Dynamic Emotion Estimation based on Physiological Signals

by

Juhuan Ye

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the M.A.Sc degree in
Electrical and Computer Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering
School of Electrical Engineering and Computer Science
Faculty of Engineering
University of Ottawa

Abstract

Affective computing is becoming more and more popular, and the need to find a user-friendly and reliable method of estimating people's emotions, in their everyday life, is growing. Traditional methods have reached their limits, and this thesis presents a new system of emotion recognition, though physiological signals. With a user-friendly, wearable device, the system can be deployed in a number of fields. A model for our emotion classification is presented and includes the following emotions: cheerfulness, sadness, erotic, horror, and neutral. An experiment of emotion elicitation is also described in this work. Three analysis models applied in our system in order to recognize emotions, including nearest neighbor, discriminant analysis, and multilayer perception, are discussed in detail. The final test results show that the system has the average recognition rates of 40%, 55.7%, and 77.34% for nearest neighbor, discriminant analysis, and multilayer perception respectively.

Acknowledgement

I would like to express my sincerest gratitude to my supervisor Dr. Abdulmotaleb El Saddik for his continuous guidance and support. With his help and supervision, it was possible to achieve the present work and current studies successfully.

I would like to express special and sincere thanks to Ph.D Hussein Al Osman and Ph.D Haiwei Dong for the invaluable assistance, guidance, and feedback he provided throughout this research. I am thankful to Juan Sebastian Arteaga Falconi, who helped me in my research. Also, I would also like to thank all colleagues in DISCOVER and MCRLab for their contributions throughout the research.

Finally, I would like to express my deepest thanks to my parents. Without their inspiration, understanding and support, I can't finish this thesis during my study. This work is dedicated to them.

Contents

Abstract **iii**

Acknowledgement **iv**

1 Introduction **2**

1.1 Background 2

1.2 Problem Statement 3

1.3 Motivation 4

1.4 Objective and Contribution 5

1.5 Thesis Organization 6

2 Background and Related Works **8**

2.1 Emotions 8

2.1.1 Categories of Emotion 9

2.1.2 Emotion Elicitation 12

2.2 Affective Computing 13

2.3 Technologies in Recognizing Emotions 14

2.4 Physiological Emotion Detection 15

2.5 Conclusions 18

3 The Purposed Emotion Recognition System **21**

3.1 Overall System Architecture 21

3.1.1 Overview 21

3.1.2 Use Case Model 22

3.2 Cloud Server 24

4 Emotion Estimation **26**

4.1 Experiment Setting and Signal Acquisition 26

4.1.1 Participants 26

4.1.2 Materials 27

4.1.3 Methodology 30

4.2 Signal Preprocessing and Normalization 34

4.2.1 Subject's Emotion Confirmation 34

4.2.2 Data Reduction 36

4.3 Emotion Feature Extraction 36

4.4 Emotion Template Matching 38

4.4.1 Nearest Neighbor (NN) 40

4.4.2 Discriminant Analysis (DA) 40

4.4.3 Multilayer Perception (MLP) 41

5 Results and Discussion **45**

5.1 Overall Performance Comparison 45

5.2 Features Selection Comparison 47

5.3 Detailed Analysis in Discriminant Analysis 47

5.3.1 Output Results 49

5.4 Detailed Analysis in Nearest Neighbor . . . 55
5.4.1 Number of Nearest Neighbors . . . 55
5.4.2 Enter Feature Selection . . . 56
5.4.3 Training and Testing . . . 56
5.4.4 Output Results . . . 56
5.5 Detailed Analysis in Multilayer Perceptron . . . 57
5.5.1 Partition Dataset . . . 60
5.5.2 Model Option . . . 60
5.5.3 Methods for Training . . . 61
5.5.4 Output Results . . . 61

6 Conclusion and Future Work **66**
6.1 Conclusion . . . 66
6.2 Future Work . . . 67

A First Appendix **69**
A.1 The Questionnaire Form . . . 70

Bibliography **72**

List of Tables

2.1 Related work comparison 19

4.2 Selected Movie Clips (a)(Carvalho et al., 2012) 28

4.4 Selected Movie Clips (b) (Carvalho et al., 2012) 29

4.5 Raw Data from the Biopeak Fusion Device 34

4.6 The List of Emotion Features 39

5.1 The Classification Results of NN 45

5.2 The Classification Results of DA 46

5.3 The MLP Classification Result 46

5.4 Models' Features Selection 47

5.5 Eigenvalues 50

5.6 Wilks' Lambda Table 50

5.7 Function at group centroids 53

5.8 ROC area under the curve 64

List of Figures

2.1 Pulkkinen's Model of Emotional Regulation (Pulkkinen, 1995) 9

2.2 Plutchik's Model(Plutchik et al., 1997) 10

2.3 Russell's Model(Russell and Feldman Barrett, 1999) 11

2.4 MAUI-Context (Lisetti and Nasoz, 2002) 14

2.5 EEG electrodes(Khalili and Moradi, 2008) 17

3.1 The emotion recognition system architecture 22

3.2 Use case diagram of the local client module 23

3.3 Use case diagram of the cloud server module 24

3.4 Entity relationship diagram 25

4.1 Biopeak Fusion Physiology Status Monitor 30

4.2 The participant wearing the device 31

4.3 The Experiment Scene 32

4.4 Self-Assessment Manikin 33

4.5 The interface of Biopeak application 33

4.6 Emotion distinction model 35

4.7 MLP Model Architecture 43

5.1 MLP's Features Importance 48
5.2 Standardized Canonical Discriminant Function Coefficients 51
5.3 Structure Matrix 52
5.4 The territorial map 54
5.5 K Selection Error log 55
5.6 The Predictor Space 58
5.7 Peers Charts 59
5.8 Predicted Pseudo-probability 62
5.9 ROC curve 63
5.10 Gain-Percentage 64
5.11 Lift-Percentage 65

Glossary Of Terms

BMI	Body Mass Index
DFA	Discriminant Function Analysis
ECG	Electrocardiography
EDA	Electrodermal Activity
EMDC	Emotion-specific Multilevel Dischotomous Classification
GSR	Galvanic Skin Response
HF	High frequency of Heart Rate Variability
HRV	Heart Rate Variability
IAPS	International Affective Picture System
LDA	Linear Discriminant Analysis
LF	Low frequency of Heart Rate Variability
MBA	Marquardt Backpropagation Algorithm
MLP	Multilayer Perception
NN	Nearest Neighbor
NN50	Number of pairs of successive normal-to-normal intervals that differ by more than 50 ms
pLDA	Extended Linear Discriminant Analysis
PNN50	Proportion of NN50 divided by total number of normal-to-normal intervals
Resp.	Respiration
RMSSD	Root Mean Dquare of Duccessive Difference
SDNN	Standard Deviation of Normal-to-Normal intervals
SVM	Support Vector Machine
VLF	Vary low frequency of Heart Rate Variability

Chapter 1

Introduction

1.1 Background

Emotion recognition is a very hot topic in the field of affective computing, as in many computing fields. For example, an intelligent driving safety system could change the interior environment of a vehicle, or a user's driving style, by changing the music, according to the driver's emotion (James, 2000).

According to the studies of the last several years, the main affective technologies are using, among others, facial expression, speech processing, various body movements, and physiological signals in order to recognize emotion. Each affective computing technology has advantages and disadvantages. Compared to other methods of emotion recognition, physiological signals have considerable advantages. The bio-signals of human being are not easy to fake, which means the emotions, if properly detected, are closer to the true feelings. Since wearable digital devices are increasingly powerful, the methods used to detect physiological signals are more convenient and user-friendly than ever before.

Some physiology research shows there is a strong relationship between a humans being's emotional state and physiological reactions. For example, some emotions such as happiness, sadness, fear or anxiety may cause a change in breathing (Homma and Masaoka, 2008). Heart rate variability (HRV) is also associated with emotion recognition and some recent studies show the HRV may have the ability to recognize emotion even with a variety of confounding variables such as body mass index (BMI), sex, depression, smoking habits, stress, anxiety, and physical activity levels (Quintana et al., 2012).

Over the past decade, many researchers have done experiments to determine the relationship between human being's emotions and bio-signals. Recently, studies show the classification analysis methods for the specific physiological emotion. The results clearly indicate that bio-signals could be utilized for emotion recognition(Wioleta, 2013).

1.2 Problem Statement

We would like to explore various methods in emotion recognition in order to achieve a higher level of affective computing. To do this, we will take advantage of the mobile wearable physiological device that is small and comfortable to wear; it will make the process of detecting bio-signals more user-friendly. Although many existing systems have achieved good results in emotion recognition, we find that there are still existing problems, as summarized in the following points:

- Limitedness of hardware: Most existing systems using a device can only be deployed in the lab, which means the devices are cumbersome and the process of signal acquisition is complex.

- Limited number of emotions: In some systems, the scale of target emotions is narrow and dependent. In our study, the scale of emotions covers most common feelings, and is independent.

- Limitedness of the experiment of emotion elicitation: Most experiments previously conducted have had few participants and few records and used methods of emotion elicitation that could still be greatly improved.

1.3 Motivation

Social networking has become a very important part of our lives. Texts, pictures, audio and video can be shared through social networks, however, corresponding emotions cannot. If multimedia could be recorded with real emotion, it would definitely help people recall the specific scenarios. Facial expressions, posture, tone of voice, and gestures are always used for emotion recognition.

However, to collect emotional data in people's normal life is not an easy task. Many researchers are working on how to use physiological signals to identify people's emotion. Bio-signals are strongly correlated to human emotion. As mobile devices become more and more powerful, it's increasingly possible to collect people's bio-signals in real-time, and share their emotion with pictures, texts or even videos, in social networks such as Facebook and Twitter.

Otherwise, the emotion recognition could improve the human-computer interaction. For example, Siri is an intelligent personal assistant application which works in Apple mobile devices. It could understands the commands which people say. If the application could recieve the commands with people's emotion, it could adapt to the user's individual preferences and personalizes results.

For this purpose, we are investigating this issue. Our work proposes a different experiment, where emotional film clips are used for emotion elicitation. We also design the methods for signals prepossessing and normalization, emotion feature extraction, and emotion recognition.

1.4 Objective and Contribution

In this thesis, we will introduce an experiment for emotion elicitation and three analysis models for emotion classification. Conducting a reasonable experiment for emotion elicitation and data collection. In our design, we obtained hundreds of records of physiological signals for each emotion, and we improved the emotion elicitation by using emotion movie database (EMDB) and self-assessment manikin (SAM). The main contributions of this research work is summarized in the following points:

- Using the physiological signals that can be easily detected by a small wearable sensor.We designed and implemented an algorithm to select and extract the important features from the limited physiological signals. The advantage of this is that our devices could be deployed in a user's normal life and will therefore make our system more practical.

- Building three different classification models for predicting emotions that people often feel in their normal life. In our study, the scale of the emotions covers most of the common ones, and they are all independent. To enhance the accuracy of classification, we designed and developed three classification models by using NN, DA, and MLP respectively.

1.5 Thesis Organization

The remainder of this thesis is organized as follow:

Chapter 2 introduces the basic definition of emotion in psychology and the main categories of emotions. Then, it presents the emotion elicitation technologies. It also describes affective computing and the technologies used in emotion recognition. After that, it shows the most relevant research in physiological emotion detection. In the end, this chapter provides a comparison table to discuss the important differences between these studies. Finally, a summary of this chapter is given along with contribution points provided by our research.

Chapter 3 shows the design of proposed emotion estimation system. At first, the framework of the system is presented. Then it describes a high level overview of the system by using use case models. At last, it introduces the four componts of cloud server in detail.

Chapter 4 explains the proposed emotion estimation system in detail. It starts with a description of the experiment setting and signals acquisition. It then provides a brief overview of the signal prepossessing and normalization, followed by a description of the emotion feature extraction. After that, it gives a detailed description of the three classification algorithms: nearest neighbor, discriminant analysis, and multilayer perception.

Chapter 5 provides the results of the experiments as well as a discussion on the predicted emotions. It starts by showing a performance comparison of the three models. The results of feature selections are then shown, as well as a detailed analysis of the three algorithms. The discriminant analysis section shows the details of the stepwise methods and the final results of the emotion classification. The nearest neighbor section describes the detailed process and its results. Finally, the work of multilayer perception is described in aspects of data set partition, model option, training methods, and output diagrams.

Chapter 6 summarizes the work done in this research, provides a conclusion , and proposes future work that can enhence the present work.

Chapter 2

Background and Related Works

2.1 Emotions

To help understand our emotion recognition system, we will first describe some important concepts.

Emotion is a complex state of feeling for a human being, for example sorrow, fear, joy, hatefulness. In physiology, emotion is often related to the arousal of the nervous system and may be accompanied by physiological changes such as increased respiration or heartbeat. The main difference between mood and emotion is the duration of the feeling. A good or bad mood may last for one or two days, but an emotion may just last for a few seconds or minutes.

According to popular research on emotions, the causes of emotions could be grouped into three main categories: physiological, neurological and cognitive. The Cannon Bard theory of emotions claims that emotions are a result of physiological reactions to events, where the thalamus sends a message to the brain(Cannon, 1927).For the cognitive theories,

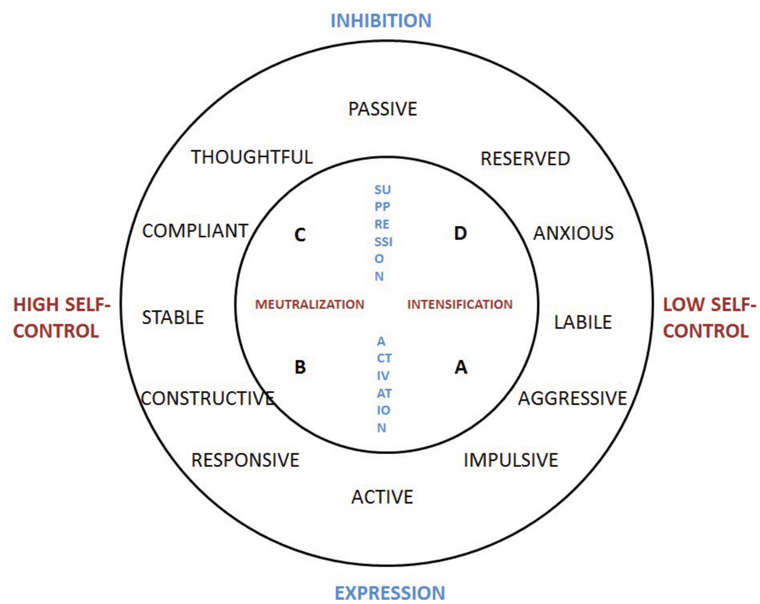


Figure 2.1: Pulkkinen's Model of Emotional Regulation (Pulkkinen, 1995)

the well-known Schachter-Singer theory is a good example that suggests that the physiological arousal happens, then the person recognizes the reason for the arousal and considers it as an emotion (Schachter and Singer, 1962).

2.1.1 Categories of Emotion

Many researchers have attempted to divide the basic emotions. A good example of this is Paul Ekman's six basic emotions: happiness, sadness, anger, disgust, fear, and surprise (Ekman, 1992). He explains that the basic emotions allow particular characteristics to be expressed in different degrees. For practical reasons, many theorists define emotional models according to various dimensions. In our study, we examined three of the most influential classification models of theoretical emotional framework: Pulkkinen's, Plutchik's, and Russell's.

