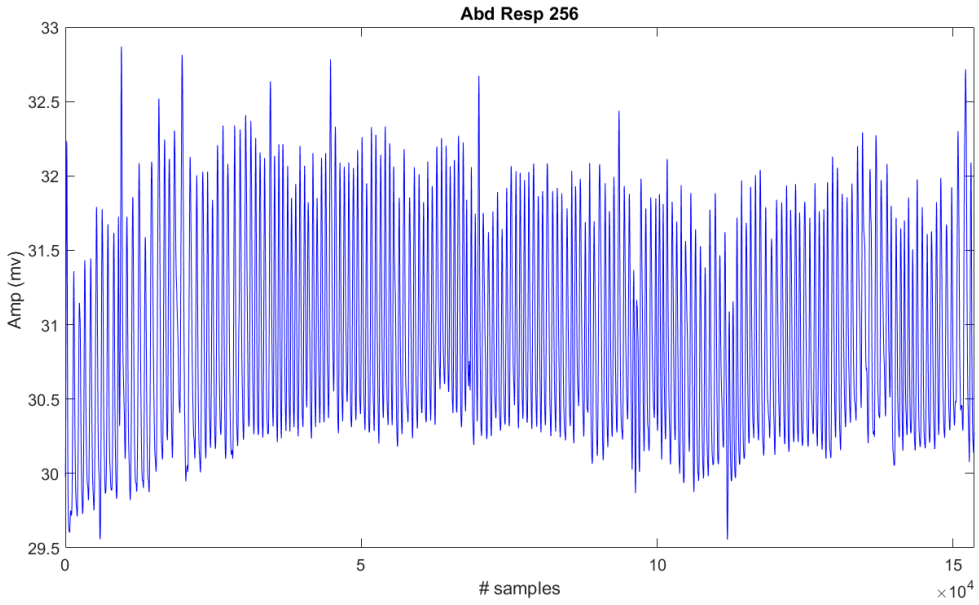


Figure 41- Abdominal Respiratory signal

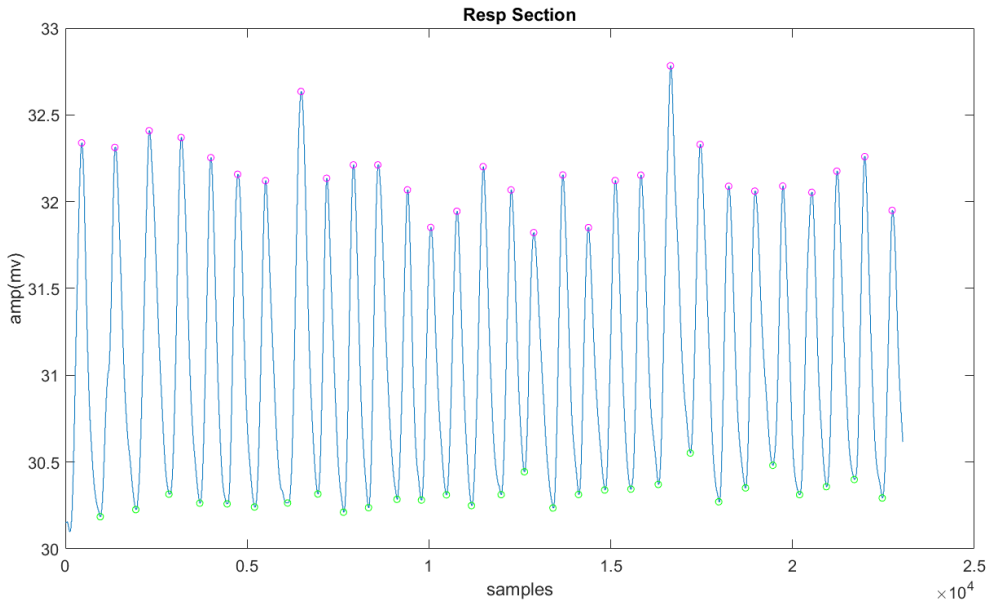


Figure 42 - Abdominal respiratory signal: peak detection

From this raw signal, detection of the maxima is performed (Fig. 42) and the respiration rate is obtained by evaluating the peak-peak distances. Then, the relative breath amplitude, its mean and its standard deviation, are extracted from the waveform.
In the frequency domain, moreover, some indices are extracted just as already described in the previous paragraph, talking about the ECG analysis.
Hence, by analyzing the spectral power of the respiration signal, the VLF, LF and HF components of the absolute power of the signal were extracted. For this scope Yule Walker method is used, through an autoregressive parametric model estimation.

3.5 Signal analysis: Electrodermal Activity

First of all, it is right to underline the difference between EDA (electrodermal activity) and SCR (Skin Conductance Response).
EDA is the signal acquired in the way we have already described.
SCR is the characteristic response, the one that highlights short waveforms and peaks.
In order to obtain SCR from the EDA acquired on each subject, there are some steps that have to be followed (Fig. 43)[33].
This algorithm starts with a downsampling from 256 Hz to 20 Hz, followed by the differentiation of the derived signal.
Then, a convolution with a 20-point Bartlett window is performed.
The Galvanic Response to internal and external stimuli is represented by peaks: to correctly identify these occurrences, two consecutive zero-crossing points (from negative to positive and vice versa) are detected, evaluating the amplitude of the response looking for the maximum in these zones.
Detected peaks with an amplitude smaller than 10% of the maximum SCR amplitude (for each area) are excluded.
Evaluated features are: mean values of amplitude, prominence and width of selected peaks, number of peaks and maximum amplitude for each session of the presentation, so there are 5 parameters for each zone.

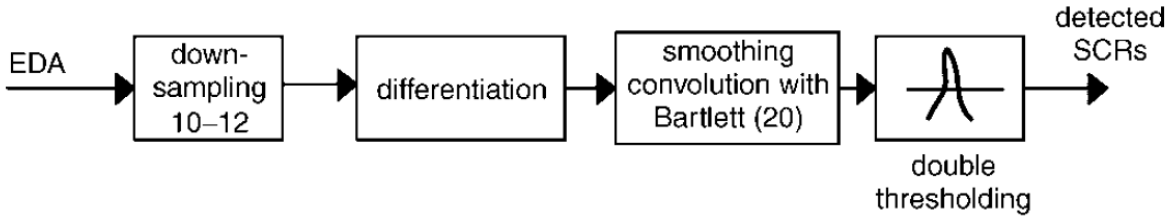


Figure 43 - EDA processing

3.6 Bivariate signal analysis

3.6.1 Coherence

Spectral coherence is used to evaluate the relation between two signals or dataset, in order to provide a power transfer estimation between an input signal and the output one, in a linear system.

The magnitude-squared coherence is represented as the following one:

$$G_{xy(f)} = \frac{\left|G_{xy}(f)\right|^2}{G_{xx}(f)G_{yy}(f)}$$

where Gxy(f) represents the cross-spectral density between x and y (signals). Cross-spectrum is normalized on the product of the two spectra Gxx(f) and Gyy(f), obtaining a coherence value between 0 and 1.

In Matlab coherence is evaluated point by point (frequency by frequency) after the evaluation of the two spectra. Significant values are considered the ones above the threshold, normally chosen as 0.5 (as you can see in the following figure, representing the coherence between the ECG signal and the respiratory one).

3.7 Statistical Analysis

3.7.1 Normality Test

Anderson-Darling test has been used to verify if data comes from a population with a normal distribution. It is slightly different from Kolmogorov-Smirnov one, but it takes more into account distribution tails [60].

AD test makes use of a specific distribution to evaluate critical values and this allows a better test sensitivity, but these values need to be calculated for each analyzed distribution.

Since this test is applied through a Matlab command (adtest), there is no need to calculate critical values, they are included in the algorithm. This command returns a decision for the null hypothesis of a normal distribution, while the alternative one is that the population comes from a non-normal distribution. Result is 1 if the test rejects the null hypothesis with a significance level of 5%, otherwise it is 0.

3.7.2 Wilcoxon test

It represents the non-parametric equivalent of the t-test Student for independent samples. It has to be used when data are distributed in a non-Gaussian way. This solution is sufficiently accurate for $n > 16$, where n is the size of the evaluated populations.

In Matlab ranksum command is used, that returns the p-value of a two-sided analysis. It tests the null hypothesis that data in x and y are samples from continuous distributions with equal medians, against the alternative hypothesis that they are not. The test assumes samples as independent, and they can also have different lengths.

3.7.3 Boxplot

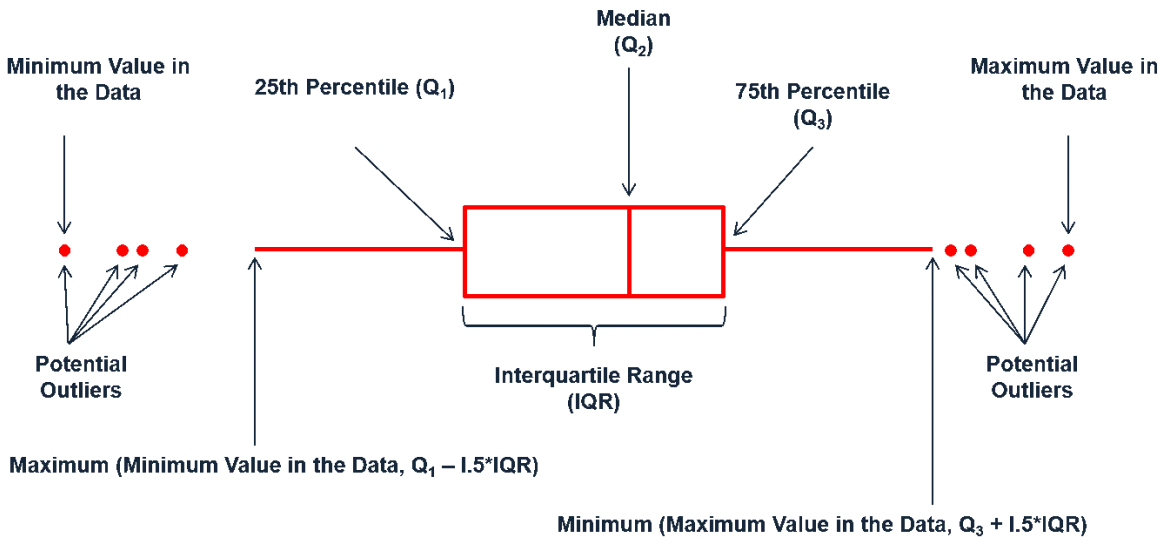


Figure 44 - Boxplot

Box plot is a graph (Fig. 44) built up starting from 5 numbers: minimum, 1° quartile (Q1), median, 3° quartile (Q3), maximum, and it describes main characteristics of the distribution.

A percentile is the measure that indicates the value below which a certain percentage of observations in a group fall. Quartiles are values that divides the data table in four groups with an equal number of values, so 25%, 50% (median), 75% and 100%.

It is obtained plotting on an axis these 5 numbers: the box has Q1 and Q3 as inferior and superior extremes respectively. Whiskers are made linking Q1 to the minimum and Q3 to the maximum value. In some graphs whisker is 1.5 times longer than the box height (Q3 - Q1, interquartile range). Comparing whiskers lengths (distances Q1-minimum and Q3-maximum) and the heights of the two rectangles that makes the box (distances Q1-median and median-Q3) info about distribution symmetries could be get: the distribution is as symmetric as rects heights are similar to each other. Moreover, whiskers highlight the presence of outliers.

Box plot is good to represent distributions in a summary way: with few info the morphology of the distribution could be get (symmetric or not) [61].

3.7.4 Error Bar

Error bar is a graphic representation (Fig. 45) of data variability, used to evaluate error or uncertainty in a measure. They give a general idea about measure precision (how much real value is far from the evaluated one). Error bar represent the standard deviation of the uncertainty, the standard error or a confidence interval. There quantities are not the same, they need to be chosen in a proper way.

Error bar are useful to visually compare two quantities, evaluating statistical differences[62].

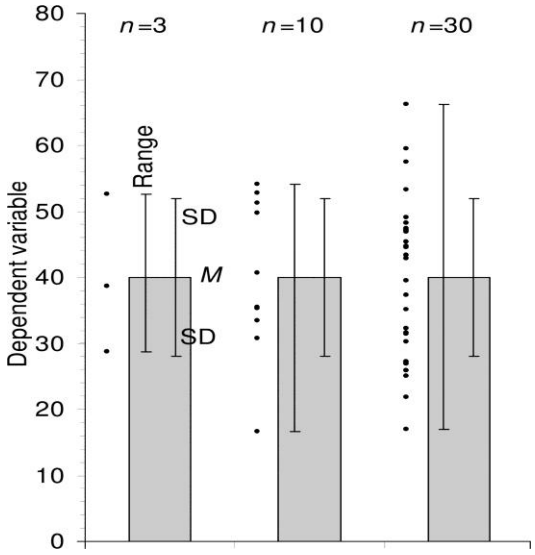


Figure 45 - Errorbar

4. Results

4.1 PHQ-9 Questionnaire

The first step to follow is the evaluation of the results obtained from the Phq-9 questionnaire. For each one the total score is calculated, to "delete" the acquired signal of that subjects with a score superior than 4. Overcoming that threshold is an index of mental disorder, or maybe an unbalanced emotional moment, that could affect and modify the emotional responses, that are spontaneous but affected from the general approach. After these evaluation, 8 subjects are deleted, within 5 were women and 3 men.

The table showed below sums up the situation of testers actually analyzed.

Figure 46 shows, for each of the 10 questions, the answers distribution.