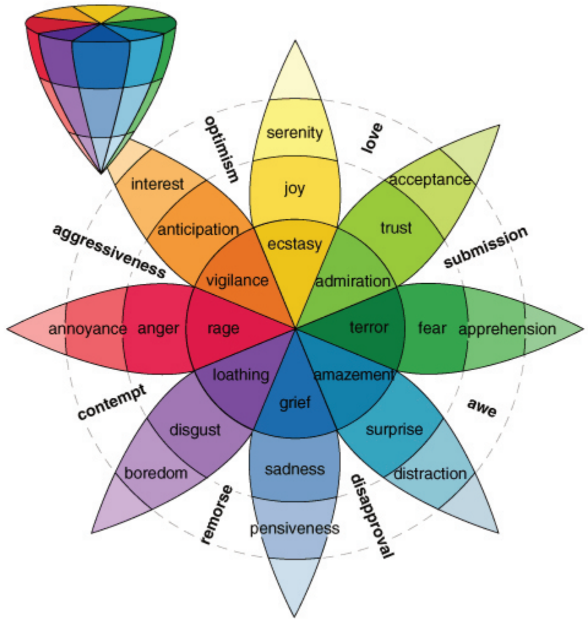


Figure 2.2: Plutchik's Model(Plutchik et al., 1997)

Pulkkinen's model of emotional regulation provides a map of emotions, which are divided into two dimensions: inhibition vs expression and high self-control vs low self-control. In Figure 2.1, we can see that the model identifies 11 emotions: passive, reserved, anxious, labile, aggressive, impulsive, active, responsive, constructive, stable, compliant and thoughtful. Pulkkinen also divides the emotions into four behavioural groups: A, B, C and D. These four patterns are used for behavioural regulation.

The theoretical framework of Plutchik's study shows a three-dimensional model containing a circle and the vertical dimension of a cone, as in Figure 2.2. In his study, he defines the eight primary emotions as: joy, trust, fear, surprise, sadness, disgust, anger and anticipation. The eight primary emotions are arranged as four pairs of opposites: joy-sadness, trust-disgust, fear-anger, and surprise-anticipation. In the model, each emotion has a different color. As we can see, the intensity of the emotions is represented by the cone's

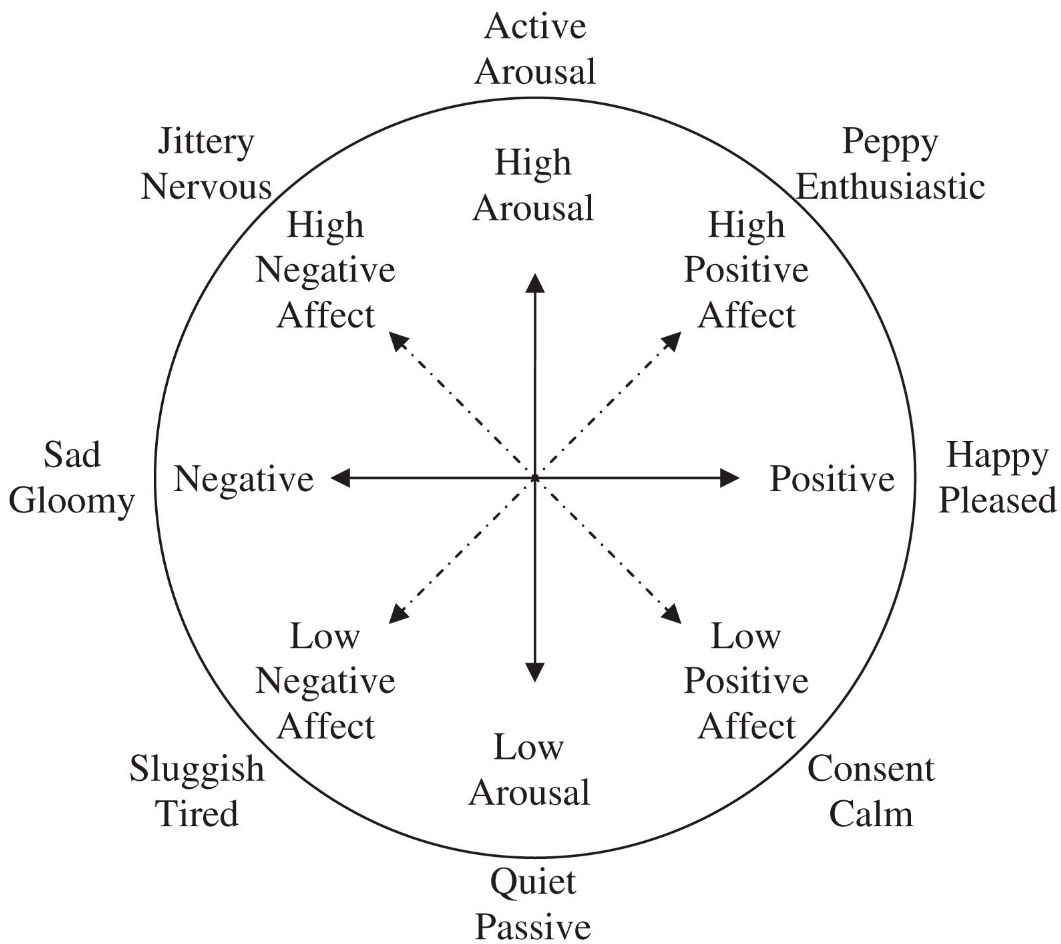


Figure 2.3: Russell's Model(Russell and Feldman Barrett, 1999)

vertical dimension, and the circle indicates the degrees of similarity among the emotions. So each emotion can be identified by the varying degrees and levels of arousal. Plutchik's model has different versions for different tasks, such as affective computing in computer science (Cambria et al., 2012).

For Russell's model, depicted in Figure 2.3,we can see that the emotions are divided by arousal and valence dimensions. The vertical axis represents arousal and the horizontal axis means valence. In this model, each emotion can be identified by its varying degree of arousal and valence, and the eight primary emotions are defined as: active/arousal, peppy/enthusiastic, happy/pleased, consent/calm, quiet/passive, sluggish/tired, sad/gloomy and jittery/nervous (Russell and Feldman Barrett, 1999). The centre of the circle means a neutral state. This model has been widely used by emotion classification tests and emotional facial expression recognition (Remington et al., 2000).

2.1.2 Emotion Elicitation

Before we identify the emotion of the subject, we must find a method of eliciting emotional responses in the laboratory. In this part, we discuss the reliability of various emotion elicitation methods, according to recent research.

The International Affective Picture System (IAPS) is a very popular system used for emotion elicitation (Lang et al., 1997). The IAPS consists of 1,000 colored pictures. IAPS pictures have been used in experiments studying emotion and attention with other physiological and neurological measurements. The system also provides a set of standardized stimuli for the elicitation of emotions.

The Emotional Movie Database (EMDB) is a film clips system for emotion elicitation (Carvalho et al., 2012). In this study, they developed a film database with 52 film clips based on the emotional stimuli (valence, arousal and dominance). After the pre-selection of these movies, they asked 113 participants to rate each one. They also conducted a psychophysiological assessment on 32 subjects in order to test these film clips to ensure that they could be used in the emotional study.

It seems film clips are more effective for emotion elicitation, since videos contain more emotional content than a single picture. In our opinion, if the film clips of the EMDB are used with audio, the effect of the experiment could be even better.

2.2 Affective Computing

Affective computing is a hot research about recognizing, interpreting, processing and simulating human affects (Tao, 2005). A number of studies have examined the relevance between human emotions and technology. The research projects on affective computing can be divided in two main categories: detecting and recognizing emotional information, and emotion in machines.

To recognize emotions, a number of computational sensors are used to capture the user's actions, for example body movement, gestures, speech, facial expressions, and many physiological signals. After gathering the user's data, the standard procedure of affective computing is to model the data and identify the emotions (Tao, 2005).

Another affective computing field of study is to build a device or system with emotional capabilities. In Heise's study, he provides a practical approach to insert the simulation of emotions in the interactions between human beings and machines, in order to improve their interactivity (Heise, 2004).

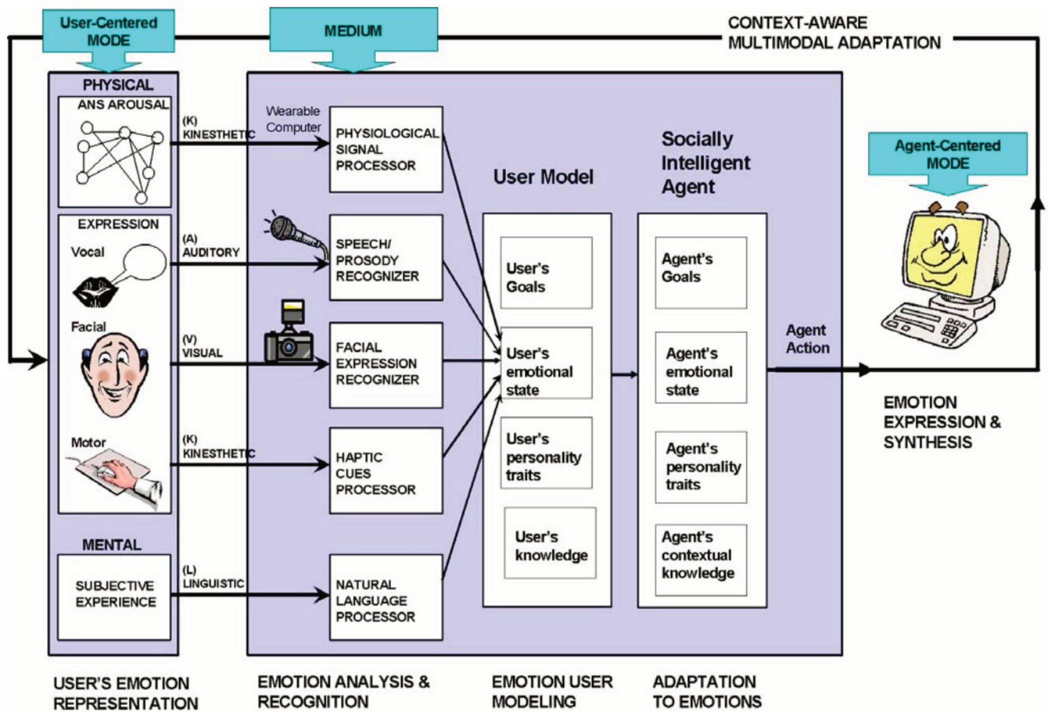


Figure 2.4: MAUI-Context (Lisetti and Nasoz, 2002)

Affective computing could provide many benefits in different fields. For example, affective computing could be used in education. When the system detects the learner is bored or pleased, the system could adjust the style of teaching (Asteriadis et al., 2009).

Since our study of affective computing belongs to the area of emotion recognition, we describe the related work in this area in greater detail.

2.3 Technologies in Recognizing Emotions

Figure 2.4 shows the overall diagram of current technologies that recognize emotions, and was developed by Lisetti (Lisetti and Nasoz, 2002). As we see in the diagram, this framework divides the user's emotional manifestations into two parts: physical and mental. The physical part includes the ANS (autonomic nervous system) arousal, the human expressions, and mental experience. In this framework, the ANS arousal belongs to physiological emotion detection and the human expressions include vocal, facial and motor information. Lisetti explained that the mental aspect in this model refers to people's feeling to a subjective experience, in a specific scene.

The second part of this framework is named Medium, and introduces the emotion analysis and recognition, as well as emotion user modeling and adaptation to emotions. For the first part, the diagram shows the basic devices used for emotion analysis. For example, the system needs wearable computers to process physiological signals.

After gathering the necessary data concerning people's emotions, they built a user model that includes the user's goals, emotional state, personality traits and knowledge.

The socially intelligent agent in the adaptation to emotions section is built according to the user model. This agent could adapt to a user's current emotional state according to the user's personality or knowledge.

This framework includes most of the main fields of affective computing: emotional speech processing, facial expression, body gestures and physiological signals. Below we introduce the related work of emotion recognition by physiological signals in detail.

2.4 Physiological Emotion Detection

Over the past decade, many researchers have conducted experiments to improve the accuracy of emotion identification. Recently, researchers have investigated classification analysis methods for specific physiological emotions.

Wearable Computers to Recognize Emotions from Biosignals The relation between affect, cognition, and the importance of emotions is discussed in detail (Lisetti and Nasoz, 2004). They developed an autonomic nervous system (including heart rate, temperature, galvanic skin response), which can recognize certain emotions. For the emotion elicitation, Lisetti and Nasoz used five movie clips (The Champ, Schindler's List, Drop Dead Fred, The Shining and Capricorn One) to elicit five specific emotions respectively: sadness, anger, amusement, fear and surprise.

Wearable physiological sensors were used to gather the physiological data including heart rate, skin temperature, and GSR. Lisetti and Nasoz designed three different supervised learning algorithms that categorize the mentioned collected data in terms of emotion. Overall, the three algorithms, k-nearest neighbors (KNN), discriminant function analysis (DFA), and marquardt backpropagation algorithm (MBP), could categorize the specific emotions with an accuracy of 72.3%, 75.0%, and 84.1%.

Using Brain and Peripheral Signals for Emotion Detection The study of Khalili proposed an emotion recognition system dependent on electroencephalographic brain signals (EEG) and other physiological signals(Khalili and Moradi, 2008). They used a set of international affective picture system (IAPS) pictures to elicit specific emotions (positively excited, negatively excited, and calm) from five participants. The device used consists of 64 electrodes, as shown in Figure , which are used to detect EEG signals. To extract the features, KNN and linear discriminant analysis (LDA) are used as classifiers. The final results show a classification accuracy of 51% for the three categories.

An Affective Computing Approach to Physiological Specificity According to the theory of Kolodyazhniy's on affective computing approaches (Kolodyazhniy et al., 2011), a

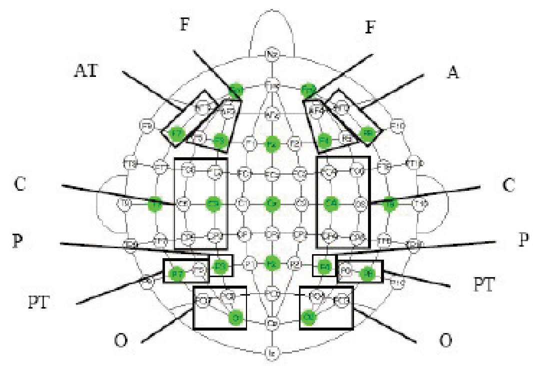


Figure 2.5: EEG electrodes(Khalili and Moradi, 2008)

pattern of classification analysis (PCA) was designed for affective computing. The PCA includes feature selection, classifier type, and cross-validation. For the emotion elicitation, 34 participants watched six 10 min. film clips to elicit fear, sadness, and a neutral emotional state. The physiological sensor collected 14 features from the subjects. In their study, they used the sequential forward selection and sequential backward selection methods for feature selection. The five common features, which are suitable for all models in their study, are: PEP (preejection period), SRR (rate of skin conductance fluctuations), pCO2 (end-tidal carbon dioxide partial pressure), CS (musculus corrugator supercili), and ZM (musculus zygomatic major).

The methods used for classifier type are linear and quadratic discriminant analysis, neural networks, and k-nearest neighbors methods. A cross-validation method (subject and stimulus (in)dependence) is also used to classify emotions. The target emotions in this study are fear, sadness, and neutral. The correct classification rate for sadness is 81%. For fear, the highest classification accuracy is achieved with stimulus-independent cross-validation (88%). The results show that emotion classification can help identify the three specific emotions.

The Emotional Movie Database Carvalho built an emotional movie database that includes 52 selected and edited film clips(Carvalho et al., 2012). These film clips are divided into 6 categories: erotic, horror, socially negative, socially positive, scenery and objects. The participants were asked to provide feedback on their feelings after watching the film clips. Psycho-physiological responses such as heart rate (HR) and skin conductance level (SCL) were also collected. In this study, HR and SCL show a better classification of high and low arousal than of positive and negative valence.

2.5 Conclusions

Based on all the different works studied on physiological emotion detection, it can be concluded that bio-signals could be used for affective computing. We compared the most relevant work, which is shown in Table 2.1. In the table, some methods give better results than others, according to the emotion classification accuracy. However, there are no conclusive results indicating which method is better. Since they tried to recognize different emotions based on different classification methods and different bio-signals, the results will vary depending on the application.

Contrary to existing research, we will design our system to possess the following properties:

- Using the physiological signals that can be easily detected by small wearable sensors. This means that our devices could be deployed in a user's normal life and will therefore make our system more practical. Most existing systems using a device can only be deployed in the lab, such as the EEG electrodes.

Predicted Emotions	Physiological Signal	Emotion Elicitation	Subjects	Devices	Methods of Classification	Results %
Sadness, fear, anger, frustration, surprise, amusement Lisetti and Nasoz (2004)	HR, ST, GSR	Film clips	29	BodyMedia SenseWear Armband	KNN, DFA, MBA	72.3(KNN), 75(DFA), 84.1(MBA)
Clam, positively excited, negatively excited Khalili and Moradi (2008)	EEG, GSR, RSP, ST, BVP	IAPS	5	-	KNN, LDA	50-70
Four pairs musical emotions(positive, negative, high and low arousal) Kim and Andre (2008)	ECG, EMG, RSP, SC	Musical induction	3	ProComp Infiniti	pLDA, EMDC	70(Subject-independent), 95(Subject-dependent)
Anger, Fear, Disgust, Sadness, Neutral, Joy Rattanyu et al. (2010)	ECG	IAPS	6	Wireless Bio Sensor RF-ECG	ANOVA, LDA	61.79
High Stress, Low Stress, Disappointment and Euphoria Katsis et al. (2008)	EDA	Film clips	12		Support Vector Machines (SVM), Adaptive Neuro-Fuzzy Inference System (ANFIS)	79.3 (SVM), 76.7 (ANFIS)

Table 2.1: Related work comparison

- Predicting the most common emotions that people often feel in their normal everyday life. This is especially important since some target emotions of existing studies are narrow and dependent, such as the studies of Khalili and Moradi (2008),Kim and Andre (2008), and Katsis et al. (2008). In our study, the scale of the emotions has covered most of the common ones, and they are all independent.

- Conducting a reasonable experiment for emotion elicitation and data collection. As we see in the Table 2.1 , some experiments have few subjects for data gathering and the methods of eliciting emotions have many aspects that need to be improved. Therefore, we gathered hundreds of records of physiological signals for each emotion, and we improved the emotion elicitation by using EMDB and SAM

- Building three different classification models to predict emotions. According to our research, we created three classification models by using NN, DA, and MLP respectively. We compared the results of the three methods to obtain the best classification accuracy.

Therefore, in this thesis, our work will help contribute to the development of affective computing using bio-signals

Chapter 3

The Purposed Emotion Recognition System

In this Chapter, the overall architecture of the proposed system is described in detail. At first, the framework of the system is presented. Then we describe a high level overview of the system by using use case models. At last, we introduce the four components of cloud server in detail.

3.1 Overall System Architecture

3.1.1 Overview

The proposed system manages the storage, transportation and display of the measured and calculated physiological data for the emotion recognition. Figure 3.1 indicates the architecture of the whole system which consists of four main parts: input sensor, local clients, cloud server, and social network. The input sensor is a device which could detect user's physiological signals and send the raw data to the local client.

The local client application could establish a connection and communication with the input sensor via a Bluetooth channel. Then it decodes the received data package and shows the data in real time. Moreover, the application relays collected information to the cloud server.

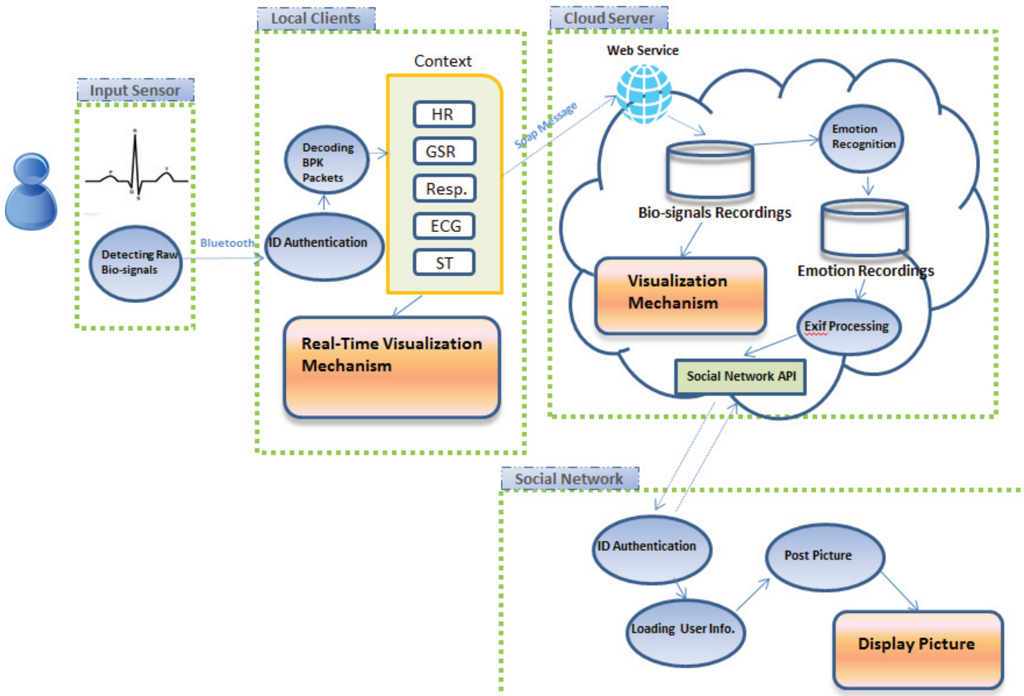


Figure 3.1: The emotion recognition system architecture

The cloud server collects all the information received from the local client module and stores them in a database. It also provides a web-interface so that stored information can be retrieved and viewed. Also, it supports to associate a dataset with the corresponding image file that was created during the same recording. The exchangeable image file format (Exif) usually includes the recorded time. The function of emotion recognition will be discussed in detail in Chapter 5.

The processed image with emotion could be shared in social network such as Facebook.

3.1.2 Use Case Model

The use case models of the local client and the cloud server show a high level overview of the system. They indicate the effective interaction between the user and the system. Based on the system design, we have defined the following actors:

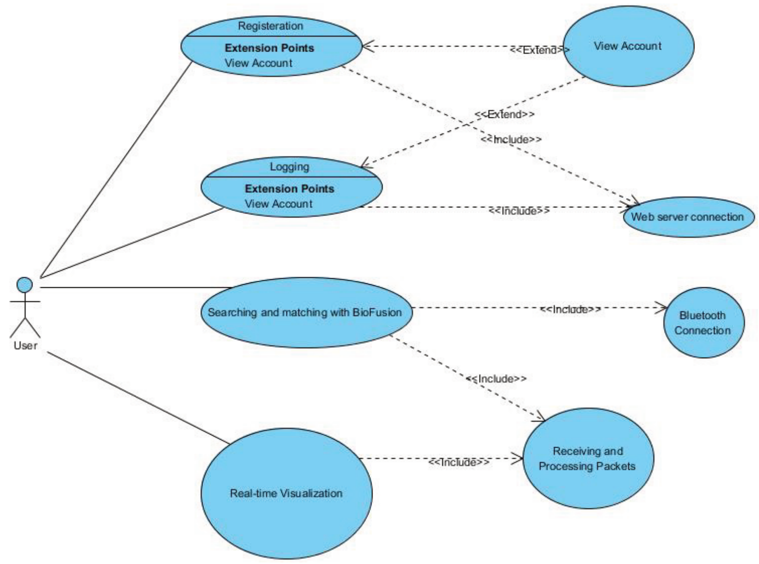


Figure 3.2: Use case diagram of the local client module

USER: The User is the person who administers the physiological measurement on the subject (note that the subject can be herself/himself). The user needs an account to record the acquired data on the cloud server. She/he has to login through the local client module in order to initiate the data communication with the cloud server.

Administrator: An administrator performs administrative tasks e.g., updating software to the server, auditing the server data, etc.

Figure 3.2 and Figure 3.3 illustrate the use-case diagrams to summarize the core requirements of the prototype system. We have discussed the following main Use Case of the models.

Registration and logging: Before using the emotion recognition system, users have to register an account and login the system. Each user has a unique id for account control. Both of the two modules have the functions.

Searching and matching device: The input sensor needs to be found before recording starts. The bluetooth connection should be established.

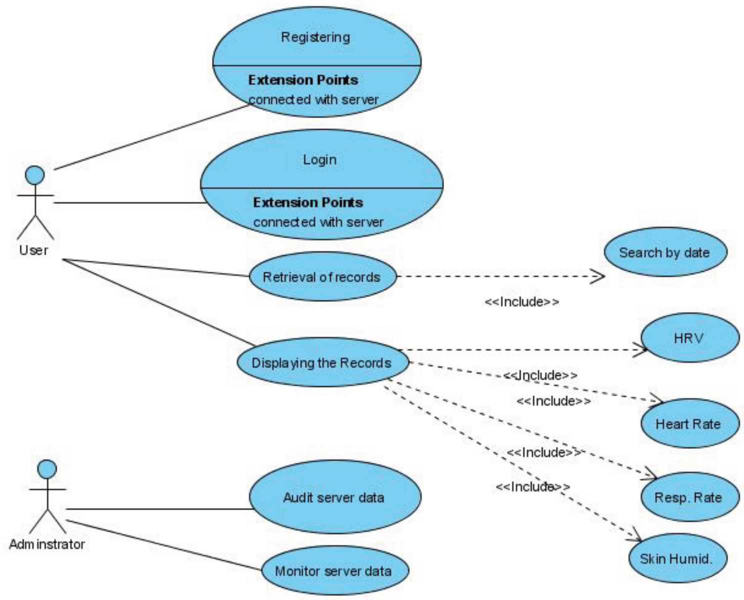


Figure 3.3: Use case diagram of the cloud server module

Display real-time visualization: For each signal that is associated with an subject (user) can generate a real-time chart on the local client.

Retrieve of records and display: User could retrieve physiological records from the database. According to the data, the cloud server would plot the records at various levels of granularity.

3.2 Cloud Server

The cloud server will be composed of four parts: persistence unit, web archive, enterprise archive, and parent project. Persistence unit is an Enterprise JavaBeans (EJB) 3 Java ARchive (JAR) file containing the entities for the database connection which shown in Figure 3.4. As we know, EJB is a popular architecture for modular construction of enterprise application.

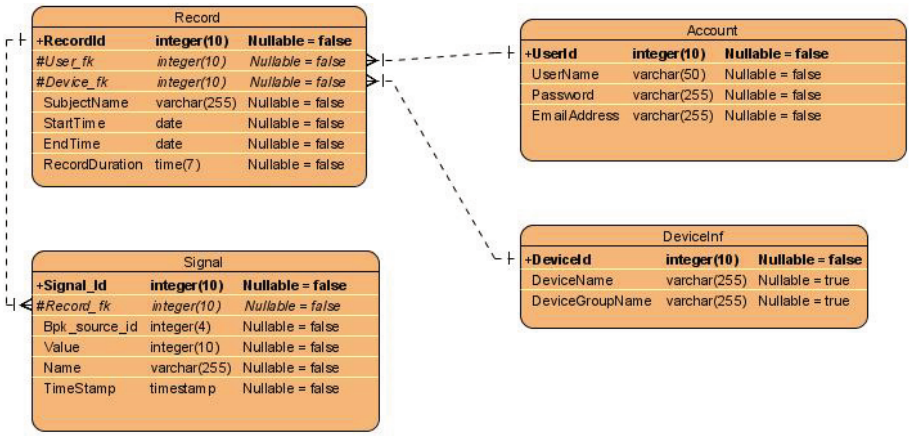


Figure 3.4: Entity relationship diagram

Web archive is a web application archive (WAR) file which contains the display logic for the application. WAR is used to distribute a series of resources such as java servlets, java classes, and XML files.

Enterprise archive (EAR) contains the WAR and the EJB3 JAR. EAR uses a deployment descriptors XML file to describe how to deploy the WAR and EJB3 JAR modules.

In our system, the parent project is a Maven root project which enables the build of the above three in the correct order. Maven is a tool which primarily used for building Java projects.

Chapter 4

Emotion Estimation

This chapter illustrates the overall design of the proposed emotion estimation system, which can detect bio-signals from the user, process the record data, and identify the emotion the user is feeling at the moment. First, the design of our experiment for emotion elicitation is presented in detail. Then, the methods used for signal prepossessing and normalization are discussed, followed by a description of the emotion feature extraction. Furthermore, the mechanism of three algorithms are introduced namely: nearest neighbor, discriminant analysis, and multilayer perception.

4.1 Experiment Setting and Signal Acquisition

4.1.1 Participants

We randomly recruited people based on these conditions:

i) Male and female

ii) Ages 18-36

iii) Not presenting any medical contra indication for athletic participation according to their answers on the Self-Assessment Manikin document (SAM).

iv) Healthy people without heart disease, hypertension or other diseases that may be caused by emotional flooding.

v) For pragmatic reasons, we recruited only people living or working in our region, since subjects needed to conduct the experiment in our lab room.

vi) Because of limited time, we recruited 20 participants based on a first come first serve basis.

We double checked the participants' health conditions and made sure that they were not at risk. The subjects could stop their experiment at anytime, for example after 5min.

In case some subjects felt uncomfortable, we have an area within the lab to rest and refresh by having water and soft drinks and a university clinic is located only about 200m away, so in an emergency

we could easily consult a physician. We also provided a list of resources and contacts for participants who need psychological help after the activity.

4.1.2 Materials

After reviewing an analyzing a number of methods of emotion elicitation, we conducted our own experiment to obtain the specific emotions related to each physiological signal. For this purpose, we used a questionnaire and the film clips chosen from EMDB (Sandra et al. 2012) to elicit the following emotions: cheer, horror, erotic, sad and neutral. The movie clips are shown in Table 4.2. .