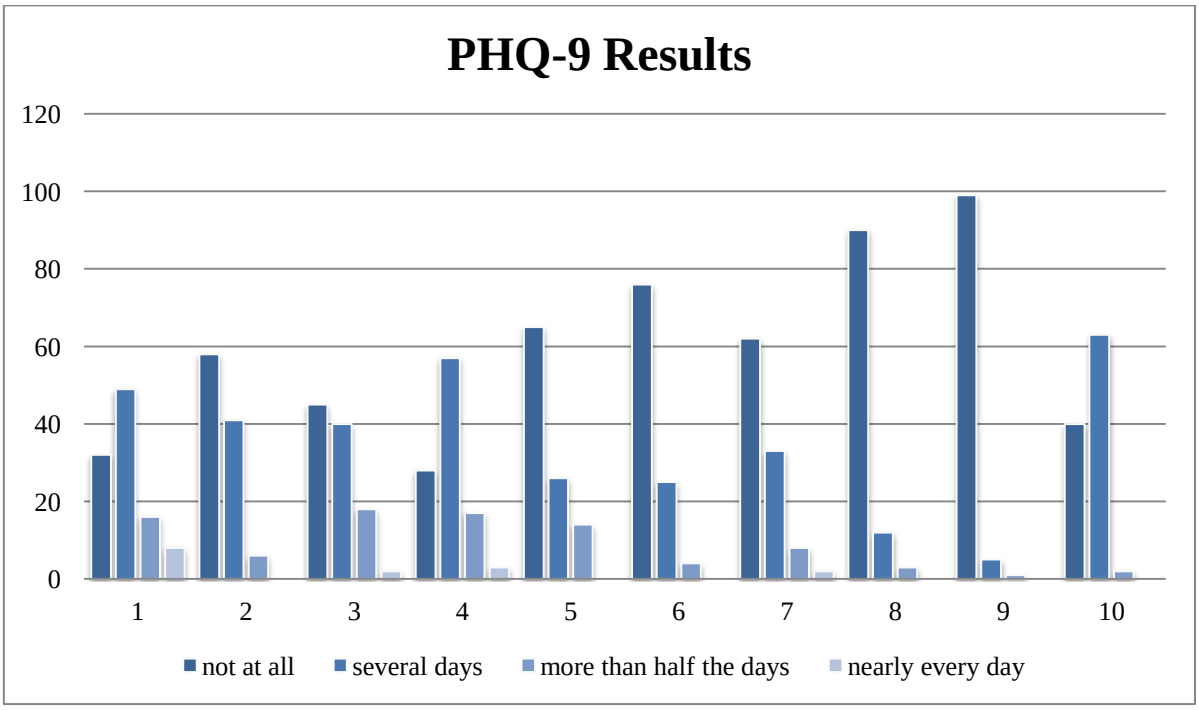


Figure 46 - PHQ-9: results of the questionnaire

The ten questions, corresponding to the horizontal axis labels in the chart, were the following:

1) Little interest or pleasure in doing things

2) Feeling down, depressed or hopeless
3) Trouble falling asleep, staying asleep, or sleeping too much
4) Feeling tired or having little energy
5) Poor appétit or overeating
6) Feeling bad about yourself – or that you're a failure or have let yourself or your family down
7) Trouble concentrating on things, such as reading the newspaper or watching television
8) Moving or speaking so slowly that other people could have noticed. Or, the opposite . being so fidgety or restless that you have been moving around a lot more than usual
9) Thoughts that you would be better off dead or of hurting yourself in some way

All questions were asked with reference to facts experienced in the previous two weeks.
For the PHQ-9 test, the "depression severity" level, linked to different scores, is defined as follows: Minimal, 0-4; Mild, 5-9; Moderate, 10-14; Moderately severe, 15-19; Severe, 20-27 [53].
In the present study, among the 105 subjects that submitted the questionnaire, 8 subjects obtained a score higher than 4 and were therefore eliminated from our analysis in order to maintain the investigated population as homogeneous as possible.
Otherwise, stimulation, could have brought to unexpected emotive reactions for those subjects, thus adversely affecting final results.
Some of the analyzed subjects are shown in Fig. 47.

Figure 47 - Some partecipants

4.2 Monovariate Signal Analysis

The starting point is a visual signal evaluation, trying to identify measurement errors, due to movement artifacts, detaching of electrodes or other problems.

Because of some experimental errors, data for some participants and measures were unavailable.

Figure 48 represents the morphology of a crop of the properly filtered signals: ECG, respiratory signal, SCR. Figure 49 also reports the sessions division.

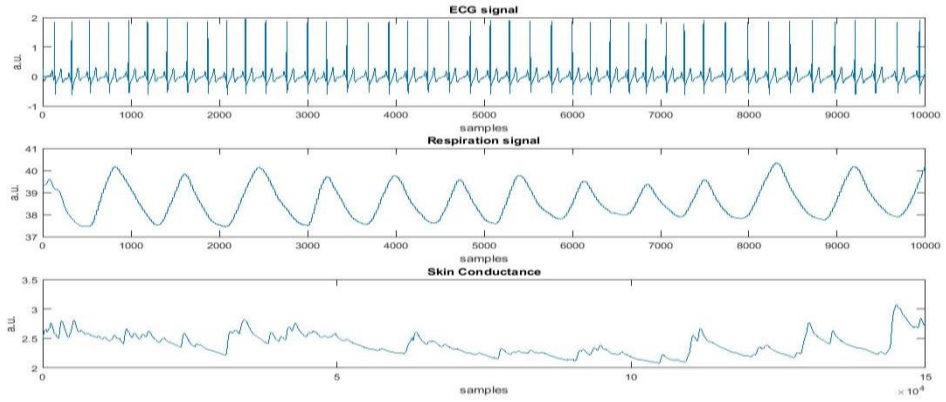


Figure 48 – Zoom of sampled signals: ECG, Respiration signal, Skin Conductance

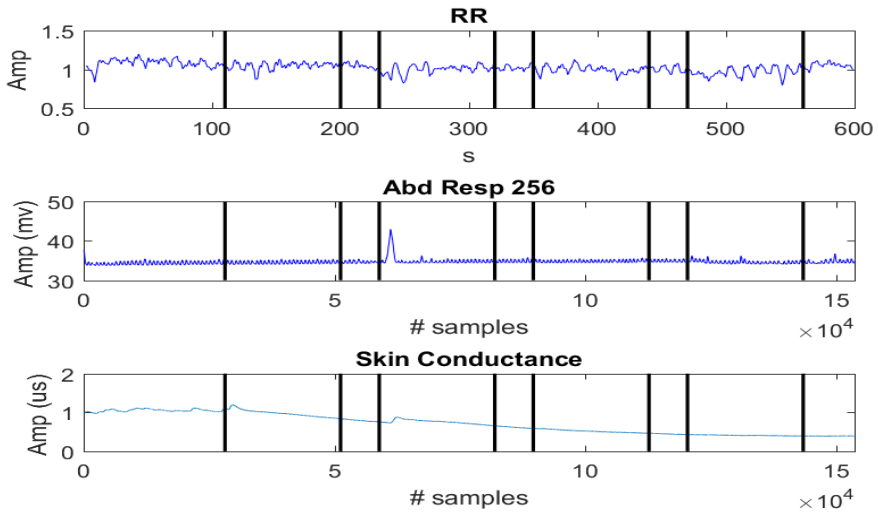


Figure 49 - Plot of sampled R-R intervals, Abdominal Respiration and Skin Conductance of a 10 minutes cycle, with sessions division (in black)

4.2.1 ECG Analysis : Time domain

As already specified, ECG has been firstly analyzed in the time domain. Peaks are detected through the Pan-Tompkins algorithm, with an output as the one shown in Figure 50.

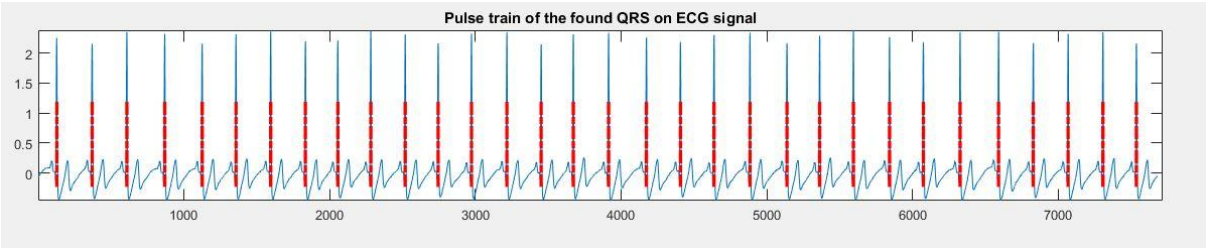


Figure 50 - Pulse train of the found QRS complexes on ECG signal

Detected peaks have been used to build the tachogram, representing RR sequence in function of beats indexes. On the RR signal features in the dime domain have been extracted, with values as the ones presented later.

The evaluation has been done for each of the emotional areas (baseline, depressed, stressed, elated, excited).

In Figure 51 the red part of the tachogram represents the emotional response in the order previously defined), while the black part is the neutral zone. The first neutral zone is the one considered as baseline.

In Figure 52 return map of RR intervals are showed, relating the value of an RR interval at the moment (i) and the following one (i+1). This is a sort of cloud that displays HRV value versus the previous one, and if a variation exists in HR, the cloud is wider, while if the HR is quite regular, the shape is similar to a small ball with a greater density.

Starting from this evaluation, all the temporal features are extracted: SDNN, pNN50, RMSSD, SD1SD2ratio, HRV triangular index, through an algorithm implemented by ourselves. Values are presented in Table 6.

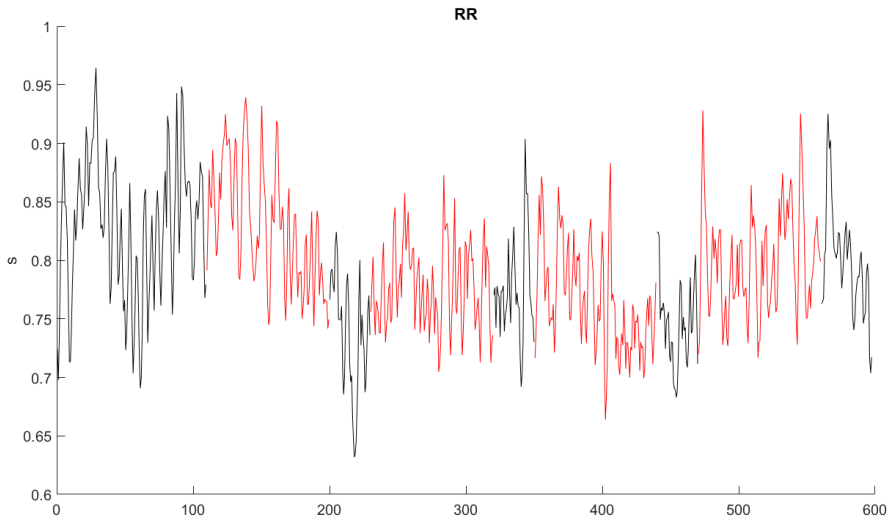


Figure 51 - R-R intervals plot in time, with sessions division: baseline, depressed, neutral, elated, neutral, excited, neutral, stressed, neutral

		Emotion				
		Baseline	**Depressed**	**Elated**	**Excited**	**Stressed**
Temporal Index	**meanRR**	0. 8243	0. 8269	0. 7769	0. 7639	0. 7972
	SDNN	0. 0596	0. 0515	0. 0355	0. 0443	0. 0440
	RMSSD	0. 0389	0. 0322	0. 0315	0. 0315	0. 0313
	pNN50	29	15	10	11	14
	TI	133	54. 500	116	117	56. 500
	SD1	0. 0145	0. 0131	0. 0128	0. 0139	0. 0128
	SD2	0. 0794	0. 0689	0. 0449	0. 0586	0. 0128
	SD12	0. 1832	0. 1898	0. 2849	0. 2367	0. 2204
	E	33. 701	27. 583	34. 986	36. 210	32. 000

Table 6 – Time domain parameters in each session of a representative subject

As shown in the previous table, there are some variations for time domain indexes in the ECG analysis, following state of art works. For instance, mean RR values decrease passing from baseline to "emotional" phases as stressed and excited one, while remains quite similar in depressed section. The same for RMSSD, where frequency increasing led to RMSSD decrease. These considerations agree with the ones explained in a previous study [34]. SDNN is considered in normal values if greater than 35 in the neutral state, showing a good regulating function of ANS. Clinically, a decrease in this index is a sign of "internal changes", due to external causes.

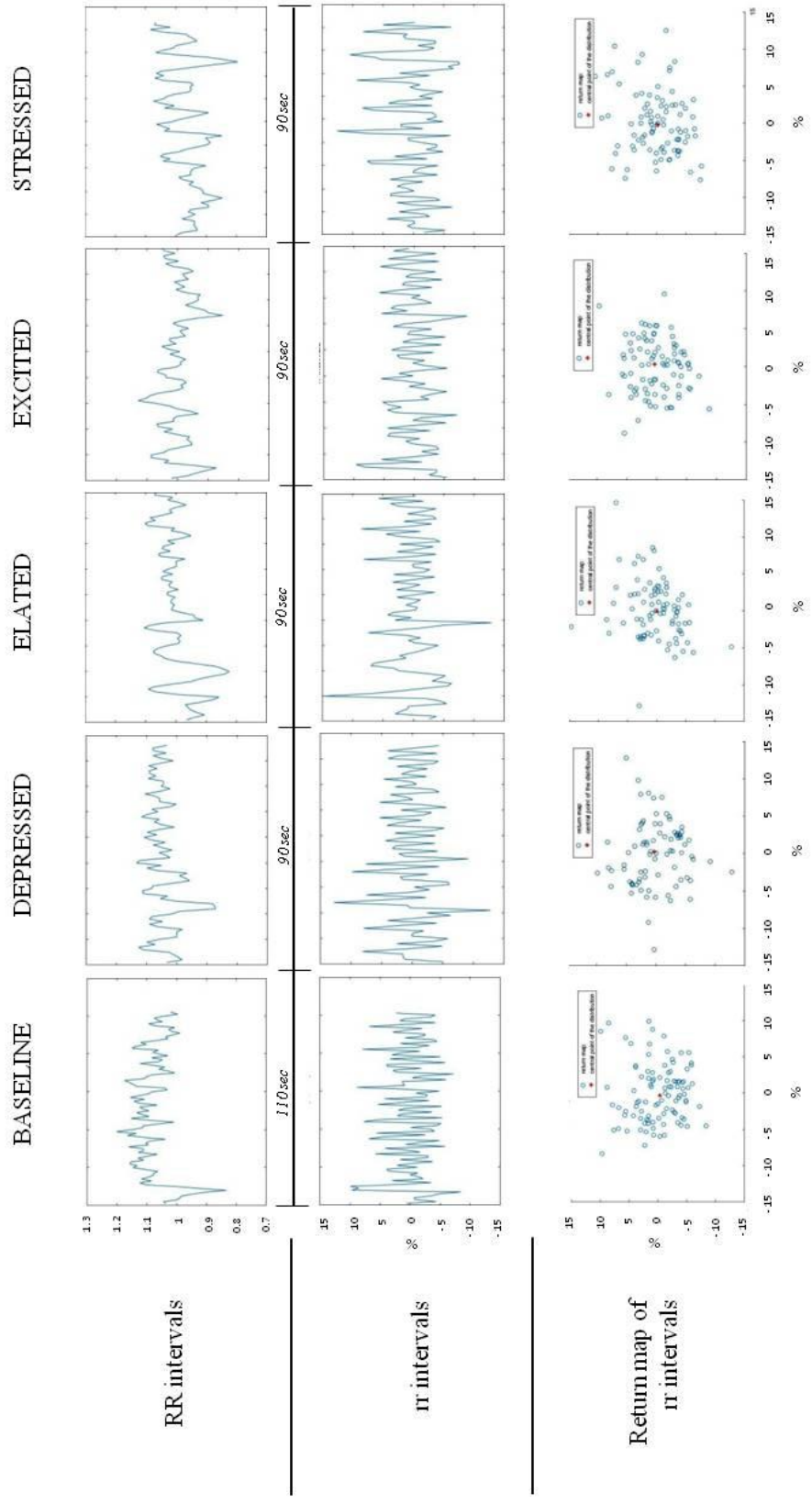


Figure 52 - RR intervals, rr intervals and return map of rr intervals in each session of a representative subject

4.2.2 ECG Analysis : Frequency domain

From the RR signal previously showed, Point process algorithm is used to evaluate instantaneous power spectra along the signal, then processed for each section.

Features in the frequency domain are then extracted: Low Frequency power spectrum (potLF), High frequency power spectrum (potHF), sympatho-vagal balance (bal), power spectra normalized on the total one (LFnorm, HFnorm).

Moreover, variance in HR and RR signal are evaluated.

Results from a single subject are shown in the Figure 53.

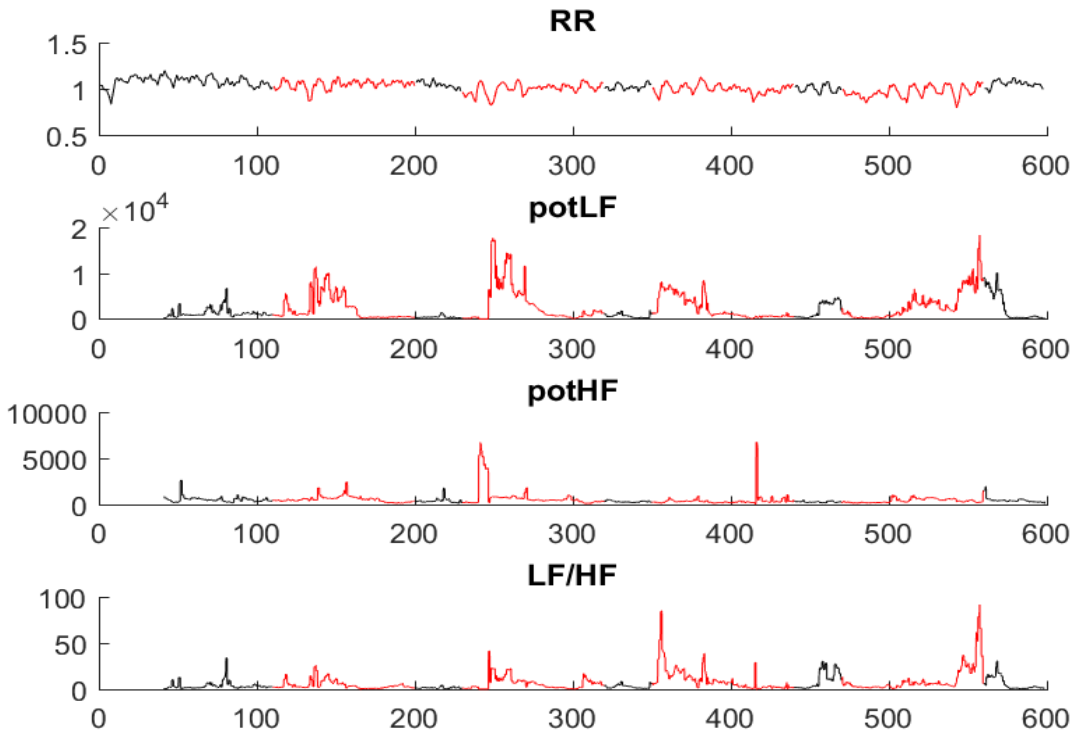


Figure 53 - Plot of the 10 minutes cycle RR, potLF, potHF and LF/HF ratio, in time, with session division (red for arousal session and black for baseline and neutral sessions)

From the Figure 53 it is easy to see how power in LF band (potLF) increases in the emotional sections (red parts), highlighting activation of sympathetic system, as the sympatho-vagal balance (LF/HF) confirms. Total power is not considered since it is a reflection of humoral effects and circadian rhythm, useful in 24h evaluations. LF power and HF power spectra of ECG signal for the 10 minutes signal are shown in Figure 54 and 55, while power spectra for each of the 4 arousal sessions and the baseline session are shown in Figure 56.

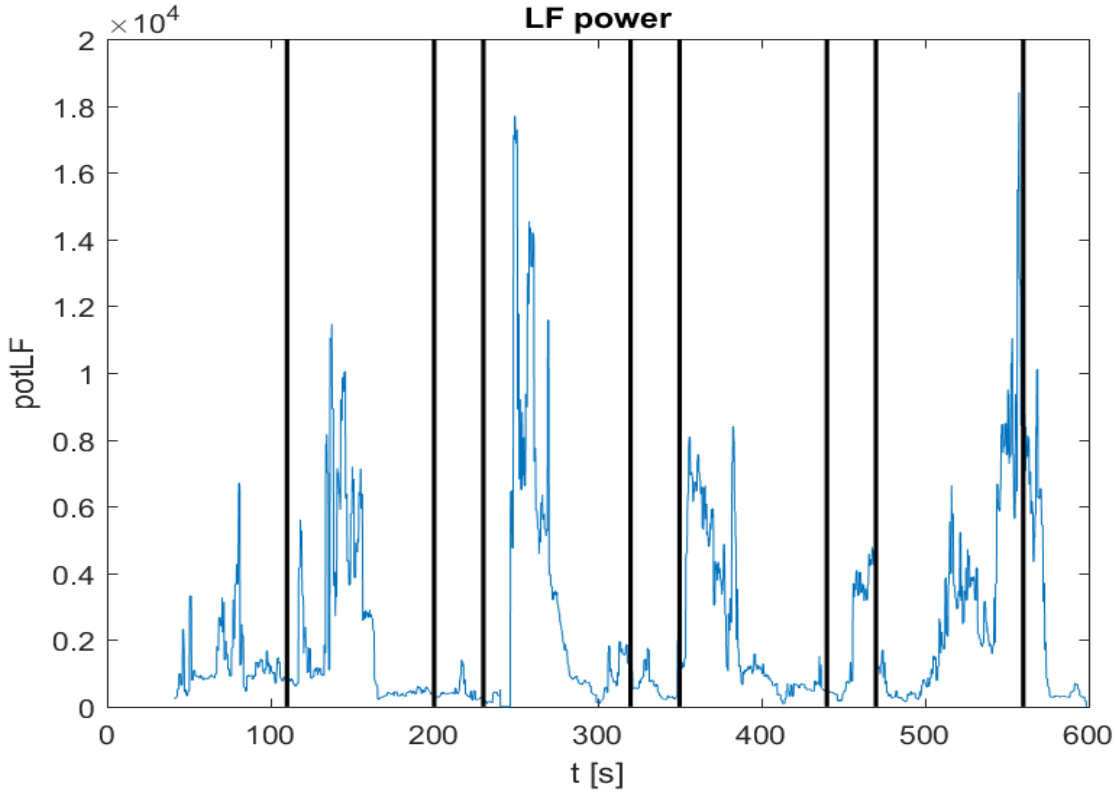


Figure 54 - LF power plot for the whole 10 minutes signal

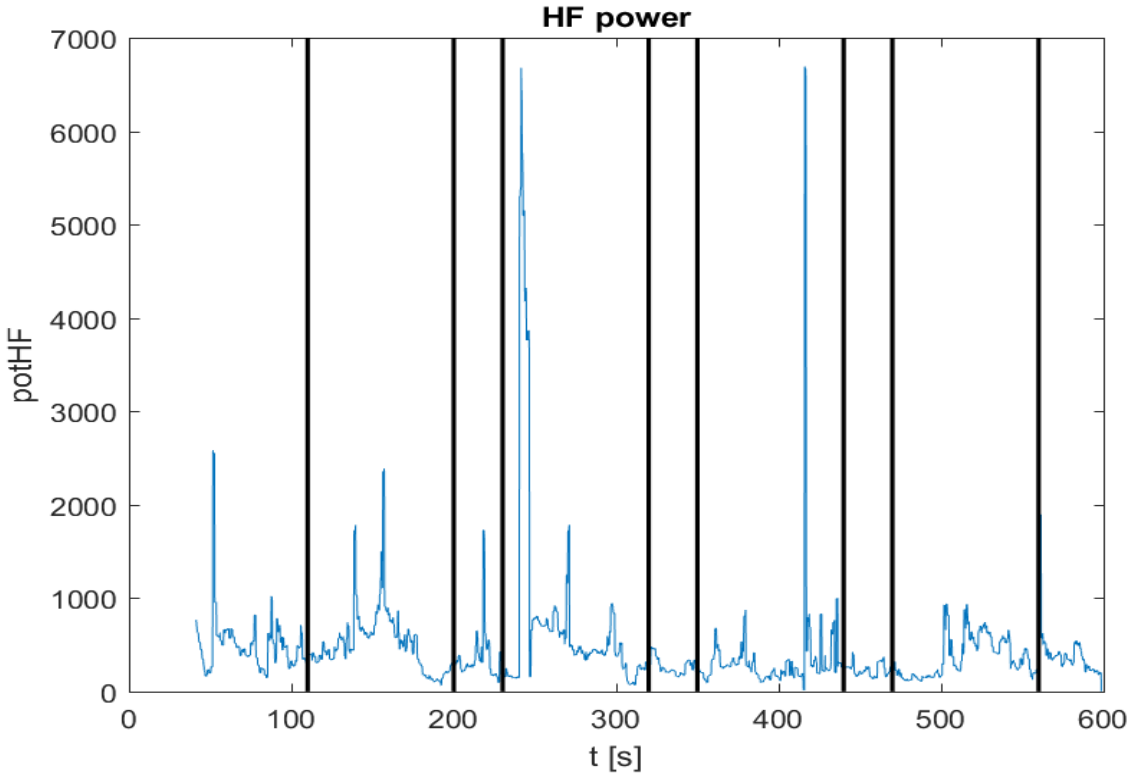


Figure 55 - HF power plot for the whole 10 minutes signal

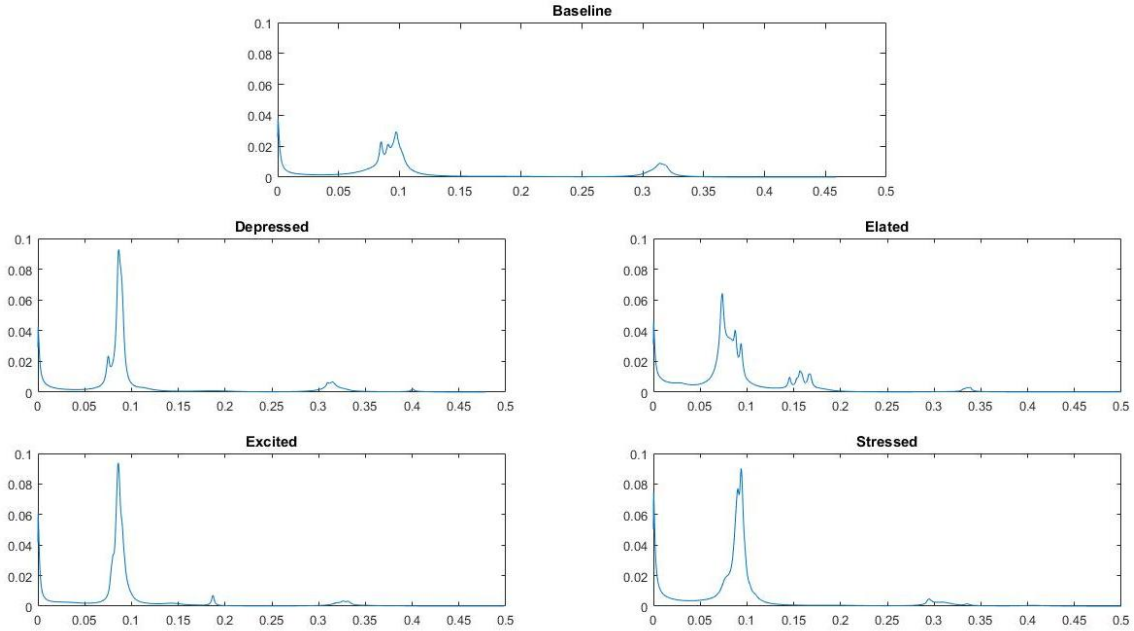


Figure 56 - R-R power spectra in the 5 sessions

4.2.3 Respiratory signal analysis

In this case too, analysis of the respiratory signal starts from the detection of peaks, evaluating their amplitude, with medium values and standard deviation for each section.
Power spectra are computed, through an autoregressive algorithm, extracting, as an informative feature, the HF component, with the peak position, its height, and the power spectrum in the HF band. Figure 57 shows power spectra for Respiratory signal in each arousal session and in the baseline session.