For psychophysiological recording, Biopeak Fusion Physiology Status Monitor (BF-PSM) was used to capture user's physiological signals. The BFPSM is a compact wearable

Type	No.	Title	Clip Description
Neutral	1001	Disney's Earth	Desert and polar scenes
	1002	Disney's Earth	Mountains with ice
	1003	Disney's Earth	Scenery with polar scenes and the dusk
	1004	Disney's Earth	Waterfalls
	1005	Disney's Earth	Flowers and trees
	1006	Disney's Earth	Sandstorm and desert
	1007	Disney's Earth	Several takes of trees
	1008	Disney's Earth	A scenery overview including trees, waterfalls and sand
	1009	Disney's Earth	Several scenes from a jungle and in the end mushroom start to grow
	1010	Disney's Earth	Clouds swirling
Sad	2001	Boogeyman	Barry Watson entering a church for his mom's funeral
	2002	The Descent	Girl crying while her beloved died in her arms
	2003	The Pianist	Emilia Fox brings a cloth to Adrien Brody, who is lying sick on the bed
	2004	Diary of a Nymphomaniac	Very sad Belen Fabra considering suicide
	2005	Mystic River	Marcia Gay Harden and Tim Robbins having an argument
	2006	Boogeyman 2	Danielle Savre crying in the arms of Matt Cohen
	2007	Bridge to Terabithia	Josh Hutcherson crying in the arms of Robert Patrick
	2008	American Beauty	Annete Benning and Thora Birch having an argument, in Thora's room
	2009	American Beauty	Wes Bentley and Chris Cooper having an argument
	2010	Mystic River	Sean Penn crying in the front porch
Cheer	3001	This Girls Life	Kip Kardue having dinner with Juliette Marquis
	3002	My Best Friend's Girls	Dane Cook and Kate Hudson funny dance in the Promenade party
	3003	Good luck Chuck	Dane Cook and Jessica Alba walking and talking at night
	3004	Ruins	Night scene at the beach, with happy people partying
	3005	Lie With Me	Lauren Lee Smith and Eric Dalfour riding a bike and then in a discotheque
	3006	Last Chance Harvey	Grooms dancing at the wedding reception
	3007	Ruins	Scene in the pool, where they are deciding if they will go to the ruins
	3008	Diary of a Nymphomaniac	Belen Favra, happy entering her new home for the first time
	3009	Diary of a Nymphomaniac	Happy Belen Favra remembering romantic scenes while talking to her friend
	3010	The Rest Stop	Couple on top of a car with violet flowers surrounding them

Table 4.2: Selected Movie Clips (a)(Carvalho et al., 2012)

Type	No.	Title	Clip Description
Erotic	4001	Underworld: Evolution	Sex scene between Kate Beckinsale and Scott Speedman
	4002	Playboy's Clip	Couple having sex: woman in astride position and with rear entry while the man is standing
	4003	9 Songs	Margo Stilley and Kieran O'Brien having sex in the living room. She is sitting on the sofa while he is standing
	4004	Killing Me softly	Joseph Fiennes and Heather Graham having bondage sex near the fireplace
	4005	Kama Sutra: the sensual art of lovemaking	Couple having sex in arch position
	4006	Kama Sutra: the sensual art of lovemaking	Couple having sex in Variant yawning, namely Fixing nail
	4007	9 Songs	Margo Stilley and Kieran O'Brien having oral sex
	4008	Monamour	Anna Jimskaia wearing a black silk dress and having oral sex and intercourse with Riccardo Marino
	4009	Diary of a Nymphomaniac	Belen Farvra having sex in the missionary position
	4010	Diary of a Nymphomaniac	Belen Favra having sex on the desk chair
Horror	5001	The ruins	Amputation scene on top of the ruins
	5002	Texas Chainsaw Massacre: The Beginning	Leatherface removing the face of Mathew Bomer
	5003	Midnight Meat Train	Vinnie Jones removing the eyes and the teeth of the victim
	5004	Hostel	Jay Hernandez is being tortured on a chair and fingers from his hand are amputated
	5005	Hostel 2	Cannibalism scene
	5006	Midnight Meat Train	Leslie Bibb inside a carriage with bodies hanging from the ceiling
	5007	Canibal Holocaust	Savage attack from a cannibal tribe on an anthropologist, dismembering him, children start to eat parts of him
	5008	Texas Chainsaw Massacre: The Beginning	Jordana Brewster very scared, hidden on a box assists to a mutilation of her boyfriend
	5009	The Rest Stop	Jaimie Alexander giving a merciful shot to the head of a police officer
	5010	Midnight Mear Train	Vinnie Jones with a vicious attack direct on a woman, which ends with her decapitation

Table 4.4: Selected Movie Clips (b) (Carvalho et al., 2012)

Figure 4.1: Biopeak Fusion Physiology Status Monitor

non-invasive sensor platform that delivers unparalleled performance in monitoring, collecting and reporting critical information on human performance and health parameters. The monitor is shown in Figure 4.1. We also use a laptop to run the application and get the signal data from the monitor via bluetooth. Participants are asked to wear the device as they watched the emotional film clips, as shown in Figure 4.2. The following subsections discuss the signal prepossessing and normalization in greater detail.

Our experiment received ethical approval by the Research Ethics Board of the University of Ottawa. All participants were informed about the content and potential risks before beginning the experiment.

4.1.3 Methodology

Before the experiment began, all the participants were told that they would see several strong emotion film clips. They were also informed that they might be shocked by some film clips and that they were at liberty of quitting at any time. The participants were given instructions for the Self-assessment Manikin (Lang 1980), as shown in Figure 4.4 : rating

Figure 4.2: The participant wearing the device

based on a 9-point Likert scale for valence, arousal and dominance. The Before watching the film clips, the participant was required to wear the Biopeak device. The Figure4.3 shows the participant wearing the physiological device and watching one of the film clips (it is not necessary to be naked during the experiment). Immediately after viewing the clip, the participant completes the questionnaire then waits for few minutes before watching the next clip. The experimental session takes approximately 30 minutes, during which each participant watches 5 random film clips from each category

During each session, we also recorded the participant's physiological signals, including the data of ECG, respiration, CO deflection, thoracic impedance and ECG quality, with the Biopeak device, through bluetooth. The interface of the recording application is shown in Figure 4.5.

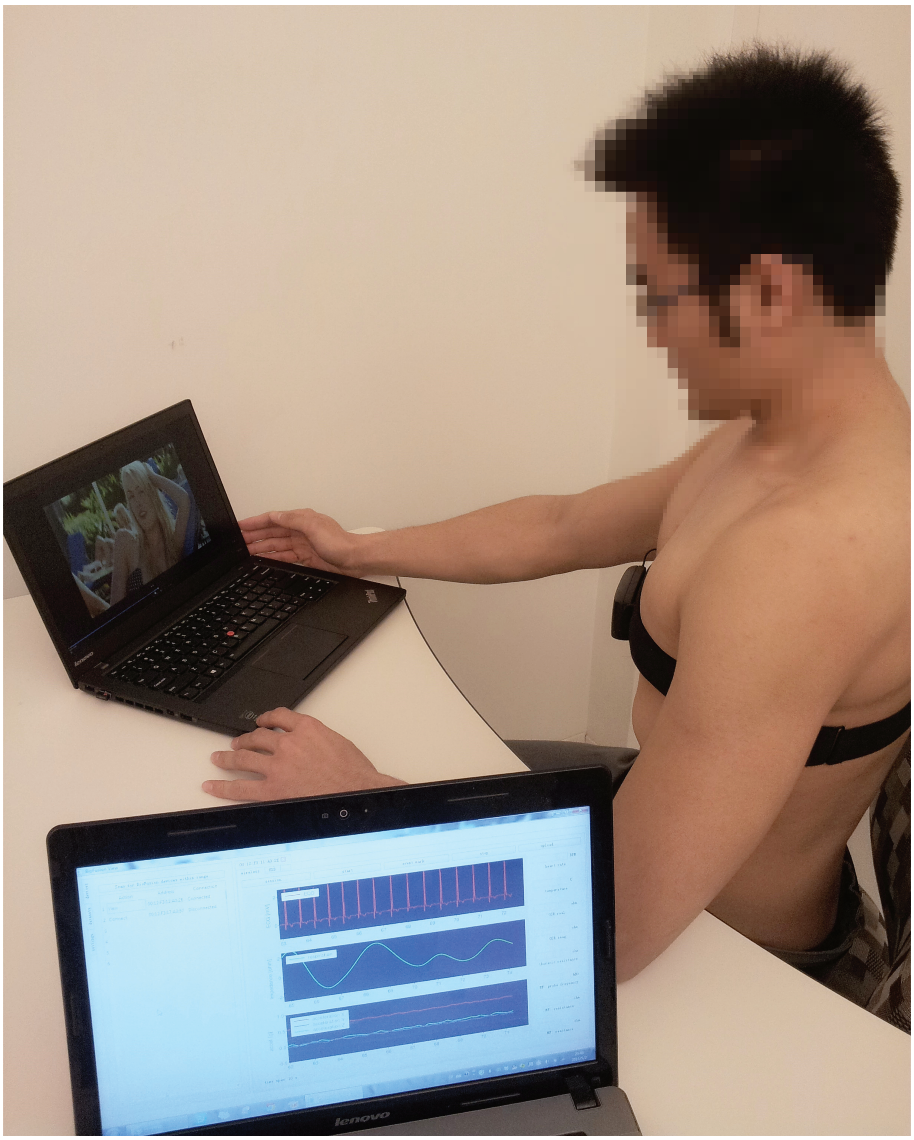


Figure 4.3: The Experiment Scene

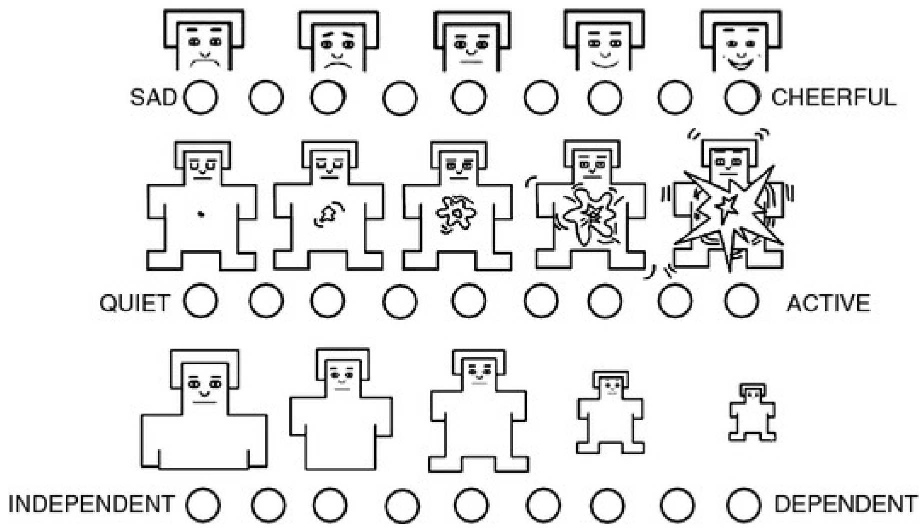


Figure 4.4: Self-Assessment Manikin

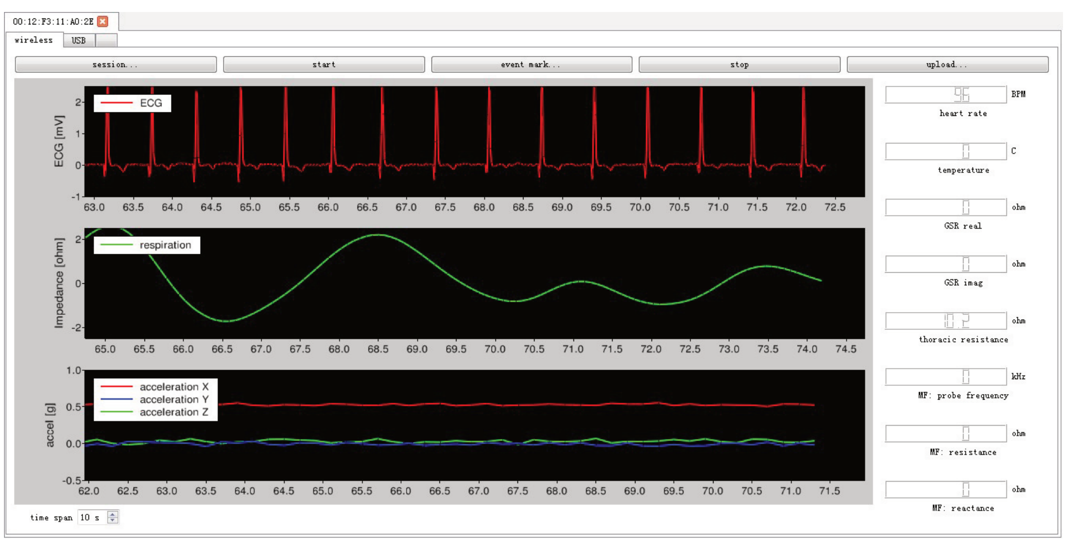


Figure 4.5: The interface of Biopeak application

Label	Physical Units	DigMax	DigMin	PhysMax	PhysMin	SamplingRate
ECG	mV	60000	0	6	-6	300
Respiration	Ohm	65535	0	5	-5	50
CO deflection	Ohm	65535	0	1.5	-1.5	50
Thoracic Impedence	Ohm	65535	0	118.433	-118.433	1

Table 4.5: Raw Data from the Biopeak Fusion Device

Server time tags are recorded according to the reaction of participants, and are used for data reduction. There are 8 files used to store signals, and 1 general data format (GDF) file in which all the information is kept. The GDF file includes file name, label, file type, physical unit, digital max value, digital min value, sampling rate, number of samples, and etc. These 8 files contain signal measurement values in decimals. The files were named based on the participant's number and film name. The real names of the participants are not recorded in the experiment.

4.2 Signal Preprocessing and Normalization

According to the rules of data conservation of the University of Ottawa, the data collected by both hard copy and electronically, must be kept in a secure manner. All the data is received by the researcher's laptop, and only the researcher and the supervisor can access it. They will be kept for a minimum of 5 years by the supervisor on the University of Ottawa campus. After the conservation period, the data will be destroyed.

4.2.1 Subject's Emotion Confirmation

As we know, the type of movie we watched is not entirely responsible for the emotions at that precise moment. For example, some people do not feel sad when they are watching

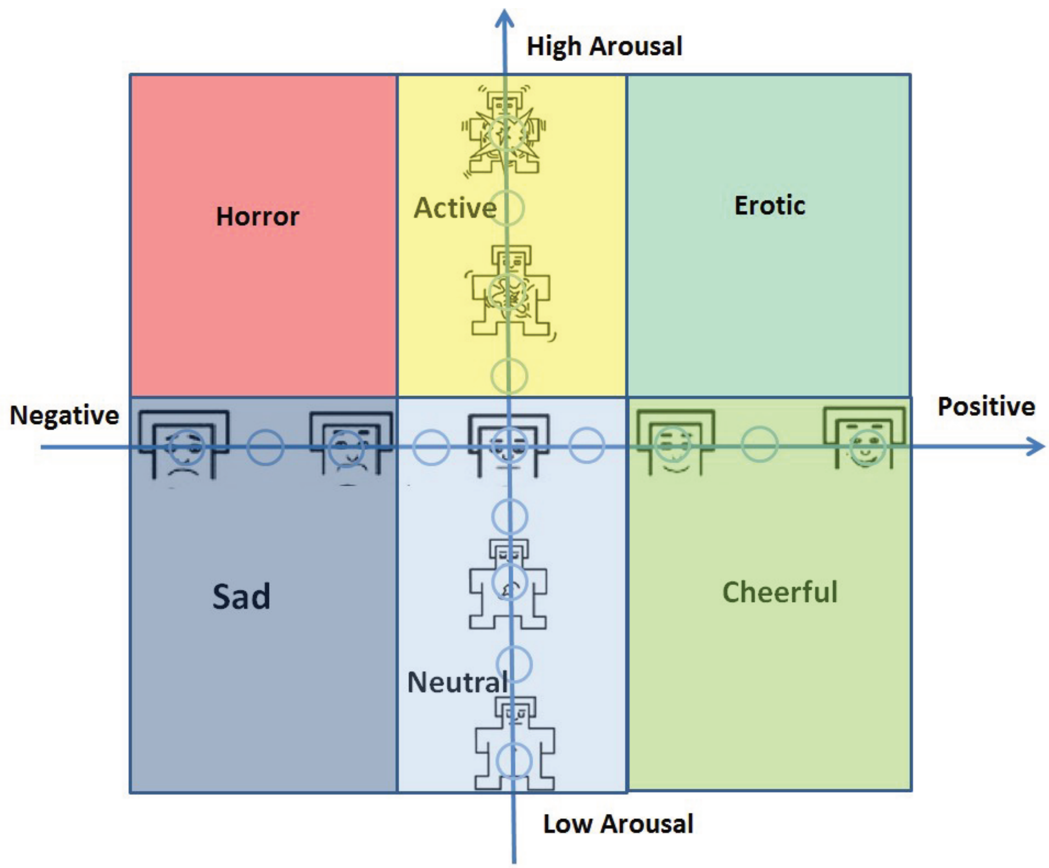


Figure 4.6: Emotion distinction model

a movie clip in the sad category, and some people even feel cheer when they watch a horror movie. In order to get the participant's true feelings, we built an emotion distinction model to help identify the emotions. As we see in Figure 4.6, the model consists of six blocks identifying six emotions with different colors. According to the SAM, the x axis corresponds to feelings from negative to positive, and the y axis represents arousal, from low to high. Moreover, there are small figures on both the x and the y axis, starting from a frowning, unhappy figure to a smiling, happy figure, and from a sleepy figure to an excited figure. As we mentioned, the target emotions include horror, erotic, sad, cheer, and neutral. The active emotion is a supplementary part of the model.

4.2.2 Data Reduction

The signal recording procedure cannot avoid getting a great deal of redundancy data. In order to reduce the redundant data and make the data process more efficient, we developed a number of Matlab functions to handle the work of data reduction. By using these functions, each type of signal record was trimmed to a fixed length of 10 seconds, according to the time tags.

4.3 Emotion Feature Extraction

The first phase of the feature extraction: ECG signal is a kind of periodic signal, and we extract its features in the frequency domain. Whereas, the respiration, CO deflection and thoracic impedence signals are nonperiodic signals, we obtain their features in the time domain. Specifically, in the frequency domain, the feature considered is mean frequency, which is defined as

$$F_{Mean}(\omega) = \frac{\int_0^\infty \omega PSD_S(\omega)\, d\omega}{\int_0^\infty PSD_S(\omega)\, d\omega} \tag{4.1}$$

where $F_{Mean}(\omega)$ is the mean frequency. $PSD_S(\omega)$ is the power spectrum density of the signal S. ω is the frequency variable. In this paper, we compute the power spectrum density by Fast Fourier Transform (FFT) as the power spectrum format is identical to the real part of the FFT, i.e.,

$$PSD_S(\omega) = FFT_S(\omega) \cdot FFT_S^*(\omega) = |FFT_S(FFT_S(\omega))|^2 \tag{4.2}$$

Here, $FFT_S(\omega)$ is the Fast Fourier Transform of the signal S. $FFT_S^*(\omega)$ is the complex conjugate of $FFT_S(\omega)$. In the time domain, the features we chose include mean and variance. Hence, we can define an emotion feature vector (i.e., the feature vector under emotion E) in our study as

$$\Phi_E = \begin{bmatrix} F_{Mean}(S_{ECG}) \\ Mean(S_{TI}) \\ Var(S_{TI}) \\ Max(S_{TI}) \\ Min(S_{TI}) \\ Mean(S_{RES}) \\ Var(S_{RES}) \\ Max(S_{RES}) \\ Min(S_{RES}) \\ Mean(S_{CO-D}) \\ Var(S_{CO-D}) \\ Max(S_{CO-D}) \\ Min(S_{CO-D}) \end{bmatrix} \tag{4.3}$$

where $F_{Mean}(\cdot)$, $Mean(\cdot)$,$Var(\cdot)$,$Max(\cdot)$ and $Min(\cdot)$ are the operators to calculate mean frequency, mean, variance, maximum value, and minimum value of signal S, respectively.

The second phase of the feature extraction HRV was calculated and acquired from raw ECG signals, is based on the sequence of RR intervals. SDNN, a straightforward and useful metric of HRV, is also calculated from RR intervals. RMSSD calculates the successive differences between neighboring RR intervals. pNN50 is defined as the mean number of times per hour in which the change in consecutive NN exceeds 50 ms. HF, LF and VLF were gained by analyzing the fluctuations in the frequency domain. LF band (0.04-0.15 Hz) and HF band (0.15-0.40 Hz) are related to the parasympathetic effects of Akselrod et al. (1981). Next, mean value, max value, min value and standard deviation, which are typical statistical values, were calculated from basic signal values. Finally, 24 features were extracted from the raw data, as shown in Table 4.6

4.4 Emotion Template Matching

By supposing we have feature vectors under five emotion categories: cheer (Φ_{Cheer}), horror (Φ_{Horror}), erotic (Φ_{Erotic}), sad (Φ_{Sad}) and neutral ($\Phi_{Neutral}$), we define the emotion template as

$$\Phi = \left[\begin{array}{ccc} \Phi_{Cheer} & \Phi_{Horror} & \Phi_{Sad}\Phi_{Erotic}\Phi_{Neutral} \end{array} \right] \quad (4.4)$$

Our objective in this subsection is to match an undefined emotion feature vector $\Phi_{Undefined}$ at moment t with the mentioned defined emotion template. In the following data processing, we apply three supervised learning methods (nearest neighbor, discriminant analysis and multilayer perception) to compare the matching performance. The overall data is divided into two datasets: training dataset and testing dataset.

No.	Features	Abbr.	Unit
1	Electrocardiography	ECG	mV
2	Heart Rate Variability	HRV	ms
3	Mean of Heart Rate Variability	HRV_Mean	ms
4	Max value of Heart Rate Variability	HRV_maxVal	ms
5	Min value of Heart Rate Variability	HRV_minVal	ms
6	Standard deviation of normal-to-normal intervals	SDNN	ms
7	Root mean square of successive difference	RMSSD	ms
8	Proportion of NN50 divided by total number of n-to-ns	pNN50	%
9	High frequency of Heart Rate Variability	HF	Hz
10	Low frequency of Heart Rate Variability	LF	Hz
11	Vary low frequency of Heart Rate Variability	VLF	Hz
12	LF over HF	LF/HF	%
13	Mean of respiration	Resp_mean	Ohm
14	Respiration Variability	Resp_var	Ohm
15	Min value of respiration	Resp_minVal	Ohm
16	Max value of respiration	Resp_maxVal	Ohm
17	Mean of CO deflection	CO_D_mean	Ohm
18	Min value of CO deflection	CO_D_minVal	Ohm
19	Max value of CO deflection	CO_D_maxVal	Ohm
20	CO deflection variability	CO_D_var	Ohm
21	Max value of thoracic impedence	ThoracicImp_maxVal	Ohm
22	Min value of thoracic impedence	ThoracicImp_minVal	Ohm
23	Mean of thoracic impedence	ThoracicImp_mean	Ohm
24	Thoracic impedence variability	ThoracicImp_var	Ohm

Table 4.6: The List of Emotion Features

4.4.1 Nearest Neighbor (NN)

The nearest neighbor method classifies an undefined emotion feature vector based on its similarity to other defined emotion features, where the similarity is measured by distance. Specifically, each feature signal in the training dataset is first normalized as

$$S_{p,i}^{N} = \frac{2\left(S_{p,i} - min\left(S\right)\right)}{max\left(S\right) - min\left(S\right)} - 1 \tag{4.5}$$

where p denotes one of the emotion features and can be ECG, HR, Resp_var or CO_D_mean. The superscript N indicates the normalization. $min\left(S\right)$ and $max\left(S\right)$ are the manipulators to calculate the minimum value and maximum value of $\left(S\right)$. We then calculate the distance between two emotion feature vectors by City Block Distance, which is defined as

$$Dis_{j,k} = \left\|\Phi_j^N - \Phi_k^N\right\|_1 \tag{4.6}$$

Thus, the categorical response of emotion classification can be simply obtained by

$$min\,Dis_{Undefined,\Phi} \tag{4.7}$$

4.4.2 Discriminant Analysis (DA)

Our discriminant analysis predictive model is composed of discriminant functions based on the physiological features obtained from the subjects. We assume our record cases are independent and that the values of features have a multivariate normal distribution. The emotion membership is assumed to be mutually exclusive, that is, no record case belongs to more than two emotions, and collectively exhaustive, which means all record cases are members of an emotion.The minimum and maximum value of the signal variable are specified by the filter of the receiver application, and all of the values are integers.

Discriminant analysis assumes that different emotions have different emotion feature data with different Gaussion distributions. It typically classifies two emotions by maximizing the ratio of the between-emotion-class scatter to the within-emotion-class scatter. For the case of multiple emotions, the two-emotion classification method can easily be extended to a multi-emotion classification case by getting a subspace that contains all the class variability. Specifically, in two-emotion classification, the within-emotion-class scatter is

$$\sigma_{within}^2 = p_1 \times cov_1 + p_2 \times cov_2 \tag{4.8}$$

where p_1 and p_2 are the the apriori probabilities of the two emotion classes 1 and 2. cov_1 and cov_2 are the covariance of class 1 and 2. The between-emotion-class scatter is

$$\sigma_{between}^2 = (S_1 - \mu_1) \times (S_1 - \mu_1)^T \tag{4.9}$$

The emotion classifier can be obtained by

$$max \, \frac{\sigma_{between}^2}{\sigma_{within}^2} \tag{4.10}$$

4.4.3 Multilayer Perception (MLP)

The multilayer perception is a learning network used to classify the emotions based on a criteria that minimizes the error between supervised classification and ground truth (i.e., subject's agreement). The mentioned network is a feed-forward network and in our work, there are two hidden layers. Specifically, we first calculate a combination of emotion features as

$$A_j = \sum_{i=1}^{8} w_{ji}\Phi_E(i) \tag{4.11}$$

where w_{ji} is the weight of each element in the feature vector and $\Phi_E(i)$ is the i-th element in Φ_E. Based on A_j, we obtain the output of MLP by the sigmoidal function

$$O_j = \frac{1}{1+e^{A_j}} \tag{4.12}$$

Then we can calculate the weights w_{ji} by minimizing the error between the calculated output emotion and the defined emotion in the training procedure, i.e.,

$$w_{ji} = argmin \sum_j \left\| O_j - \Phi_{defined} \right\|_2 \tag{4.13}$$

A detailed architecture diagram of our MLP network with one input layer, two hidden layers and one output layer is shown in Figure 4.7. In the input layer, the values of features are transmitted as the input values (i.e. the input values of ECG, HRV or other physiological signals for each record in our model) to the computing units in the first hidden layer.

In our case, the final result of the output layer is the estimated emotion. The activation function in the hidden layers is chosen as the Hyperbolic Tangent function, i.e.,

$$y(a) = tanh(a) = \frac{e^a - e^{-a}}{e^a + e^{-a}} \tag{4.14}$$

In the training process, to estimate the weights, the MPL model takes a random record dataset and uses the simulated annealing. After the training procedure of the random records, the model gets the initial weights and calculates the derivative of the Sum of

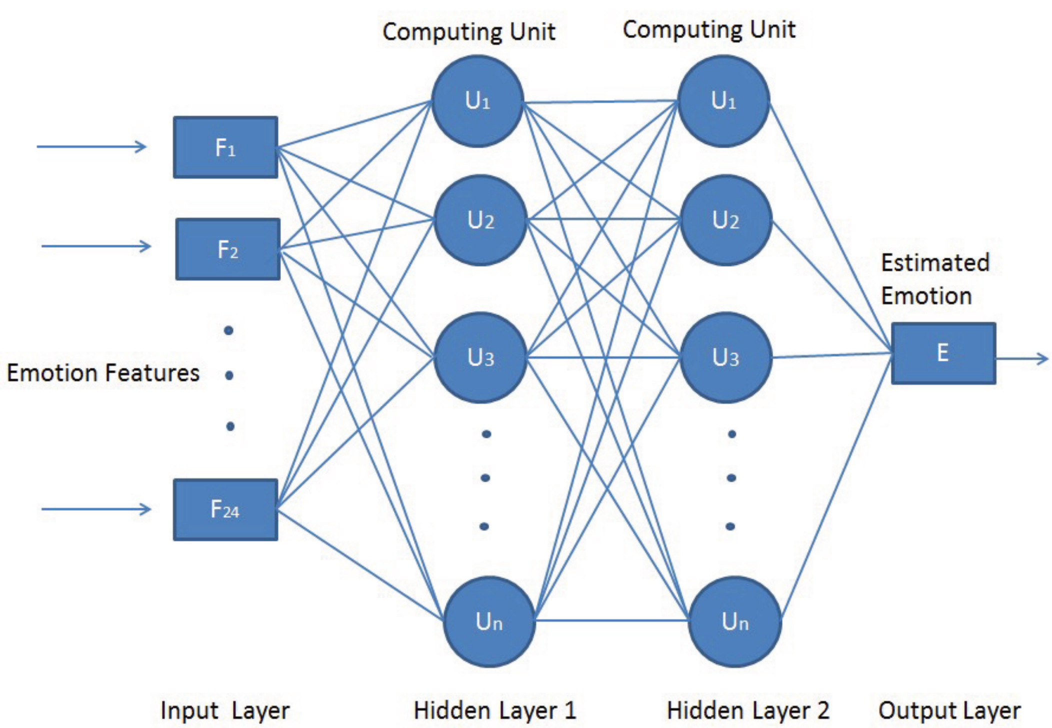


Figure 4.7: MLP Model Architecture

Squares. The Sum of Squares is computed by first calculating the differences between each of the scores and their mean.

At the end of the training step, the model applies the Gradient Descent method to update the estimated weights. The Gradient Descent algorithm is known as a first-order optimization algorithm and is used to find the local minimum of a function.

Chapter 5

Results and Discussion

5.1 Overall Performance Comparison

In our research, we used the IBM Statistical Package for the Social Sciences (SPSS) Statistics for statistical analysis.

The Table 5.1 shows the cross-classification of the observed emotions versus the predicted emotions of the target subjects in the NN model, when k =3. In the table, erotic has the best recognition, with a rate of 66.7%. However, the model is not good at recognizing horror and sadness, and couldn't predict cheer and neutral.

OBSERVED	PREDICTED					
	Cheer	Erotic	Horror	Neutral	Sad	Percent Correct
Cheer	0	3	3	1	3	0.0%
Erotic	1	10	2	1	1	66.7%
Horror	2	8	3	0	2	20.0%
Neutral	0	5	3	0	0	0.0%
Sad	0	8	2	0	3	23.1%

Table 5.1: The Classification Results of NN

OBSERVED	PREDICTED					
	Horror	Erotic	Sad	Cheer	Neutral	Percent Correct
Horror	9	1	1	2	2	60.0%
Erotic	1	9	3	1	1	60.0%
Sad	0	1	6	5	1	46.2%
Cheer	1	1	1	6	1	60.0%
Neutral	1	1	1	1	4	50.0%

Table 5.2: The Classification Results of DA

OBSERVED	PREDICTED					
	Cheer	Erotic	Horror	Neutral	Sad	Percent Correct
Cheer	5	1	1	0	0	71.4%
Erotic	1	8	0	1	0	80.0%
Horror	1	0	8	0	0	88.9%
Neutral	0	1	1	5	0	71.4%
Sad	0	0	1	1	6	75.0%

Table 5.3: The MLP Classification Result

In the Table 5.1, the emotion recognition results of the DA model are displayed. Horror, erotic, and cheer have the same best recognition rates of 60%. The correct detection percentage of neutral and sad are 50% and 46.2 % respectively.