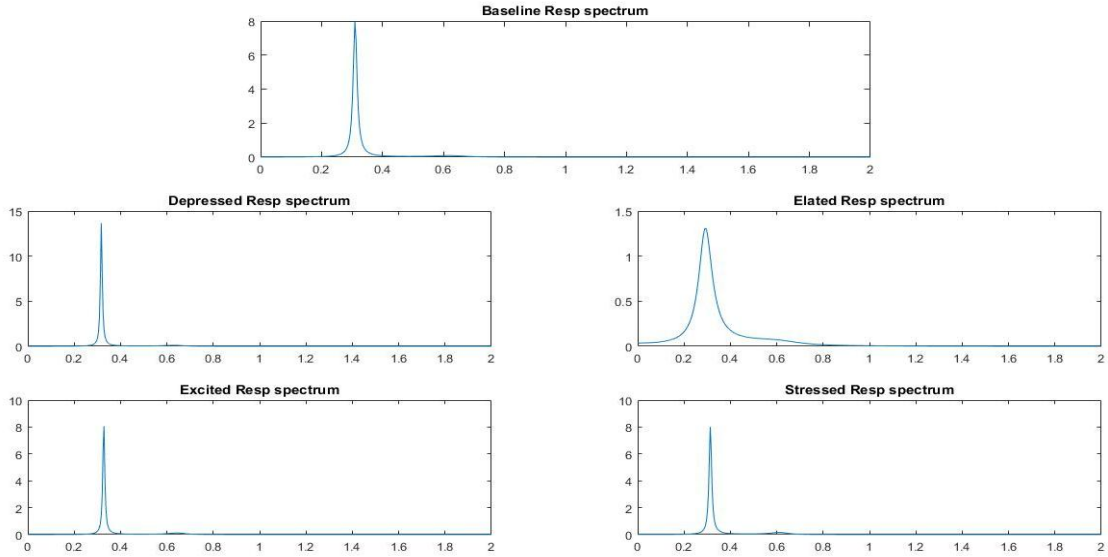


Figure 57 - Respiration signal power spectra in the five sessions

Just to stress what said before, in the depressed, stressed and excited sessions it is clear how the LF band is strongly stimulated, eliciting an activation in the sympathetic system that is visible from the increasing in the peak and the power in this frequency area.

4.2.4 Skin Conductance analysis

Through the algorithm explained in Materials and Methods section, following results have been obtained: Figure 58 shows the raw acquired signal, Figure 59 shows the resampled one (20 Hz), Figure 60 shows the signal after differentiation, Figure 61 the results of the convolution with a 20-points Bartlett window and Figure 62 shows the zero crossing detection.

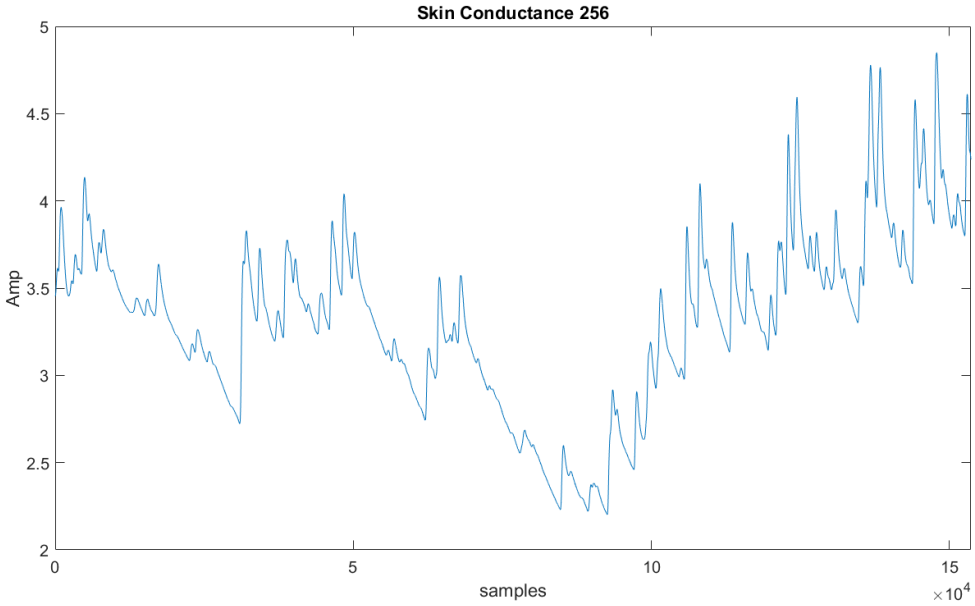


Figure 58 - Skin Conductance signal (10 minutes)

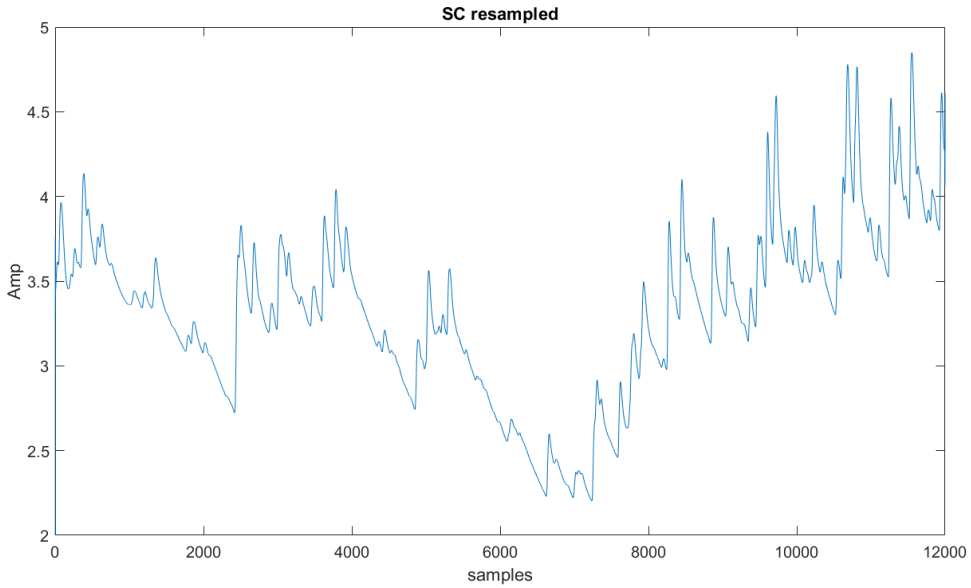


Figure 59 - Skin Conductance signal resampled at 20 Hz (10 minutes)

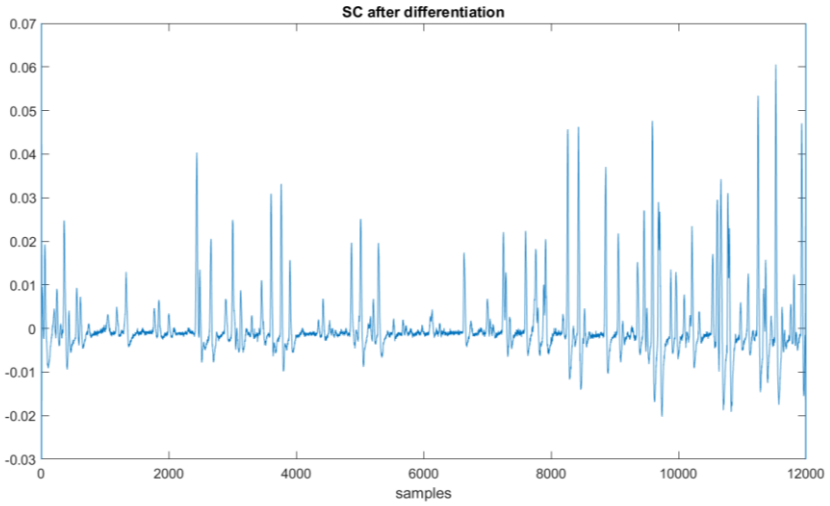


Figure 60 - Skin Conductance signal after differentiation

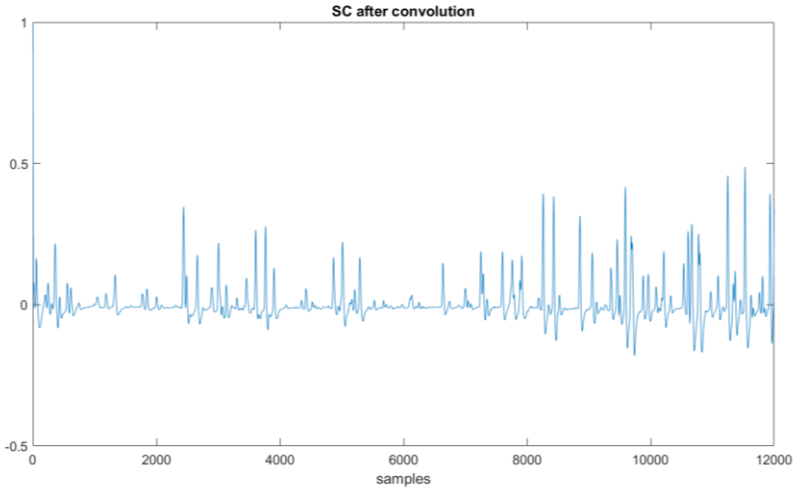


Figure 61 - Skin conductance signal after convolution with Bartlett window

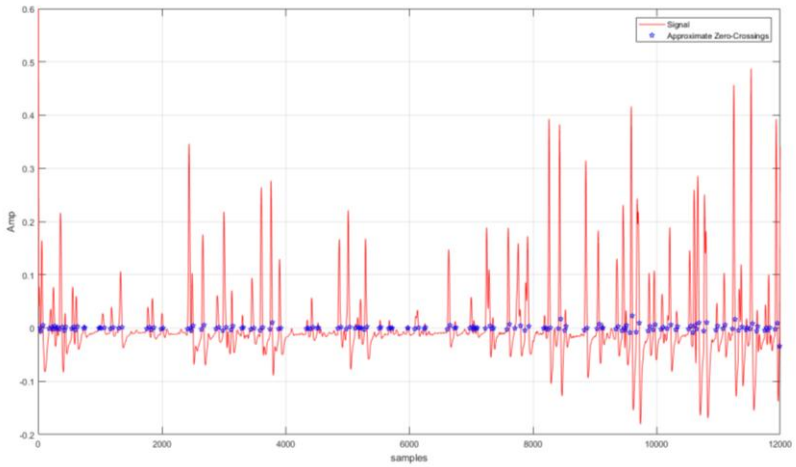


Figure 62 - Skin Conductance: zero crossing detection

On the last signal peaks detection is evaluated, for each of the section of interest, evaluating features and parameters: peaks occurrences in each section, maximum peak height, mean values of prominence and width (at half-prominence). In Figure 63, peaks detection is represented: peaks with an amplitude lower than 10% of the maximum are not considered in the analysis. Then, in Figure 64, width and prominence values are represented.

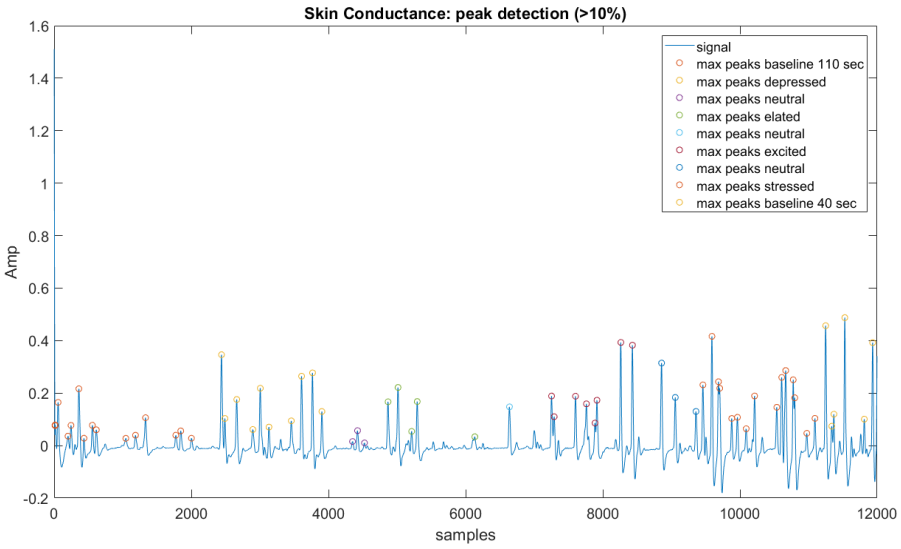


Figure 63 - Skin Conductance signal: Peak Detection

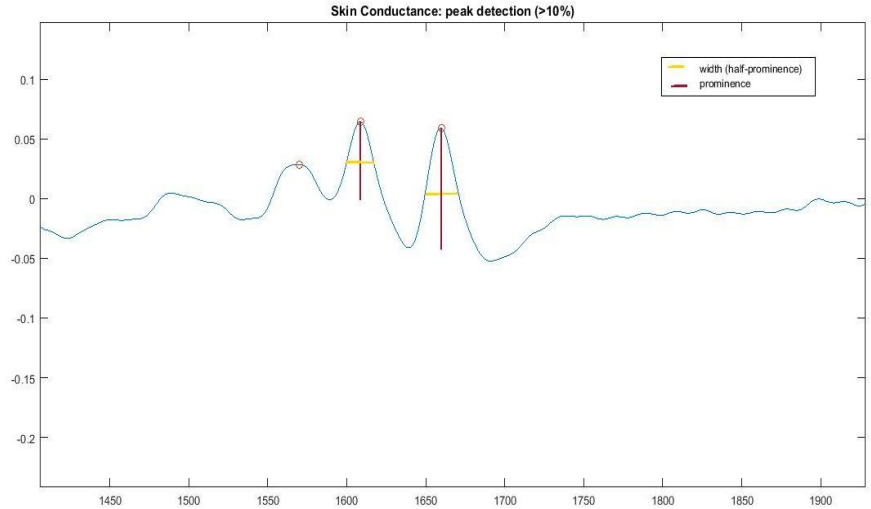


Figure 64 - Skin Conductance peaks: width at half prominence and prominence height

4.3 Bivariate Signal Analysis

4.3.1 ECG & Respiration

Bivariate analysis was performed, between cardiac signal and respiration signal; then the coherence was calculated.