The Table 5.1 displays the results of using the MLP analysis for each categorical dependent variable, by partition and overall. In the table, we can see that the recognition rates of cheer, erotic, horror, neutral, and sad are of 71.4%, 80.0%, 88.9%, 71.4%, and 75.0% respectively. The number of cases classified incorrectly for each emotion category is also shown in the table.

Analysis Model	Features Selection
NN	HF, LF_over_HF, HRV_mean, SDNN, RMSSD, HRV_maxVal, LF, VLF
DA	HRV_maxVal, SDNN, RMSSD, HF, LF, VLF, HRV_mean, LF_over_HF
MLP	LF, Resp_mean, Resp_Var, CO_deflection_mean, HF, HRV_mean, LF_over_HF

Table 5.4: Models' Features Selection

The average recognition rates for NN, DA, and MLA are of 40%, 55.7%, and 77.34%, respectively. In the following sections, we explain the detailed analysis of the three models.

5.2 Features Selection Comparison

The Table 5.4 shows the features selection for the three models. For the NN model, the features (HF, LF_over_HF, HRV_mean, SDNN, RMSSD, HRV_maxVal, LF, VLF) are more important than the other features for emotion recognition.

In the DA model, we get the best feature set from its Structure Matrix, which indicates HRV_maxVal, SDNN, RMSSD, HF, LF, VLF, HRV_mean, and LF_over_HF.

In the MLP model, the importance of all features can be calculated in the process, as shown in Table 5.1. We obtained the best feature set, which includes LF, Resp_mean, Resp_Var, CO_deflection_mean, HF,and HRV_mean. In the following sections, we discuss feature selection in great detail.

5.3 Detailed Analysis in Discriminant Analysis

As we metioned in the Chapter 3, we use the discriminant analysis to create a model for predicting emotions. Dependent on linear combinations of the emotion feature sets, the model could provide a discrimination between the five emotions. During the analysis, we

Independent Variable Importance

	Importance	Normalized Importance
Resp_Max	.043	85.2%
Resp_Min	.045	88.3%
Resp_mean	.046	91.4%
Resp_Var	.048	94.8%
CO_deflection_Max	.041	82.0%
CO_deflection_Min	.046	90.8%
CO_deflection_mean	.047	93.9%
CO_deflection_Var	.045	88.7%
Thoracic_max	.046	91.2%
Thoracic_min	.045	88.8%
Thoracic_mean	.046	91.1%
Thoracic_var	.047	92.8%
VLF	.045	89.3%
LF	.050	100.0%
HF	.048	94.8%
LF_over_HF	.047	93.2%
HRV_mean	.048	95.4%
SDNN	.044	88.1%
RMSSD	.047	92.3%
pNN50	.039	77.8%
HRV_maxVal	.045	88.5%
HRV_minVal	.043	84.9%

Figure 5.1: MLP's Features Importance

create the functions based on the training data for which the emotion is known. Then the functions could be applied to the test case that have unknow emotion type.

Before we run the analysis, we have to define the range of our data. If a value of one feature is outside of the range, this feature will be not used in the discriminant analysis. Then we need to select the good features for the analysis. In our case, we used the wilks' lambda method for entering or removing the features. The following subsection describe the detailed procedure of discriminant anaylsis.

5.3.1 Output Results

The target of feature selection is to figure out which features are most relevant to each emotion. It helps to reduce the data redundancy and computational cost. The presence of noisy or irrelevant features may degrade the accuracy of the algorithms. We use the stepwise discriminant analysis to compare and test the relevance between features and classes. The stepwise regression could be used for entering or removing new features. The stepwise regression starts with a model that doesn't include any of the features.

At each step, the model add the features with the largest F to Enter value which larger than the entry criteria. The features which have F to Enter values smaller than the entry criteria are moved out from the model. The stepwise regression select features based upon the statistical data. Finally, we got two function groups respectively. Group 1 includes HRV_maxVal, SDNN, RMSSD, HF, LF, VLF, HRV_mean, and LF_over_HF. Group 2 has pNN50 and CO_deflection_mean.

The Figure 5.5 shows the variance of each function groups and the relative efficacy of each discriminant function. The eigenvalue means the importance of each dimension. When the eigenvalue is larger, it means the more of the variance in the emotion is explained

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	.758	49.0	49.0	.657
2	.410	26.5	75.5	.539
3	.249	16.1	91.6	.447
4	.129	8.4	100	.338

Table 5.5: Eigenvalues

Test of Functions	Wilks' Lambda	Chi-square	df	Sig
1 through 4	.286	61.976	64	.548
2 through 4	.503	34.037	45	.884
3 through 4	.709	17.031	28	.948
4	.885	6.022	13	.945

Table 5.6: Wilks' Lambda Table

by the function. In this model, function group 1, 2 and 3 are much more important than function group 4. We can fairly ignore the forth safely, because the percentage of variance is only 8.4%. The percentage of variance also indicates the variance of each function. The percentage of function 1 is 49% and we consider this feature set has the most irrelevance with the emotions in our model. The forth column shows the cumulative percentage and the fifth displays the canonical correlation. The canonical correlation is the measure of association between the discriminant function and the emotions.

In the Table 5.6 , the first column means each function is tested on condition that the means of the functions are equal across groups. The second column contains the values of Wilks' Lambda of the functions. Wilks' lambda is a measure of how well each function separates features into groups. The lower the value is, the more important the function is. It's the proportion of the total variance. The third column includes the value of chi-square test.

Standardized Canonical Discriminant Function Coefficients

	Function			
	1	2	3	4
SDNN	1.993	1.393	.624	-.785
RMSSD	.082	-.202	.115	.267
pNN50	-.168	-.619	.256	-.284
HRV_minVal	1.390	1.503	.586	.966
HRV_maxVal	-1.317	-1.724	-1.392	2.527
HRV_mean	.438	-.459	5.649	-4.451
HF	-.131	1.074	-4.927	1.322
LF_over_HF	-.176	-.455	-.562	-.100
CO_deflection_Var	-2.688	.942	-.245	-1.049
CO_deflection_mean	.320	.755	-.077	.421
CO_deflection_Min	-1.753	.144	.555	-.288
CO_deflection_Max	.856	-1.147	.479	1.242
Resp_Var	1.752	.112	1.412	1.708
Resp_Min	.328	.510	.610	.671
Resp_Max	-.825	.621	-.146	-.335
Resp_mean	1.543	-.352	.773	.897

Figure 5.2: Standardized Canonical Discriminant Function Coefficients

As we know, the associated chi-square test is used to get the variance of a normally distributed population which based on the value of Wilks' Lambda. In the column of significance, the value of function 1 is the least. It indicates that the smaller significance value gets better result of separating the emotions.

The Figure 5.2 shows the features measured on different scales. The absolute values of coefficients indicate the relative importance of the features in discriminating emotions. The greater the absolute value is, the better discriminating result is. In function group 1, the most five features are CO_deflection_var, SDNN, CO_deflection_Min,

Structure Matrix

	Function			
	1	2	3	4
HRV_maxVal	.366*	.042	.109	.057
SDNN	.363*	.002	-.030	.066
RMSSD	.309*	-.008	.129	-.006
HF	.306*	.128	.110	-.085
LF[a]	.306*	.128	.110	-.085
VLF[a]	.306*	.128	.110	-.085
HRV_mean	.300*	.126	.177	-.100
LF_over_HF	-.232*	-.097	-.052	-.085
pNN50	.075	-.574*	.288	-.114
CO_deflection_mean	.196	.387*	.119	.191
Resp_mean	.117	-.191	.420*	-.102
CO_deflection_Min	-.102	.202	.378*	-.196
HRV_minVal	.025	.110	.132*	-.009
CO_deflection_Max	.036	-.111	-.096	.513*
CO_deflection_Var	-.113	-.080	-.040	.348*
Resp_Var	-.125	.063	-.112	.308*
Resp_Min	.002	.011	.290	-.297*
Resp_Max	-.049	-.045	-.048	.290*

Figure 5.3: Structure Matrix

Resp_Var and Resp_mean. For function group 2, the HRV_maxVal, HRV_minVal, SDNN, CO_deflection_max and HF are more important than the others. For function group 3, the absolute values of HRV_mean, HF, Resp_Var, HRV_maxVal and Resp_mean are the top five. As we mentioned, the function group 4 doesn't play a important role in the model. We ignored the data of function group 4 safely.

The structure matrix as shown in Figure 5.3. The figure shows the correlations of each feature with each discriminant function. In the column of each function, some values are

EmotionNum	Function 1	Function 2	Function 3	Function 4
1	1.194	.518	.182	.138
2	-.916	.595	-.442	.088
3	-.612	-.563	.590	.306
4	-.145	-.068	.272	-.750
5	.654	-1.087	-.810	.015

Table 5.7: Function at group centroids

marked with a star. It means they are all lager than .30 and picked as best features for the function. The correlations serve as factor loading and the ordering is different from that in the standardized coefficients table due to the collinearity among the features.

Finally, we got the best features for each function group: The HRV_maxVal, SDNN, RMSSD, HF, LF, VLF and HRV_mean are the best features for function 1. The pNN50 and CO_deflection_mean are the best features for function group 2. The resp_mean, CO_deflection_min and HRV_minVal are the best features for function group 3. The CO_deflection_max, CO_deflection_var, Resp_var, Resp_min and Resp_max are nearly useless for emotion classification in our model.

Acquired these three function groups, canonical discriminant analysis is used to get the possible multiple correlation with the emotions. The Territorial Map uses the first function and the secord function as x and y axis. At this step, function group 1 and 2 can be taken as the first and the second canonical correlation. We got the relationship between the emotions and the discriminant functions from the territorial map, as described in Figure 5.4. The result we obtained is that the function group 1 is good at distinguishing Horror and Erotic,

```
                                            Territorial Map
                          (Assuming all functions but the first two are zero)
Canonical Discriminant
Function 2
       -8.0      -6.0      -4.0      -2.0        .0       2.0       4.0       6.0       8.0
          +---------+---------+---------+---------+---------+---------+---------+---------+
     8.0 +                                          21                                    +
         I                                          21                                    I
         I                                          21                                    I
         I                                          21                                    I
         I                                          21                                    I
         I                                          21                                    I
     6.0 +          +         +         +         + 21      +         +         +         +
         I                                         21                                     I
         I                                         21                                     I
         I                                         21                                     I
         I                                         21                                     I
         I                                         21                                     I
     4.0 +          +         +         +         +21       +         +         +         +
         I                                         21                                     I
         I                                         21                                     I
         I                                         21                                     I
         I                                         21                                     I
         I                                         21                                     I
     2.0 +          +         +         +         +21       +         +         +         +
         I                                         21                                     I
         I                                         21                                     I
         I                                         21                                     I
         I                                   *     21   *                                 I
         I                                       2241                                     I
      .0 +          +         +         +     222*4 41      +         +         +         +
         I                             222222233344 4111                                  I
         I                       2222223333333 * 334455511111                             I
         I                 222222333333            35*  5555511111                        I
         I          2222222333333                 35         5555511111                   I
         I    2222223333333                      35              5555511111              I
    -2.0 +2222333333+         +         +       35+         +         +5555511111         +
         I3333                                  35                          5555511111    I
         I                                     35                              5555511111I
         I                                    35                                    55555I
         I                                    35                                          I
         I                                   35                                           I
    -4.0 +          +         +         +   35    +         +         +         +         +
         I                                  35                                            I
         I                                 35                                             I
         I                                35                                              I
         I                                35                                              I
         I                               35                                               I
    -6.0 +          +         +         35        +         +         +         +         +
         I                             35                                                 I
         I                             35                                                 I
         I                            35                                                  I
         I                           35                                                   I
         I                           35                                                   I
    -8.0 +                          35                                                    +
          +---------+---------+---------+---------+---------+---------+---------+---------+
       -8.0      -6.0      -4.0      -2.0        .0       2.0       4.0       6.0       8.0
                                   Canonical Discriminant Function 1
```

Figure 5.4: The territorial map

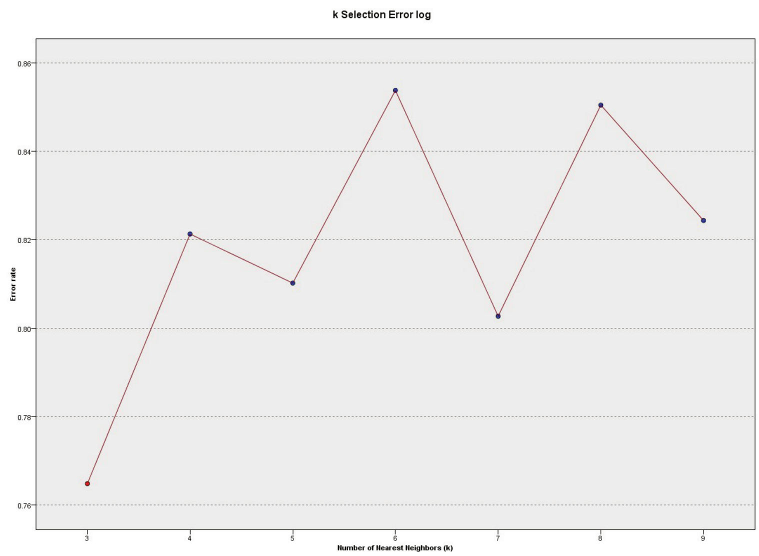


Figure 5.5: K Selection Error log

Sad and Neutral. Function group 2 separates Horror and Neutral, Erotic and Cheer. Sad is not easy to be decided in this model. The classification results show in the Figure **??**

5.4 Detailed Analysis in Nearest Neighbor

5.4.1 Number of Nearest Neighbors

As we know the different value of K has different classification result and the best choice of K depends upon the data which we collected. For our study, six different values of K (K=3-9) are applied to calculate the classification results. In order to specify the number of nearest neighbors, we use the V-fold cross-validation to capture the “best” value of K. First, 7 sub-samples are created by the cross-validation. Each sub-sample generates a nearest neighbor model with different number of nearest neighbors. Compared with each error rate, we get the best number of nearest neighbor.

The Figure of K Selection Log tells the error rates for each value of K. In the graph, X and Y axes respectively represents the values of K and Error rate. The points show the error rates according to different values of K and no evidence shows K and error rate have a linear relation. In our models, the model with 3 nearest neighbors produced the lowest error rate. In our case , the best result came out with K = 3.

5.4.2 Enter Feature Selection

To improve the result of emotion classification, we have to specify the most relevant features from the feature database. In the NN model, we used the forward selection to decide the best feature set. First, we added the features which calculated by discriminant analysis into the model. We tried to capture all the important features by increasing the number of features to select , which maybe increase the error rate of the model. At the same time, we built a tight model by decreasing the number of features, which may cause missing some important features. After the process of feature selection, we get the feature set as shown in Figure 5.4.

5.4.3 Training and Testing

Before the data analysis, the data group need to be divided into training and testing samples by a specific method. The training partition is used to build the nearest neighbor model by comparing the signal data. The testing partition is the other part of the data records. In our model, we specify 70% of cases to assign to the training partition. The rest 30% are assigned to the testing partition.

5.4.4 Output Results

The Figure 5.6 shows the feature space of predictors of the Nearest Neighbor model. The predictor space chart convey a lot information. Each axis represents a predictor in the model. In the graph, X, Y and Z axes respectively represents the values of HRV_mean, HF and LF_over_HF. The points show the values of these predictor and different colors mean different kind of emotions. The circle points belong to the training data and the triangle points means the testing data. The red focal points are shown linked to their k nearest neighbors. In the case we can see the nearest points of the focal points and the focal emotion is close to the related ones.

The Figure 5.7 shows the focal emotion selected in the feature space and their K nearest neighbor on each feature. In the chart, we can clearly see every 1-dimensional slice of the feature space. The title of each peers chart means the name of feature and the value of each point is displayed at the left of the chart. In order to simplify the charts, we make the tag of each point be the number in stand of the name of emotion. The 1, 2, 3, 4 and 5 represent horror, erotic, sad, cheer and neural. It clearly tells each focal points have more closest points than other ones. For example, in the SDNN slice, the value of the points of neutral are higher than the other emotion values. These slices simply provide some more details about our nearest neighbor model.

5.5 Detailed Analysis in Multilayer Perceptron

We use the Multilayer Perceptron (MLP) to build a predictive model for the five emotions based on the values of our features. The multilayer perception is a neural network which can distinguish data are not linearly separable. Our multilayer perceptron model consists of three layers: an input layer, two hidden layers and an output layer. The input layer includes 22 features, such as ECG, HRV, SDNN and so on. The hidden layer use the hyperbolic tangent as the activation function. The output layer takes the emotion category as the dependent variable.

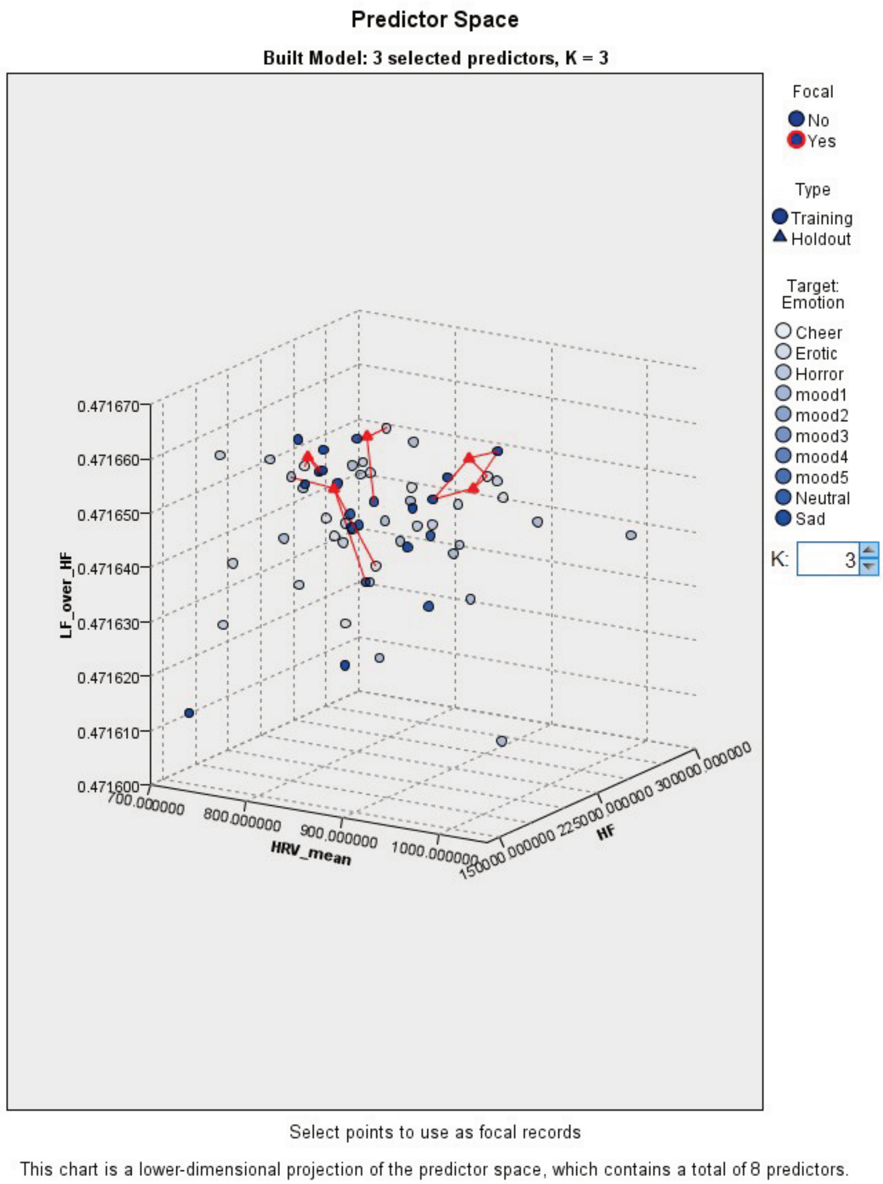


Figure 5.6: The Predictor Space

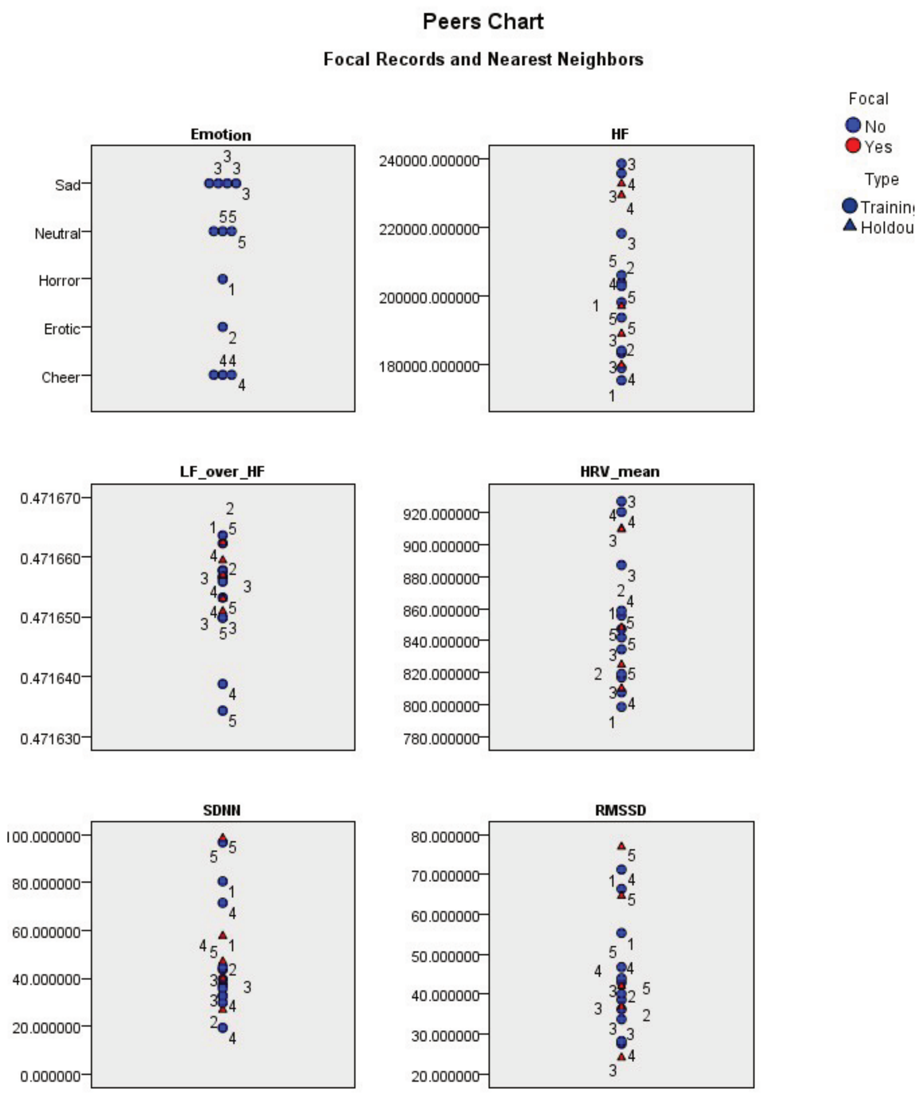


Figure 5.7: Peers Charts

5.5.1 Partition Dataset

As we mentioned in nearest neighbor analysis, the whole data need to be divided by the training and testing dataset. We used a different stratagem to get the two partitions. First, we created a partition variable to assign cases. The values of the partition value only could be 0 or 1. In our model, the cases with partition value equals 0 are set as the training case and the cases with a value of 1 are assigned to the testing sample.

In order to assign the cases fairly, a random number need to be produced before the assigning procedure. The simulated annealing algorithm is used in automatic architecture selection. However, to reproduce the same randomized results, we use the same initialization value for the random number generator. If the partition variable is missing, the case will be treated as invalid. In our model,

5.5.2 Model Option

To get the best result of our test, we specified the structure of the multilayer perceptron. Three parts of the model can be costumed: the hidden layers, the output layer and the dependent variables. The hidden layers of the model contain the network units which are unobservable. The network units are the activation functions which contain the weighted values of input. We used two hidden layers in the process. The second hidden layer is almost the same as the first one but its networks units get the weighted sum of the unit of the first layer. Both activation functions of the two layers are the same.

We used the hyperbolic tangent for the hidden layers. The hyperbolic tangent transforms all the arguments to the range (-1,1). For the output layer, we also used the Hyperbolic tangent function as activation function. Except for the Hyperbolic, the sigmoid is also apply for our model. The main difference between these two functions is the range of the result of the sigmoid function which is (0,1). Since our dependent variables are the name of emotion, we don't rescale the dependent variables.

5.5.3 Methods for Training

Different training types may produce different results. Three training types could be adopted in our model: Batch, Online and Mini-batch. The common point of the three methods is that all of them determine how the model processes the feature records. The difference is when to update the synaptic weights. The batch method renews the synaptic weights after getting all training feature records. The online method updates them after checking each signal record. The method of mini-batch is a bit more complex. It divides the training data into same size groups and updates the weights after checking each single group.

We consider the online method is the best way to train our records among these methods according to the efficiency. Also we can change the arguments of the scaled conjugate gradient algorithm for improving the result: initial lambda, initial sigma, interval center and interval offset. The scale of initial lambda and initial sigma parameter are separately (0,0.000001) and (0, 0.0001). The parameter of interval center and interval offset are used for weight initialization and architecture selection. In our model, the values of these parameters are 0.0000005, 0.00005, 0 and 0.5 respectively.

5.5.4 Output Results

All the possibilities for each emotion are displayed in Figure 5.8. In the predicted-by-observed chart, each emotion is displayed as clustered box plots of predicted pseudo-probabilities for the combined five emotions. The x axis represents the emotions and the y axis represents the correct predictions. It is shown that the scale of y axis is between

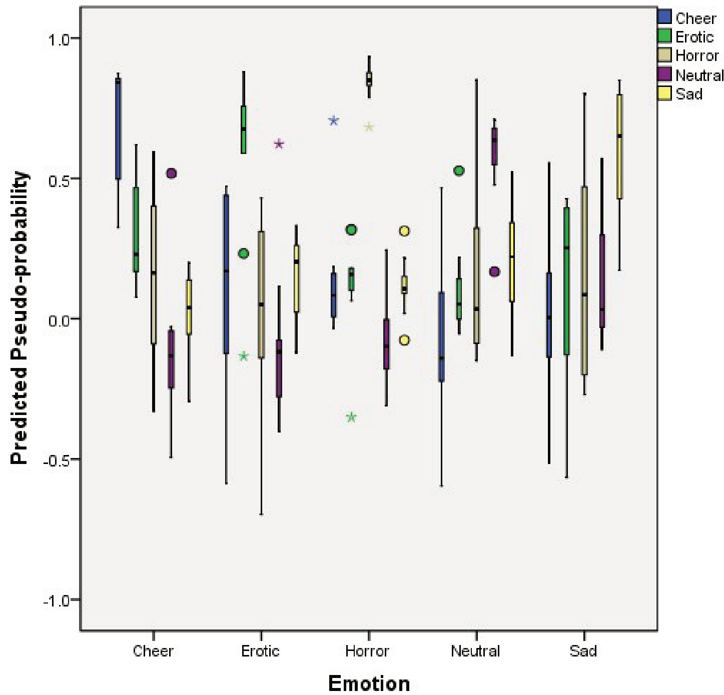


Figure 5.8: Predicted Pseudo-probability

(-1,1) not (0, -1), since the pseudo-probabilities need to be rescaled by dividing by their sum. If the value of pseudo-probability is negative, this value will be added to all pseudo-probabilities with its absolute value before the rescaling. After this procedure, they will be rescaled to be between (0,1). Some diagrams are created depended on the pseudo-probabilities, such as cumulative gains, lift charts, and ROC.

The ROC (Receiver Operating Characteristic) curve is a kind of chart which illustrates the sensitivity and specificity for each categorical dependent variable. For each given emotion, the ROC chart displays one curve for each one with different color and the area under the curve means the probability that predicted pseudo-probability of being in that category.

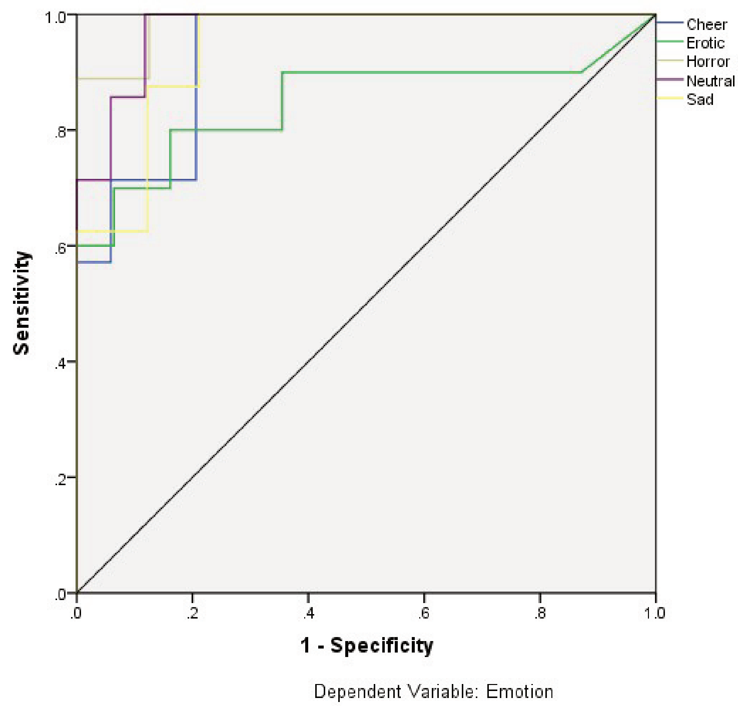


Figure 5.9: ROC curve

In our ROC chart, the x axis represents the value of (1 - specificity) and the y axis represents the sensitivity. Since our dependent variable has five categories, "each curve treats the category at issue as the positive state versus the aggregate of all other categories".

The Table 5.8 is a summary of the ROC curve. It shows the area of each emotion category, which represents the probability that the predicted emotion result for a randomly chosen positive case will exceed the result for a randomly chosen negative case.