In Figure 65, 66, 67, 68, 69, for each arousal session and for the baseline, power spectrum of RR signal, power spectrum of respiration signal and the computed coherence between these two, are shown

Coherence overcomes threshold value (0.5) in each represented session, showing a clear interdependence between the two considered signals, and an evident influence of the respiratory signal on the ECG one.

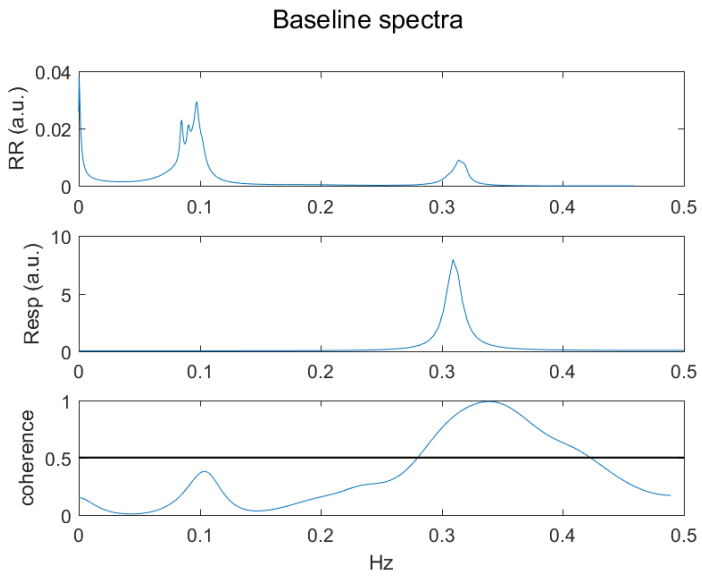


Figure 65 - top: RR baseline power spectrum; middle: Respiratory signal baseline power spectrum; down: coherence

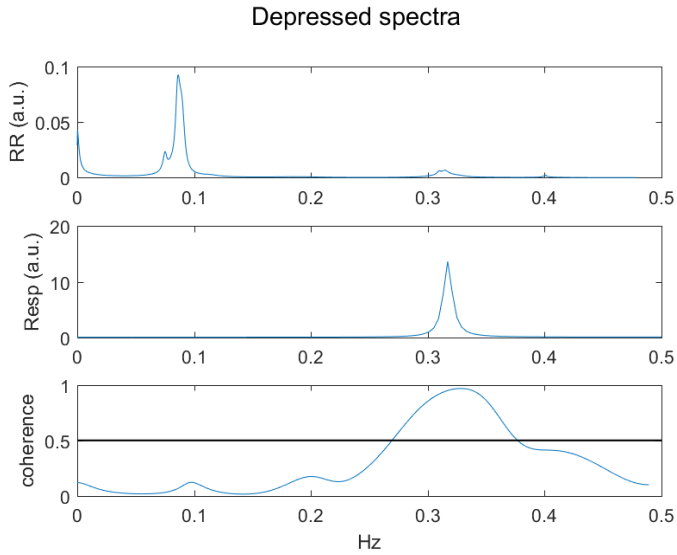


Figure 66 - top: RR depressed power spectrum; middle: Respiratory signal depressed power spectrum; down: coherence

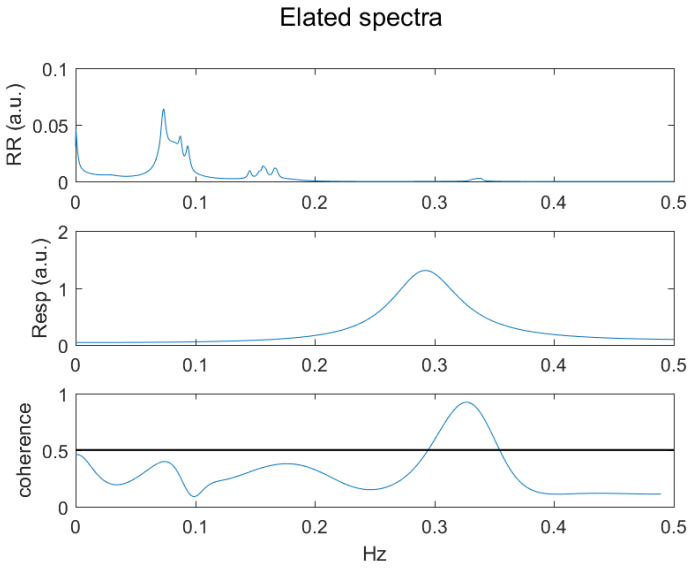


Figure 67 - top: RR elated power spectrum; middle: Respiratory signal elated power spectrum; down: coherence

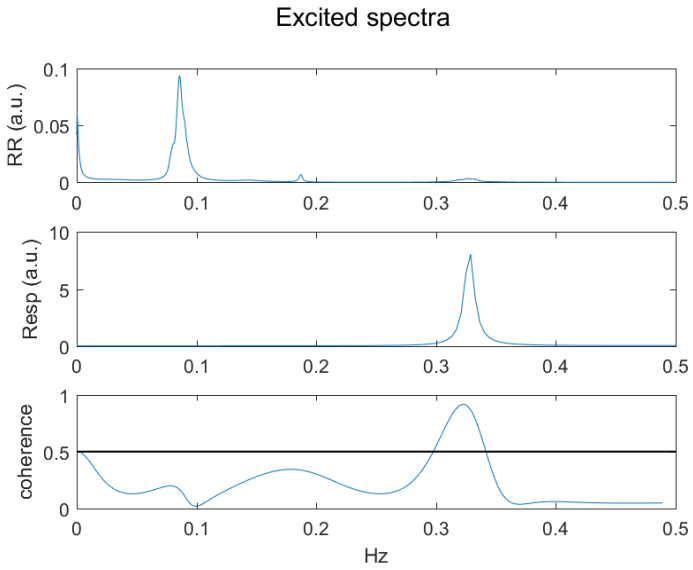


Figure 68 - top: RR excited power spectrum; middle: Respiratory signal excited power spectrum; down: coherence

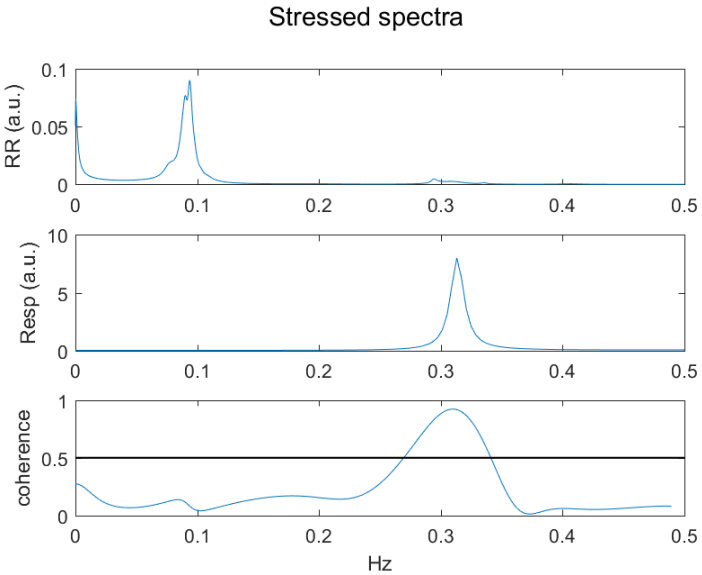


Figure 69 - top: RR stressed power spectrum; middle: Respiratory signal stressed power spectrum; down: coherence

4.4 Multiple Subjects analysis

The analysis presented in the previous section has been developed on each single subject. In this way variables with extracted features have been made in each area of interest (the five sections of our protocol).

Then a "vertical" analysis has been made, and for each variable the mean value is calculated, using all the values of the same variables in all the samples, trying to find a trend or a precise behavior.

All these variables are then plotted using boxplot, table and error-bars, to underline or identify these informative trends, eventually useful to interpret emotional responses of a subject.

4.4.1 ECG: Time domain analysis

		Emotion				
		Baseline	**Depressed**	**Elated**	**Excited**	**Stressed**
Temporal Index	**meanRR**	0.8243 ± 0.0904	0.8056 ± 0.0928	0.7649 ± 0.0906	0.7775 ± 0.0913	0.7802 ± 0.1022
	SDNN	0.0547 ± 0.0154	0.0501 ± 0.0128	0.0454 ± 0.0149	0.0458 ± 0.0143	0.0480 ± 0.0148
	RMSSD	0.0386 ± 0.0081	0.0329 ± 0.0130	0.0266 ± 0.0078	0.0315 ± 0.0091	0.0311 ± 0.0095
	pNN50	25 ± 15	15 ± 11	10 ± 7	11 ± 10	13 ± 10
	TI	75 ± 19	76 ± 27	103 ± 29	64.5 ± 13	95 ± 23.5000
	SD1	0.0152 ± 0.0035	0.0141 ± 0.0051	0.0119 ± 0.0035	0.0139 ± 0.0049	0.0141 ± 0.0037
	SD2	0.0731 ± 0.0183	0.0629 ± 0.0171	0.0603 ± 0.0176	0.0586 ± 0.0175	0.0141 ± 0.0037
	SD12	0.2436 ± 0.0515	0.2562 ± 0.0507	0.2248 ± 0.0269	0.2431 ± 0.0441	0.2128 ± 0.0358
	E	30.1883 ± 11.9733	29.9524 ± 12.3310	35.1402 ± 11.9699	35.0811 ± 11.0903	33.1075 ± 13.6626

Table 7- Time domain parameters of the population (values expressed as X=median(X) ± MAD(X))

As explained for the single subject analysis, time domain evaluation follows state of the art findings: frequency increases in stressed and excited sections, while in depressed session remains pretty similar to baseline value. RMSSD follows mean RR behaviour, as expected. Values of temporal parameter are shown in Table 7 and are clearly represented in the boxplots (Figs. 71-74) and errorbars (Figs. 75, 76), where these values are compared between baseline and each emotional response. Figure 77 shows RMSSD of a group of subjects with a meanRR value in a specific range [0.6 - 0.92 seconds]: in each of the 4 plot include in the figure, regression lines for baseline and emotion are plotted and the value of the slope is indicated.