The Figure 5.10 is a cumulative gains chart that shows the percentage of all the records in a given category "gained" by targeting a percentage of the total number of records. For example, the first point on the curve for the Erotic category is approximately at (10%, 50%), meaning that if you get a dataset and sort all of the cases by predicted pseudo-probability

Emotion	Area
Cheer	.933
Erotic	.848
Horror	.986
Neutral	.975
Sad	.943

Table 5.8: ROC area under the curve

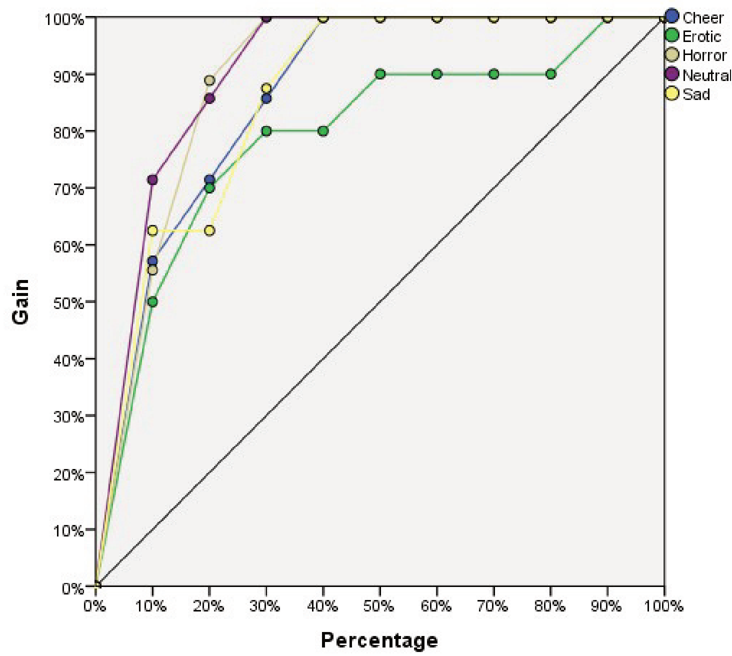


Figure 5.10: Gain-Percentage

of Erotic, you would expect the top 10% to contain approximately 50% of all of the cases that actually take the category Erotic. Likewise, the top 20% would contain approximately 70% of the defaulters, the top 30% of cases would contain 80% of defaulters, and so on. If you select 100% of the scored dataset, you obtain all of the defaulters in the dataset. As we know, the farther above the diagonal line a curve lies, the better the gain.

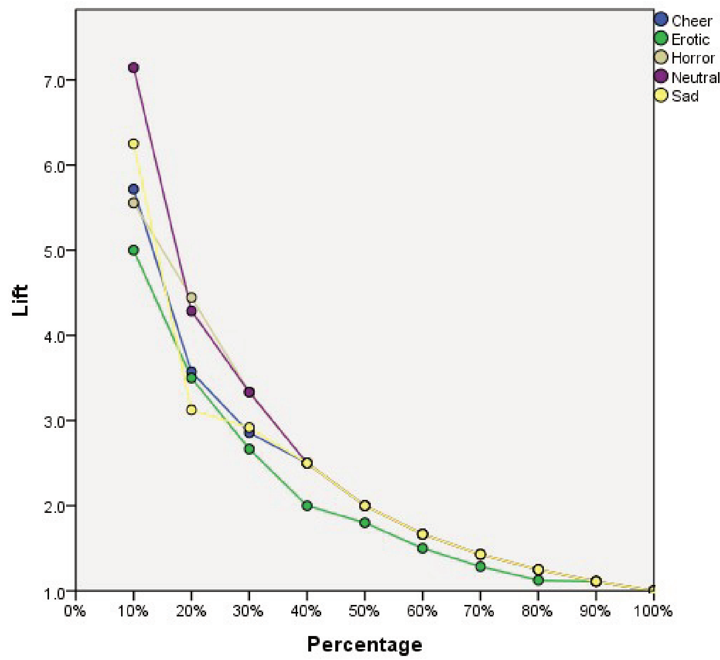


Figure 5.11: Lift-Percentage

The Figure 5.11 is a lift chart which calculated from the cumulative gain chart. In this chart, the x axis correspond to the percentage of all cases and y axis corresponds to the ratio of the cumulative gain for each curve to the diagonal line. For example, for the Erotic category, the lift at 10% is approximately 50%/ 10% = 5.0. We could say that the life chart provide a different way to check the information in the above cumulative gains chart.

By using the multilayer perceptron model, we could predict the probability that our system to recognize a unknown emotion based on the physiological signals. The model results are comparable to those obtained using discriminant analysis and nearest neighbor.

Chapter 6

Conclusion and Future Work

In this chapter, we summarize the work presented in this thesis and provide a conclusion. Then, we discuss certain topics that we did not treat and propose future work that could contribute to this field.

6.1 Conclusion

The detection of people's emotions in their everyday life is a new and very interesting topic. We first compared several technologies for affective computing, such as facial expressions, posture, tone of voice, and gestures. Since wearable digital devices are increasingly powerful, the methods used to detect physiological signals are more convenient and user-friendly than ever before. We proposed the use of physiological signals to identify people's emotions.

Based on the device and the physiological signal receiver application, we conducted a reasonable experiment for emotion elicitation and data collection. In our design, we obtained hundreds of records of physiological signals for each emotion, and we improved the emotion elicitation by using EMDB and SAM. Moreover, our system and experiment could be deployed in different environments, which means the process of signal acquisition is not complex.

In order to make the emotion elicitation more reasonable and practical, we proposed an emotion model for our emotion classification, including the emotions: cheer, sad, horror, erotic, and neutral. The scale of emotions covers the most common feelings and is independent. Moreover, we developed some functions of Matlab for signal prepossessing, data normalization, and emotion feature extraction. We successfully obtained 24 features from the processing. To the best of the author's knowledge, some sets of the features are the first to be used in predicting emotions.

We built three emotion classification models: nearest neighbor, discriminant analysis, and multilayer perception, in order for our system to achieve the highest level of accuracy for emotion recognition. The model of discriminant analysis not only got a satisfactory level of accuracy for emotion recognition, but it also helped us understand that some feature sets play a greater role in predicting our five emotions.

The model of nearest neighbor had a poor result compared with other two models, however, it showed the correlation between each feature and each emotion. Finally, the model of multilayer perception achieved the best results among the three models. The average recognition rates are of 40%, 55.7%, and 77.34%, for NN, DA, and MLA respectively.

6.2 Future Work

Emotions are a complex state of feeling for a human being, and the physiological signals related to these emotions also vary greatly depending on the person. We therefore think that personal information such as age, gender, job, hobbies, and culture, could be very helpful for emotion recognition. In further work, the concept of Big Data could be included in the work, to help build a magnanimity database. We believe it will help improve our system's recognition rate.

On the other hand, we would also like to take advantage of our emotion recognition system in other technological fields, such as computer games, special education, and social networks. Some computer games could benefit from emotion detection; based on the player's emotions, the game could adjust the level of difficulty by itself in order to satisfy the different levels of players. Also, if the system detects that the player is feeling anxious or tired, the system could remind the player to take a break.

The emotion recognition system could also be applied to e-learning. For example, the system could get the feedback on the student's emotions and send the information to the teacher. Based on the feedback, the teacher could determine the student's level of acceptance for the content and adjust the teaching method accordingly. It will therefore benefit both the teacher and the student.

At last, as we previously mentioned, the people's true emotions are still not being shared in social networks. As mobile devices become more and more powerful, we could implement the emotion analysis model in the mobile device, for example, an android mobile phone.

The system may combine the emotions and the multimedia, and post both to the social network. For example, when the user takes a photo, the system could calculate the emotion at the moment and attach the emotion with the photo. Once the user posts the photo to the social network, people could see the user's emotion in the photo, at that moment.

Appendix A

First Appendix

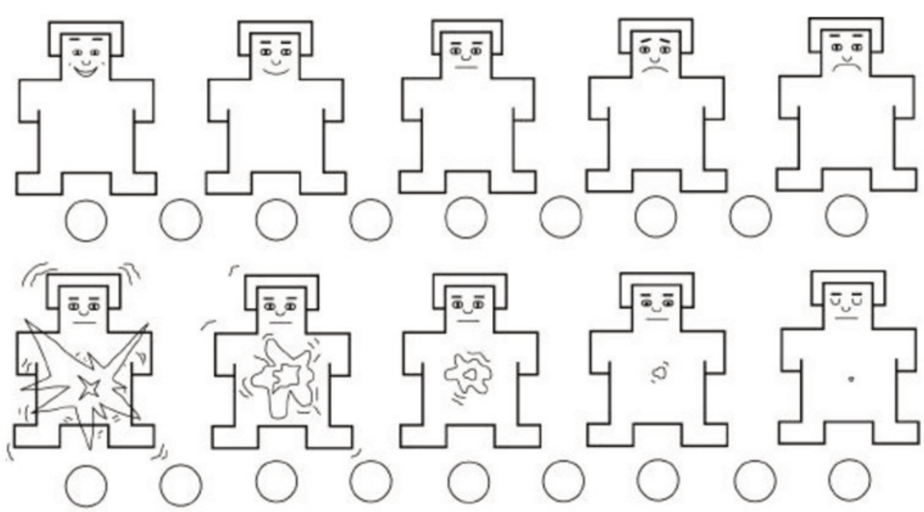
A.1 The Questionnaire Form

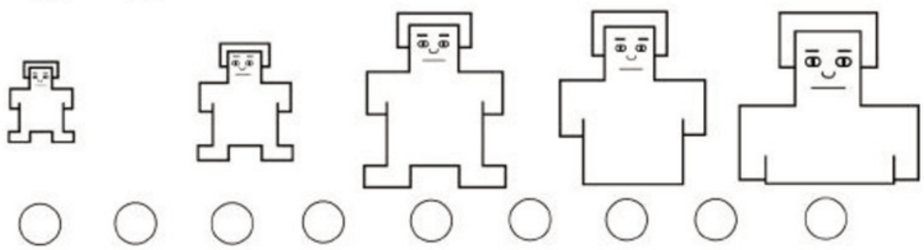
The Emotion Questionnaire

Subject No.: Gender: Age:
Movie Flip No.:

Q1: Have you watched this movie before?

Q2: Have you closed your eyes or look away during the clip presentation?





For **Pleasure**, SAM ranges from a smiling, happy figure to a frowning , unhappy figure;

For **Arousal**, SAM ranges from sleepy with eyes closed to excited with eyes open.

For **Dominance** scale shows SAM ranging from a very small figure representing a feeling of being controlled or submissive to a very large figure representing in-control or a powerful feeling.

Bibliography

Solange Akselrod, David Gordon, F. Andrew Ubel, Daniel C. Shannon, A. Clifford Berger, and Richard J. Cohen. Power spectrum analysis of heart rate fluctuation: A quantitative probe of beat-to-beat cardiovascular control. *Science*, 213:220–222, 1981.

Bradley M. Appelhans and Linda J. Luecken. Heart rate variability as an index of regulated emotional responding. *Review of General Psychology*, 10:229–240, 2006.

S. Asteriadis, P. Tzouveli, K. Karpouzis, and S. Kollias. Estimation of behavioral user state based on eye gaze and head pose ¡ªapplication in an e-learning environment. *Multimedia Tools and Applications*, 41:469–493, 2009.

Erik Cambria, Andrew Livingstone, and Amir Hussain. *Erik Cambria; Andrew Livingstone and Amir Hussain (2012). "The Hourglass of Emotions"...* Springer, 2012.

W. B. Cannon. The james-lange theory of emotions: a critical examination and an alternative theory. *The American Journal of Psychology*, 39:106–126, 1927.

Sandra Carvalho, Jorge Leite, Santiago Galdo-A lvarez, and Oscar F. Goncalves. The emotional movie database (emdb): A self-report and psychophysiological study. *Appl Psychophysiol Biofeedback*, 37:279–294, 2012.

Paul Ekman. An argument for basic emotions. *Cognition and Emotion*, 6:169–200, 1992.

David Heise. Enculturating agents with expressive role behavior. *In Trappl, Robert. Agent Culture: Human-Agent Interaction in a Mutlicultural World. Lawrence Erlbaum Associates*, pages 127–142, 2004.

Ikuo Homma and Yuri Masaoka. Breathing rhythms and emotions. *Experimental Physiology*, 93:1011–1021, 2008.

L. James. *Road Rage and Aggressive Driving: Steering Clear of Highway Warfare*. Amherst, 2000.

C.D. Katsis, N. Katertsidis, G. Ganiatsas, and D.I. Fotiadis. Toward emotion recognition in car-racing drivers: a biosignal processing approach. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 38:502–512, 2008.

Z. Khalili and M. H. Moradi. Emotion detection using brain and peripheral signals. *Biomedical Engineering Conference CIBEC*, pages 1–4, 2008.

Jonghwa Kim and Elisabeth Andre. Emotion recognition based on physiological changes in music listening. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30: 12, 2008.

Yelin Kim and Emily Mower Provost. Emotion classification via utterance-level dynamics: a pattern-based approach to characterizing affective expressions,. *ICASSP*, 26-31:3677–3681, 2013.

Vitaliy Kolodyazhniy, Sylvia D. Kreibig, James J. Gross, Walton T. Roth, and Frank H. Wilhelm. An affective computing approach to physiological emotion specificity: toward subject-independent and stimulus-independent classification of film induced emotions. *Psychophysiology*, 48:908–922, 2011.

P.J. Lang, M.M. Bradley, and B.N. Cuthbert. International affective picture system (iaps): technical manual and affective ratings. 1997.

C. Lisetti and F. Nasoz. Maui: a multimodal affective user interface. *In Proceedings of the ACM Multimedia International Conference*, 2002.

Christine latitia Lisetti and Fatma Nasoz. Using noninvasive wearable computers to recognize human emotions from physiological signals. *EURASIP Journal on Applied Signal Processing*, 11:1672–1687, 2004.

Margaret M. Bradley and Peter J. Lang. Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25:49–59, 1994.

Daniel S. Quintana, Adam J. Guastella, Tim Outhred, Ian B. Hickie, and Andrew H. Kemp. Heart rate variability is associated with emotion recognition: direct evidence for a relationship between the autonomic nervous system and social cognition. *International Journal of Psychophysiology*, 86:168–172, 2012.

Kanlaya Rattanyu, Michiko Ohkura, and Makoto Mizukawa. Emotion monitoring from physiological signals for service robots in the living space. *International Conference on Control, Automation and Systems*, pages 580–583, 2010.

N. A. Remington, L. R. Fabrigar, and P. S. Visser. Re-examining the circumplex model of affect. *Journal of Personality and Social Psychology*, 79:286–300, 2000.

Elisabeth Ruiz-Padial, John Sollers III, J. Jaime Vila, and Julian F. Thayer. The rhythm of the heart in the blink of an eye: emotion-modulated startle magnitude covariates with heart rate variability. *Psychophysiology*, 40:306–313, 2003.

James A. Russell and Lisa Feldman Barrett. Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *Journal of Personality and Social Psychology*, 76:805–819, 1999.

S. Schachter and J.E. Singer. Congitive, social, and physiological determinants of emotional state. *Psychological Review*, 5:379–399, 1962.

Jianhua; Tieniu Tan Tao. Affective computing: a review. *Affective Computing and Intelligent Interaction*, pages 981–995, 2005.

Szwoch Wioleta. Using physiological signals for emotion recognition. *Sopot*, pages 06–08, 2013.