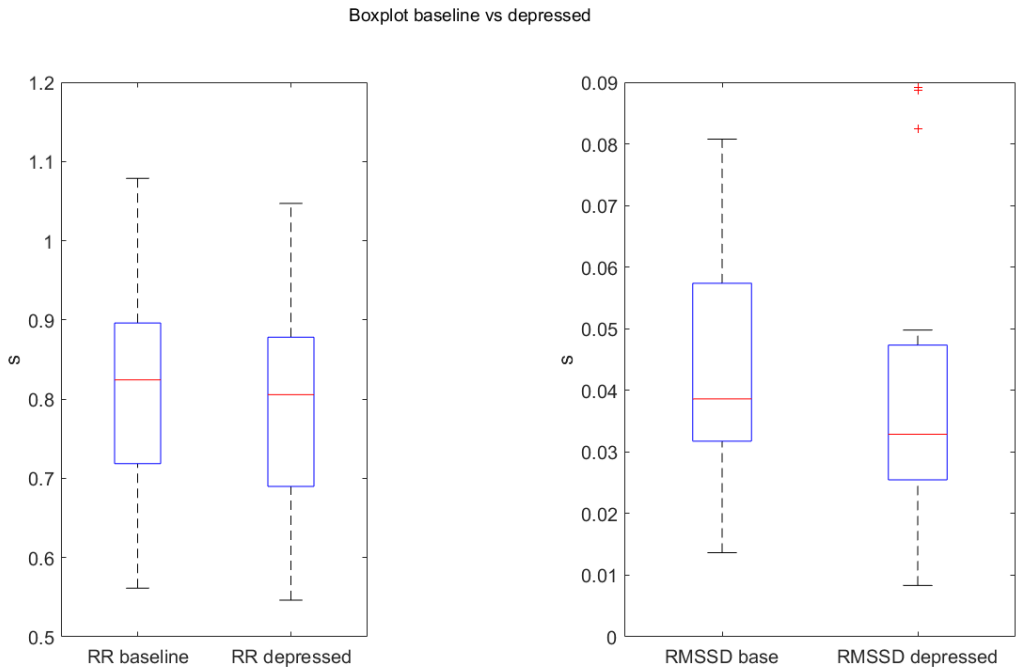


Figure 70 - Boxplots for baseline and depressed sessions comparison

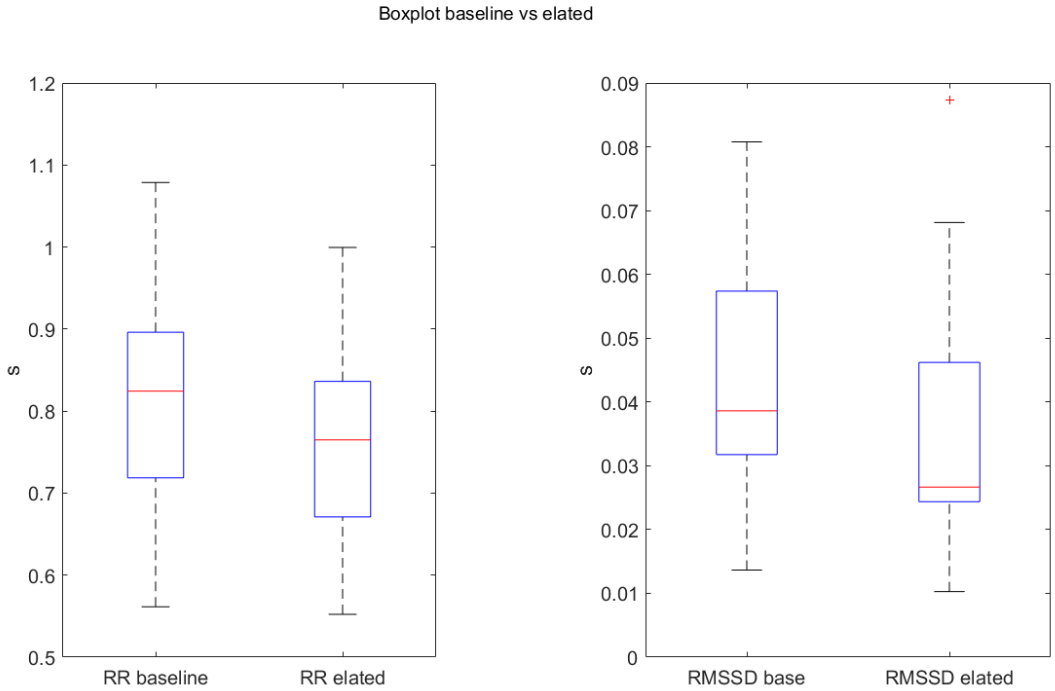


Figure 71 - Boxplots for baseline and elated sessions comparison

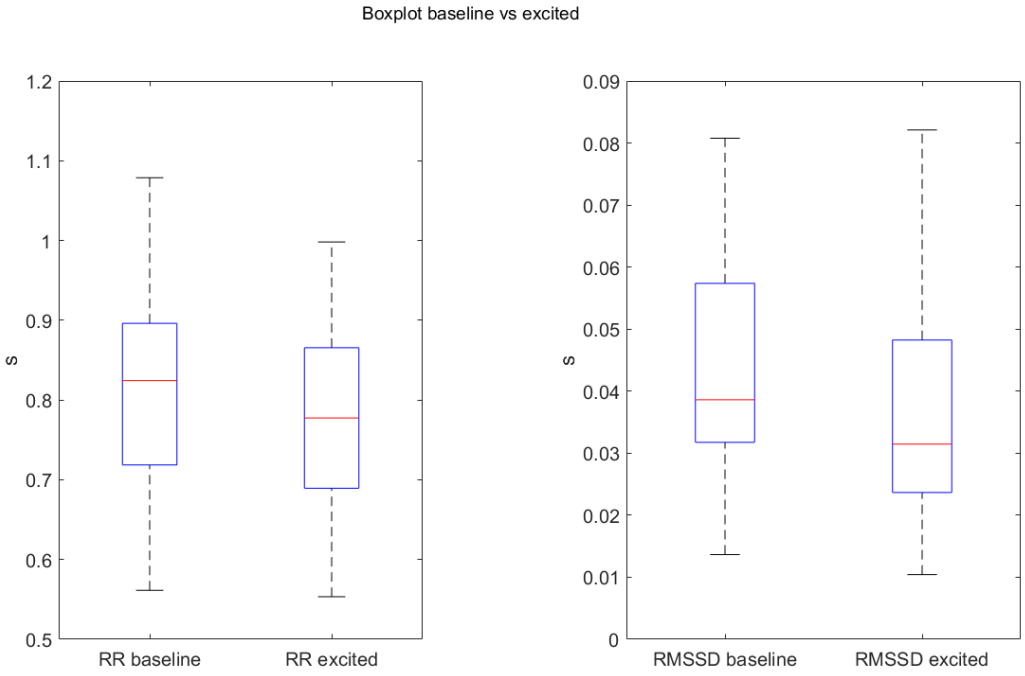


Figure 72 - Boxplots for baseline and excited sessions comparison

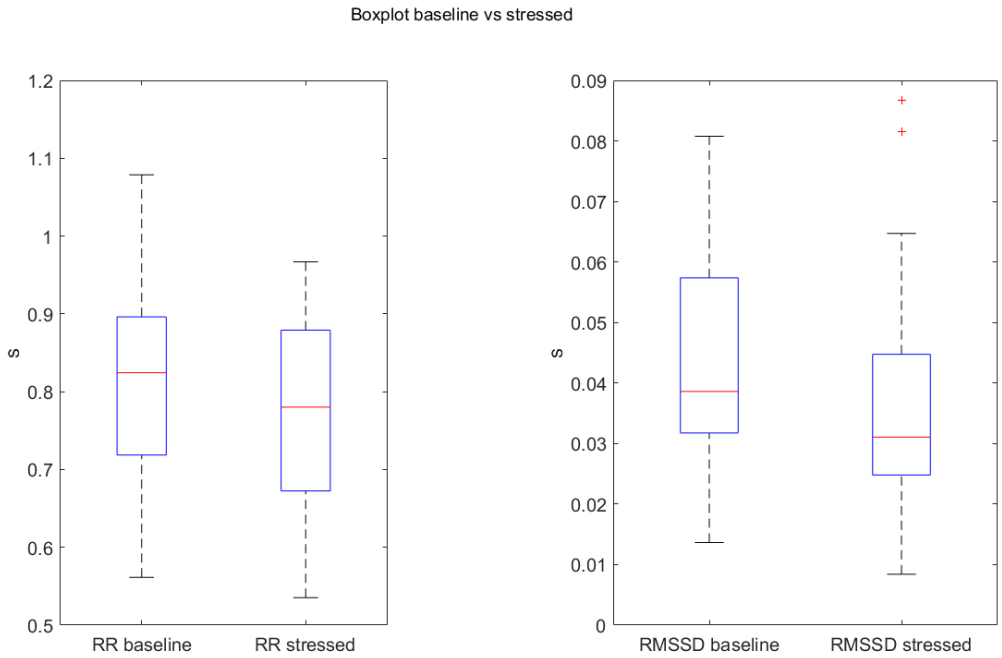


Figure 73 - Boxplots for baseline and stressed sessions comparison

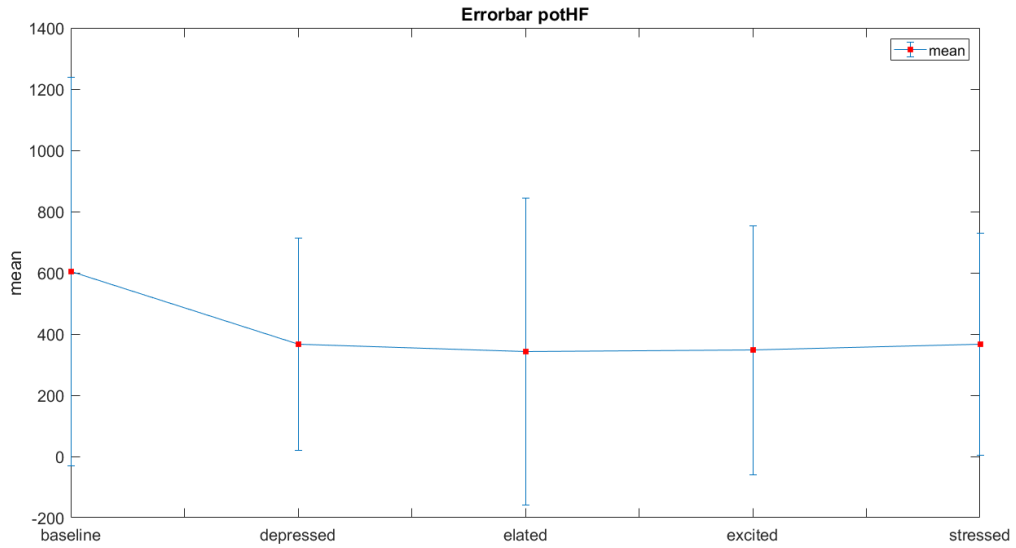


Figure 74 - Errorbar of potHF parameter in the 5 sessions

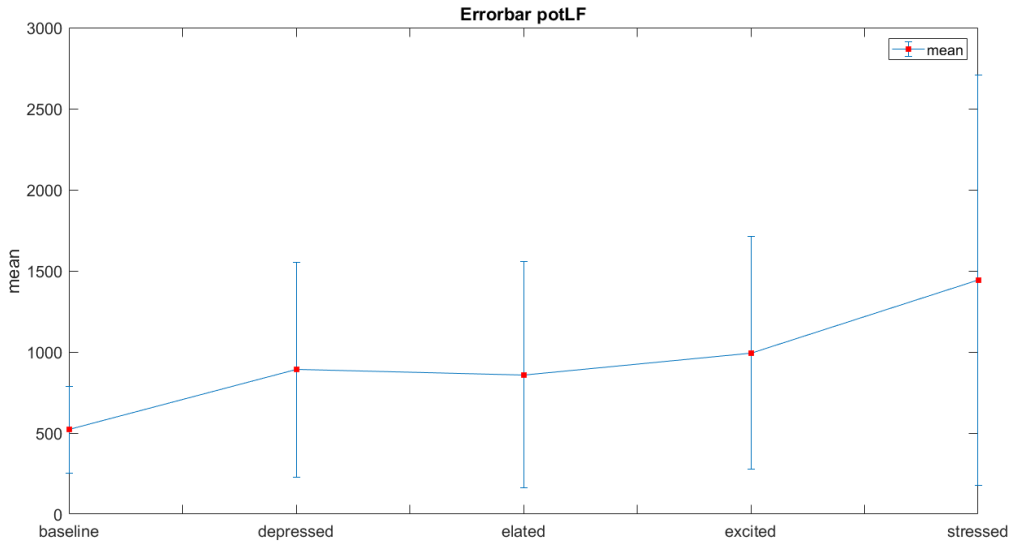


Figure 75 - Errorbar of potHF parameter in the 5 sessions

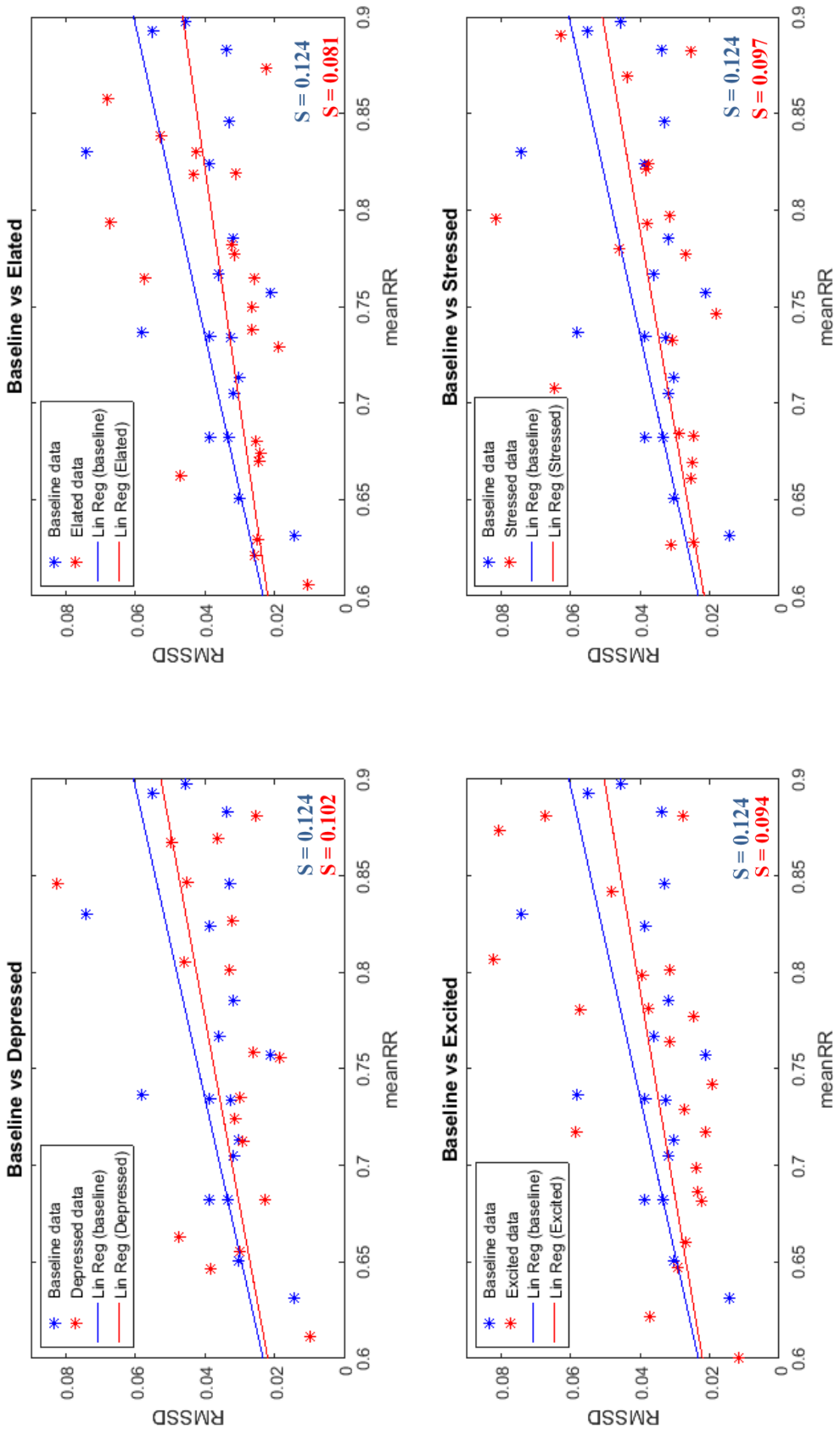


Figure 76 - Representation of RMSSD a group of subjects with a meanRR value in a specific range [0.6 - 0.92 seconds]. Regression line slope values are indicated as "s" at the bottom-right of each figure

4.4.2 ECG: Frequency domain analysis

Frequency domain features are represented in Table 8, through their mean values and standard deviations, for each of the investigated emotions. Same values are represented through boxplot and error-bar distributions.

		Emotion				
		Baseline	Depressed	Elated	Excited	Stressed
Spectral Index	varRR	194.4172 ± 89.7000	206.5047 ± 115.1415	191.9823 ± 94.1960	191.3231 ± 108.9942	227.0053 ± 126.5356
	varHR	2.4907 ± 0.9209	2.6823 ± 0.9135	2.6294 ± 1.1067	2.7260 ± 0.8693	3.2356 ± 1.1994
	potLF	351.1406 ± 122.1345	678.3191 ± 455.3015	463.0558 ± 208.3211	498.7588 ± 276.4605	655.5100 ± 316.7809
	potHF	208.7474 ± 129.0672	224.4643 ± 153.1774	177.5722 ± 120.9561	194.8624 ± 144.3933	310.9350 ± 200.5481
	potLFnorm	0.6426 ± 0.1658	0.6450 ± 0.1108	0.7188 ± 0.1734	0.6965 ± 0.1378	0.6906 ± 0.1068
	potHFnorm	0.3843 ± 0.1251	0.4046 ± 0.1065	0.4105 ± 0.1701	0.3668 ± 0.1331	0.4246 ± 0.1186
	potbal	2.9264 ± 2.0930	4.0242 ± 2.4784	4.0696 ± 2.5205	3.7074 ± 2.5898	3.6527 ± 2.2624

Table 8 - ECG spectral parameters of the population (values expressed as X=median(X) ± MAD(X))

Frequency indexes are evaluated for a group of people, choosing the one with a baseline value considered as “normal” with respect to state of art. Some subjects present abnormal values due to “experimental artefacts, with alterations probably caused by caffeine, smoke or time. Nevertheless, results reflect state of art and what previously described for the single subject analysis, that is an increasing in the LF power for the exciting and stressed states, and for depressed section too. After developing a normality test, verifying the non - normality of examined populations, a Wilcoxon test has been performed. In this test emotional responses are evaluated in a coupled analysis, starting from the comparison between baseline and each of the other state. So in Table 9, “b vs d” stands for baseline vs depressed, “b vs el” for baseline vs elated, “b vs ex” for baseline vs excited, “b vs s” for baseline vs stressed, and so on with the others represented in the table above.

As expected, baseline is different in a statistically significant way from the other emotions, but not from the elated state. This is underlined, considering LF power, by a p-value less than 0.5 in the test with excited and stressed response, and inferior to 0.7 with reference to depressed one. Considering normalized power in LF and HF bands, and balance too, there is statistically significant difference between baseline and the other emotions with a very very low p-value (less than 0.03 in all these evaluations).

	Index						
	varRR	**varHR**	**potLF**	**potHF**	**potLFnorm**	**potHFnorm**	**bal**
b vs d	0, 7710	0, 6979	0, 0681	0, 4151	0, 0117	0, 0257	0, 0137
b vs el	0, 9845	1, 0000	0, 1031	0, 1511	0, 0093	0, 0104	0, 0046
b vs ex	0, 7858	0, 6979	0, 0283	0, 2523	0, 0046	0, 0104	0, 0055
b vs s	0, 3417	0, 4151	0, 0010	0, 3618	0, 0008	0, 0270	0, 0008
d vs s	0, 5605	0, 6276	0, 1160	0, 9227	0, 2366	0, 9536	0, 1511
d vs ex	0, 9073	0, 9845	0, 7269	0, 4492	0, 6695	0, 5737	0, 7124
el vs s	0, 4151	0, 4609	0, 0653	0, 2604	0, 3224	0, 7562	0, 3721
el vs ex	0, 7562	0, 7710	0, 4971	0, 7269	0, 8614	0, 9845	0, 9073
s vs ex	0, 5346	0, 5219	0, 2860	0, 4377	0, 5737	0, 7562	0, 4377

Table 9 - p-values of spectral features of ECG signal, after ranksum test

4.4.3 Respiration signal analysis

On the respiratory signal extracted features are: high frequency power spectrum, maximum position and maximum height. Their mean values (and standard deviations) are represented in Table 10.

	Emotion				
	Baseline	**Depressed**	**Elated**	**Excited**	**Stressed**
Peak freq	0. 2709 ± 0. 0317	0. 2861 ± 0. 0287	0. 2879 ± 0. 0283	0. 2982 ± 0. 0248	0.2764 ± 0.0311
Pot	0. 1675 ± 0. 0860	0. 1604 ± 0. 0872	0. 1817 ± 0. 1027	0. 1722 ± 0. 0962	0.1740 ± 0.0900
Peak Amp	5. 3834 ± 3. 8224	6. 9770 ± 4. 9501	7. 0107 ± 4. 7260	7. 5332 ± 4. 9777	8.0507 ± 5.4259

Table 10 - Respiratory signal spectral parameters of the population (values expressed as X=median(X) ± MAD(X))

In this case too, Wilcoxon test is performed, to evaluate the independence among the population of subjects in each examined parameter. As showed in Table 11, baseline is statistically different from the other emotional responses, with p-values much lower than the threshold value (0.05) in all the comparisons between that and each emotion.
Stressed and excited responses also show a significant pvalue.

	Index		
	freq	pot	amp
b vs d	0,0006	0,4644	0,4308
b vs el	0,0003	0,5158	0,3262
b vs ex	0,0000	0,8370	0,5668
b vs s	0,0174	0,4831	0,0824
d vs s	0,2770	0,1650	0,3037
d vs ex	0,1566	0,6235	0,8043
el vs s	0,1897	0,9515	0,5173
el vs ex	0,2111	0,4014	0,6318
s vs ex	0,0256	0,3632	0,2387

Table 11 - p-values of spectral features of Respiration signal, after ranksum test

4.4.4 Skin Conductance analysis

Looking for a meaningful trend in the five sections, extracted parameters are presented here in Table 12 , with mean values and standard deviations. These parameters are then compared using boxplot and error-bars.

	Emotion				
	Baseline	Depressed	Elated	Excited	Stressed
Number of Peaks	7 ± 2	6 ± 2	8 ± 2	10 ± 2	9 ± 4
Mean Amplitude	0.0565 ± 0.0483	0.0642 ± 0.0483	0.0524 ± 0.0359	0.0610 ± 0.0416	0.0528 ± 0.0327
Max Amplitude	0.1364 ± 0.1096	0.1336 ± 0.0877	0.1161 ± 0.0792	0.1168 ± 0.0778	0.1087 ± 0.0598
Mean Width	9.5592 ± 2.802	9.3786 ± 2.930	8.6033 ± 2.606	9.9547 ± 3.518	11.4312 ± 2.931
Mean Prominence	0.0076 ± 0.0070	0.0081 ± 0.0075	0.0029 ± 0.0024	0.0074 ± 0.0060	0.0080 ± 0.0062

Table 12 - Skin Conductance signal spectral parameters of the population (values expressed as X=median(X) ± MAD(X))

SCR is an important physiological signal, also because its response is fast, but the shape and height of its peak do not change in a significant way: what is important to notice is the number of peaks for each section, that usually changes. Following the expectations, in the excited and the stressed part, number of peaks increase, while in the depressed one decreases.

Evaluating the Wilcoxon test, p-values confirm what previously said. The only parameter that changes in a significant way is the number of peaks, that allows to distinguish, from a statistical point of view, baseline from excited and stressed response, while it has a value greater than the threshold for the test with the depressed group. Interesting to notice how depressed reactions "overcomes" this test with a p-value lower than the threshold when compared to excited and stressed section, underlining how it is different the reaction of the skin conductance to negative and positive stimuli (Table 13).

	Index				
	peaks number	**mean amplitude**	**max amplitude**	**mean width**	**mean prominence**
b vs d	0,6496	0,7164	0,8899	0,9311	0,9448
b vs el	0,0655	0,4162	0,7293	0,9036	0,9448
b vs ex	0,0003	0,9173	0,7035	0,3683	0,3326
b vs s	0,0097	0,5918	0,7687	0,1462	0,5221
d vs s	0,0097	0,3776	0,6159	0,1560	0,3870
d vs ex	0,0019	0,7819	0,9586	0,4363	0,2835
el vs s	0,3137	0,9036	0,8763	0,1368	0,3414
el vs ex	0,1373	0,4999	0,4261	0,3074	0,2535
s vs ex	0,7811	0,6038	0,5918	0,5564	0,9586

Table 13 - - p-values of Skin conductance signal features after ranksum test

4.5 Classifier

Two different databases (Fig 84), with different characteristics, were firstly created; 37 subjects have been chosen for database building, after questionnaire evaluation, signal visualization (looking for measurement errors), outlier elimination and grouping of a pool of subjects with a normal mean RR, with no bradycardic or tachicardic disorders.

Each database contains temporal and frequency indexes of different physiological signals.

In the first one ECG and Respiratory signal indexes have been included, but only the one that result statistically significant after the ranksum test: for the ECG signal, LF power, LFnorm and HFnorm power and balance have been chosen; for the Respiratory one, frequency in HF band.

The second database also includes SCR, with the number of peaks as related index.

	Database	Signal		
		ECG	RSP	SCR
median	Database 1	✓	✓	✓
	Database 2	✓	✓	X

Figure 78 - Database for classification

	Classification Learning			
Type/Data	**LDA**	**SVM**	**Logistic Regression**	**Ensembles**
Baseline vs depressed	> 70	75	> 70	> 65
Baseline vs Excited	> 70	> 75	78	> 70
Baseline vs Stressed	> 70	> 70	> 75	75
Depressed vs Stressed	> 65	> 70	> 70	> 65
Depressed vs Excited	> 65	> 70	> 70	> 70

Table 14 - Classification Learning results

Classification has been performed using "Classification Learner App in Matlab", using selected features from the ranksum test. Dataset have been compared between the most different responses. K-fold cross validation has been used, with k=5. Results are shown in Table 14.

Classification results have been furthermore investigated for the baseline vs excited and baseline vs stressed comparisons. Accuracy points out values greater than 80% thanks to a stronger stimulation of nervous system reaction.

Results are shown in Table 15.

	LDA	SVM Quadratic	KNN (medium KNN)	Ens(bagged trees)	AUC
Baseline vs Excited	77,10	81,3	81,3	83,3	0,89
	QDA	**SVM Linear**	**Logistic Regrssion**	**Ens(Subspace Discr)**	**AUC**
Baseline vs Stressed	83,30	79,2	72,9	79,2	0,83

Table 15 - Classification Learning results; LDA: Linear Discriminant Analysis; QDA: Quadratic Discriminant Analysis; SVM: Support Vector Machine; KNN: K-nearest neighbors; Ens: ensamble AUC: Area Under the Roc Curve

5. Discussion and Conclusion

All the data have been examined starting from the info extracted from the signals, either in the time and in the frequency domain; soon after a statistical analysis to test the independence of "emotional" populations, some parameters have been chosen. These features are the ones that result as statistically significant and these are the ones used for classification learner.

First analysis is the one between baseline and the four emotional responses in a one-to-one comparison, showing an accuracy percentage from 70% to 80% in almost all of the situations, with different classificatory types. Best results have been obtained in the evaluation of baseline with respect to stressed and excited sessions, and this is what state of art suggests and our first aim.

As described in the state of the art, talking about emotions, stress and excitation emotive responses are enthralling, also from a physiological point of view, because they activate our nervous system, causing reaction of easy detection. On the other side it is not as much simple ensuring a proper condition for the detection of the neutral state (baseline). For this very reason, each test required an initial await phase, during which the subject was asked to sit down and remain still in a comfortable position and relax before starting the test.

This resulted to be necessary to avoid the recording of misleading signals, related to situations occurred before the test. Each acquisition lasted 30 minutes, during which the subject was trained about the test conditions and dynamics, electrodes and sensors were placed and the quality of the signal was assessed. Then the presentation started. This process, conducted on 105 subjects, required not an indifferent work in terms of organization and precision. However, a similar experimental study constitutes a great result, thanks to the huge quantity of collected data which can allow, in future, to perform further analysis and classifications.

This is the starting point of our study, baseline is effectively distinguished from the other emotional sessions and this represents a good stimulus to go over with this protocol and these tests.

Important to notice is surely the good accuracy in separating stressed response from depressed one.

As is evident from our classification results, stressed and excited sessions are easily distinguished from baseline one, and discrimination from depressed session seems to be possible with a further analysis of extracted significant features.

Either in the comparison between baseline/stressed, and in baseline/excited one, accuracy reaches values greater than 80%, using QDA and ensemble (bagged trees type) respectively. The area under the Receiver Operating Characteristic Curve (ROC), indicated as AUC, reaches significant values (>0.8) showing great ability of discrimination; this means that our test correctly classifies populations with different reactions, in the case of the above mentioned emotional responses.

Stressed and excited condition, indeed, are characterized by an higher arousal value, highlighted by a stronger sympathetic activation: recognition of these responses, therefore, turned out to be easier than the one of depressive and elated stimulation. These last two emotions show lower differences from the baseline condition due to their weaker ability to trigger nervous system response.

Going deeper in the classification analysis, exploring different methods and techniques, starting from our results, could lead to a total discrimination among the five examined emotional states.

Results show and underline protocol potentiality, not yet fully exploited; this experimental method could let to the definition and correct identification of these four precise emotional responses: depression, stress, excitation and elation.

Signal analysis followed the expectations and state of art findings: statistical analysis confirmed good results, classification gave the last answer.

This last “act” of our work will be the starting point of a new work, trying to improve some aspects that could have influenced and affect correct acquisition and data processing.

First of all, subjects have been chosen from the same cultural field (most of the students are Biomedical Engineering ones), but variability in aging, gender, origins and background cannot be underestimated.

For instance, some people did not “overcome” the PHQ-9 Questionnaire, showing an unstable (maybe temporary) emotive and mental condition.

Then, it is proper to consider the experimental environment. There are some requirements to follow doing tests of this type, especially when physiological signals are involved.
Tests have been carried out in a university Lab (i.e. SPiNLabs, Politecnico of Milano), an environment full of people working, where social interaction probably low concentration, changing emotional factors.
"Boundary conditions" are always the main "features" to take into account: lights, voices, noises do not facilitate being focused on what to see, listen, feel.
Music has probably helped in reaching the "best" concentration level, but not the proper one. It has a powerful stimulus, but not sufficient.
Other aspects to be considered are the one regarding people habits: drinking coffees, smoking cigarettes, eating before the test. Trying to encourage them to follow these schema is not enough, sometimes timetables and planning have not helped to give people the best condition to be an appropriate tester.
Physiological reactions could change due to our habits, and digestion too, so another improvement could concern acquisition time, trying to follow circadian rhythm of testers.
Before each test, subject has been asked to not move or talk, giving the time to find the best sitting position and to relax and being comfortable.
Our first conclusion is to carry out same experiments, with the same set-up, in a different environment, isolated, with no bright lights or people talking; then, fingers sensors have to be changed, because of their low adherence to tester skin. Measurement errors are always present in an experimental protocol, but with better means better results are easy to be reached.
This is a new protocol, combining a more recent pictures database (GAPED one), a video (GoPro) and a music selection: first results state that there are great possibilities to get more, to pretend more, to reach more. Growth potential is evident, improvements could be easily obtained, starting from this protocol and this algorithm.

5.1 Conclusion

Our aim is to find a protocol that could let us to perfectly distinguish emotions, starting from the ones here examined, going over spreading our scopes and interests.

5.1.1 Future protocol developments

Test will be repeated in the same modality in a different environment, with and without music, trying to compare influence of boundary conditions, and music too.
Then, classification will be trained and examined again, with data acquired in this first work and the new ones, looking for differences, improvements, and other suggestions.
Once accuracy classification greater then 85-90% is reached, the second step will involve high-level model building, trying to obtain an algorithm able to prevent and identify people reactions, since emotion is considered as a "mental creation", a sort of brain interpretation of different physiological events.

5.1.2 High-level model & Application

This new model will be used in a neuromarketing test, involving wine tasting and preferences interpretation. Different protocol hypothesis will be tested, trying to discriminate people tastes without their reports.

5.1.3 Clinical applications

It is not only about neuromarketing. It is a clinical study too.
People emotions and their developments, our mind potential, brain interpretations, feeling, everything that make people what they are: this is a powerful field of study and applications, since many diseases, functional and cerebral disorders are linked to physiological changes and permanent modifications.
These studies are useful to assess physical and mental condition, to evaluate treatment developments and prognosis, to get a guideline, to help people.
From the cardiovascular diseases (myocardial infarction, angina pectoris, coronary disease…), to brain ones (epilepsy, brain injuries…), passing through depression, panic disorders, anxiety. Science can help people, looking for something is the only way to find it. And this is what research does.

Emotions have no voice, but we can get them to talk.

6. Bibliography

[1] W. B. Cannon, «The James-Lange Theory of Emotions: A Critical Examination and an Alternative Theory,» *The American Journal of Psychology,* vol. 39, n. 1/4, pp. 106-124, 1927.

[2] W. James, The Principles Of Psychology Volume II, 1890.

[3] R. S. L. Craig A. Smith, «Emotion and Adaptation (Chapter 23),» in *Handbook of Personality: Theory and Research*, New York, L. A. Pervin, 1990, pp. 609-637.

[4] A. R. Damasio, Descartes' Error: Emotion, reason, and the Human Brain, New York: Avon Books, 1994.

[5] E. V. a. J. H. Rosalind W. Picard, «Toward Machine Emotional Intelligence: Analysis of Affective Physiological State,» *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 23, n. 10, p. 2001.

[6] P. S. a. D. R. C. John D. Mayer, «Emotional Intelligence: New Ability or Eclectic Traits?,» *American Psychologist,* vol. 63, n. 6, pp. 503-517, 2008.

[7] S. D. Kreibig, «Autonomic nervous system activity in emotion: A review,» *Biological Psychology,* n. 84, pp. 394-421, 2010.

[8] M. G. a. A. A. Chiba, «The Amygdala and Emotion,» *Current Opinion in Neurobiology,* n. 6, pp. 221-227, 1996.

[9] L. Pessoa, «Emotion and Cognition and the Amygdala: From "what is it?" to "what's to be done?",» *Elsevier, Neuropsychologia,* vol. 48, pp. 3416-3429, 2010.

[10] G. Richter-Levin, «The Amygdala, the Hippocampus, and Emotional Modulation of Memory,» *The Neuroscientist,* vol. 10, n. 1, 2004.

[11] B.-j. P. M.-S. P. S.-H. K. a. J.-H. S. Eun-Hye Jang, «analysis of physiological signals for recognition of boredom, pain, and surprise emotions,» *Journal of Physiological Anthropology,* vol. 34, n. 25, 2015.

[12] M. C. b. N. C. a. P. J. L. Margaret M. Bradley, «Emotion and Motivation I: defensive and Appetitive Reactions in Picture Processing,» *American Psychological*

Association, vol. 1, n. 3, pp. 276-298, 2001.

[13] P. Redding, «Feeling, Thought and Orientation: William James and The Idealist Anti-Cartesian Tradition,» *Parrhesia,* n. 13, pp. 41-51, 2011.

[14] R. B. M. M. V. R. a. L. M. Francesco Onorati, «Characterization of affective states by pupillary dynamics and autonomic correlates,» *Frontiers in Neuroengineering,* vol. 6, n. 9, 2013.

[15] j. A. R. a. B. S. P. Jonathan Posner, «The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology,» *Development and Psychopathology,* vol. 17, pp. 715-734, 2005.

[16] F. N. E. A. &. R. K. Hauke Egermann, «Continuous Measurement of Musically-Induced Emotion: A Web Experiment,» *International Journal of Internet Science,* vol. 4, n. 1, pp. 4-20, 2009.

[17] P. S. a. F. R. Julian Kampfe, «The Impact of Background Music on Adult Listenters: A meta-analysis,» *Psychology of Music,* vol. 39, n. 4, pp. 424-448, 2010.

[18] P. N. J. a. P. Laukka, «Expression, Perception, and Induction of Musical Emotions: A review and a Questionnaire Study of Everyday Listening,» *Journal of New Music Research,* vol. 33, n. 3, pp. 217-238, 2010.

[19] J. Panksepp, «The Emotional Sourves of Chills induced by Music,» vol. 13, n. 2, pp. 171-207, 1995.

[20] P. N. Juslin, «Music and Emotion: seven Questions, Seven Answers,» in *Music and the Mind: Essays in Honour of John Sloboda*.

[21] A. L. E. P. S. G. Valenza, «The Role of Nonlinear Dynamics in Affective Valence and Arousal Recognition,» *IEEE Transactions of Affective Computing,* vol. 3, n. 2, 2012.

[22] A. L. a. E. P. S. Gaetano Valenza, «Improving Emotion Recognition Systems by embedding Cardiorespiratory Coupling,» *IOPscience,* vol. 34, pp. 449-464, 2013.

[23] K. R. Elise S. Dan-Glauser, «The Geneva affective picture database (GAPED): a new 730-picture database focusing on valence and normative significance,» vol. 43, pp. 468-477, 2011.

[24] P. J. Lang, «The Emotion Probe. Study of Motivation and Attention,» *American Psychologist,* vol. 50, n. 5, pp. 372-385, 1995.

[25] S. H. A. L. a. P. J. L. Margaret M. Bradley, «Brain potentials in perception: Picture complexity and emotional arousal,» *Psychophysiology*, vol. 44, pp. 363-373, 2007.

[26] P. J. L. Margaret M. Bradley, «The International Affective Picture System (IAPS) in the Study of Emotion and Attention,» in *Handbook of Emotion Elicitation and Assessment*, J. A. Coan, and J. J. B. Allen, 2007, pp. 29-46.

[27] L. C. A. L. E. P. S. a. R. B. G. Valenza, «A Nonlinear Heartbeat Dynamics Model Approach for Personalized Emotion Recognition,» in *35th Annual International Conference of the IEEE EMBS*, 2013.

[28] L. C. A. L. E. P. S. a. R. B. Gaetano Valenza, «Revealing Real-Time Emotional responses: a Personalized Assessment based on Heartbeat Dynamics,» *Scientific Reports*, 2014.

[29] l. L. a. J.-h. Chen, «Emotion Recognition Using Physiological Signals,» pp. 437-446, 2006.

[30] R. W. L. L. c. F. H. W. a. J. J. G. Iris B. Mauss, «The Tie That Binds? Coherence Among Emotion Experience, Behavior, and Physiology,» *American Psychological Association*, vol. 5, n. 2, pp. 175-190, 2005.

[31] P. E. a. W. V. F. Mark G. Frank, «Behavioral Markers and Recognizability of the Smile of Enjoyment,» *Journal of Personality and Social Psychology*, vol. 64, n. 1, pp. 83-93, 1993.

[32] E. L. R. a. P. Ekman, «Coherence between Expressive and Experiential Systems in Emotion,» *Cognition and Emotion*, vol. 8, n. 3, pp. 201-229, 1994.

[33] S. W. B. S. R. K. K. H. Kim, «Emotion recognition system using short-term monitoring of physiological signals,» *Medical & Biological Engineering & Computing*, vol. 42, pp. 419-427, 2004.

[34] M. Vollmer, «A robust, simple and reliable measure of Heart Rate Variability using relative RR intervals,» in *Computing in Cardiology Conference (CinC), 2015*, Nice, France, 2015.

[35] D. G. a. K. R. S. Marcel Zentner, «Emotions Evoked by the Sound of Music: Characterization, Classification, and Measurement,» *American Psychological Association*, vol. 8, n. 4, pp. 494-521, 2008.

[36] M. H. T. William B. Davis, «The Influence of Preferred Relaxing Music on Measures of State Anxiety, Relaxation, and Physiological Responses,» *The National Association for Music Therapy,* vol. XXVI, n. 4, pp. 168-187, 1989.

[37] J. F. T. a. L. J. P. V. D. Ivan Nyklicek, «Cardiorespiratory differentiation of musically-induced emotions,» *Journal of Psychophysiology,* vol. 11, pp. 304-321, 1997.

[38] W. E. J. K. a. N. S. Rickard, «Relaxing Music Prevents Stress-Induced Increases in Subjective Anxiety, Systolic Blood Pressure, and Heart Rate in Healthy Males and Females,» *Journal of Music Therapy,* vol. XXXVIII, n. 4, 2001.

[39] T. F. D. Y. v. C. K. M. a. A. D. F. Stefan Koelsch, «Investigating Emotion With Music: An fMRI Study,» *Wiley-Liss,* n. 27, pp. 239-250, 2006.

[40] P. E. C. D. S. a. J. A. S. a. W. V. F. Richard J. Davidson, «Approach - Withdrawal and Cerebral Asymmetry: Emotional Expression and Brain Physiology I,» *Journal of Personality and Social Psychology,* vol. 58, n. 2, pp. 330-341, 1990.

[41] K. R. Scherer, «What are emotions? And how can they be measured?,» *Social Science Information,* vol. 44, n. 4, pp. 695-729, 2005.

[42] L. G. T. John Cacioppo, The Handbook of Psychophysiology, Cambridge: Cambridge University Press, 2007.

[43] A. Barreto, «Non-Intrusive Physiological Monitoring for Affective Sensing of Computer Users,» in *Human Computer Interaction: New Developments,* Kikuo Asai, 2008.

[44] a. B. N. N. A. R. D. Pierre Rainville, «Basic emotions are associated with distinct patterns of cardiorespiratory activity,» *International Journal of psychophophysiology,* n. 61, pp. 5-18, 2006.

[45] V. M. P. C. L. M. E. B. S. C. M. V. a. R. B. M. Mauri, «Psychophysiological Signals Associated with Affective States,» *NIH-PA Author Manuscript,* n. 3562-3566, 2012.

[46] R. W. L. W. V. F. Paul Ekman, «Autonomic Nervous System Activity Distinguishes Among Emotions,» *Science,* vol. 221, n. 4616, 1983.

[47] J. T. B. S. C. e. a. A. John Camm, «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,» *European Heart Journal,* vol. 17, pp. 354-381, 1996.

[48] S. F. Y. A. R. W. Picard, «Multiple Arousal Theory and Daily-Life Electrodermal Activity Asymmetry,» *Emotion Review,* vol. 8, n. 1, pp. 62-75, 2016.

[49] R. W. Levenson, «Blood, Sweat and Fears The Autonomic Architecture of Emotion,» pp. 348-366, 2003.

[50] M. D. R. Iris B. Mauss, «Measures of Emotion: A review,» *Psychology Press,* vol. 23, n. 2, pp. 209-237, 2009.

[51] F. B. R. a. C. M. K. Gouizi, «Emotion recognition from physiological signals,» *Journal of Medical Engineering & Technology,* n. 35, pp. 300-307, 2011.

[52] E. C. M. A. A. A. a. E. N. B. Riccardo Barbieri, «A point-process model of human heartbeat intervals: new definitions of heart rate and heart rate variability,» *Am J Physiol Heart Circ Physiol 288: H424–H435,* 2004.

[53] R. L. S. J. B. W. W. Kurt Kroenke, «The PHQ-9: Validity of a Brief Depression Severity Measure,» vol. 16, 2001.

[54] S. L. M. R. B. Benedek Kurdi, «Introducing the Open Affective Standardized Image Set (OASIS),» *Behav Res,* vol. 49, pp. 457-470, 2017.

[55] D. G. a. S. Scheibe, «Age-related differences in valence and arousal ratings of pictures from the International Affective Picture System (IAPS): Do ratings become more extreme with age?,» *Behavior Research Methods,* vol. 40, n. 2, pp. 512-521, 2008.

[56] S. S. Daniel Gruhn, «Age- related differencies in valence and arousal ratings of pictures from the International Affective Picture System (IAPS): Do ratings become more extreme with age?,» *Behaviour research Methods,* vol. 40, n. 2, pp. 512-521, 2008.

[57] J. P. a. W. J. Tompkins, «A Real-Time QRS Detection Algorithm,» *IEEE Transactions on Biomedical Engineering,* Vol. %1 di %2BME-32, n. 3, 1985.

[58] L. C. E. P. S. R. B. Gaetano Valenza, «Point-Process Nonlinear Models with Laguerre and Volterra Expansions: Instantaneous Assessment of Heartbeat Dynamics,» *IEEE Transactions on Signal Processing,* vol. 61, n. 11, 2013.

[59] K. P. J. N. K. C. M. L. a. J. S. S. U. Rajendra Acharya, «Heart Rate Variability: a

review,» in *Medical & Biological Engineering & Computing*, 2006, pp. 1031-1051.

[60] M. A. Stephens, «EDF statistics for goodness of fit and some comparisons,» *Journal of the American Statistical Association,* vol. 69, n. 347, pp. 730-737, 1974.

[61] J. W. T. a. W. A. L. Robert McGrill, «The American Statistician,» vol. 32, n. 1, pp. 12-16, 1978.

[62] F. F. a. D. L. V. Geoff Cumming, «Error bars in experimental biology presented by Sahana,» *The Journal of cell Biology,* 2015.