

ADAPTIVE INTELLIGENT USER INTERFACES WITH EMOTION RECOGNITION

by

FATMA NASOZ

B.S. Bogazici University, Turkey, 2000

M.S. University of Central Florida, USA, 2003

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the School of Computer Science
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Summer Term
2004

Major Professor:
Dr Christine L. Lisetti

© 2004 by Fatma Nasoz

ABSTRACT

The focus of this dissertation is on creating Adaptive Intelligent User Interfaces to facilitate enhanced natural communication during the Human-Computer Interaction by recognizing users' affective states (i.e., emotions experienced by the users) and responding to those emotions by adapting to the current situation via an affective user model created for each user. Controlled experiments were designed and conducted in a laboratory environment and in a Virtual Reality environment to collect physiological data signals from participants experiencing specific emotions. Algorithms (k-Nearest Neighbor [KNN], Discriminant Function Analysis [DFA], Marquardt-Backpropagation [MBP], and Resilient Backpropagation [RBP]) were implemented to analyze the collected data signals and to find unique physiological patterns of emotions. Emotion Elicitation with Movie Clips Experiment was conducted to elicit Sadness, Anger, Surprise, Fear, Frustration, and Amusement from participants. Overall, the three algorithms: KNN, DFA, and MBP, could recognize emotions with 72.3%, 75.0%, and 84.1% accuracy, respectively. Driving Simulator experiment was conducted to elicit driving-related emotions and states (panic/fear, frustration/anger, and boredom/sleepiness). The KNN, MBP and RBP Algorithms were used to classify the physiological signals by corresponding emotions. Overall, KNN could classify these three emotions with 66.3%, MBP could classify them with 76.7% and RBP could classify them with 91.9% accuracy. Adapta-

tion of the interface was designed to provide multi-modal feedback to the users about their current affective state and to respond to users' negative emotional states in order to decrease the possible negative impacts of those emotions. Bayesian Belief Networks formalization was employed to develop the User Model to enable the intelligent system to appropriately adapt to the current context and situation by considering user-dependent factors, such as: personality traits and preferences.

To my parents, Kamile and Sadan Nasoz, who were always there for me...

ACKNOWLEDGMENTS

I would like to gratefully thank my advisor Dr Christine Lisetti for her endless support, help, motivation and patience, and for teaching me everything I know about how to conduct research. Besides being the best advisor and teacher I could dream of, she is also a great person who is supporting her students in every possible way.

I would very much like to thank Dr Erol Gelenbe for his precious advices during my studies at UCF.

I would also like to especially thank my dissertation committee members Dr Michael Georgiopoulos, Dr Kien Hua, and Dr Annie Wu for their support and understanding.

I would like to thank Dr Sebastian van Delden and Dr Mark Llewellyn for kindly helping me to find participants for my simulation experiments.

Thanks to all my student friends who worked with me and helped me while designing and conducting my experiments, especially Kaye Alvarez, Onur Ozyer, Kivanc Oner, Ning Yu, Andreas Marpaung, Zoran Nikoloski, George Harris, and Dahai Guo. I also would like to thank all the people who accepted to participate in my experiments.

I also would like to deeply thank my parents, Kamile and Sadan Nasoz for supporting and motivating me in every stage of my life.

TABLE OF CONTENTS

List of Tables	xii
List of Figures	xiii
CHAPTER 1 INTRODUCTION	1
1.1 Problem Statement	3
1.2 Non-invasive Multimodal Sensing Devices	6
1.3 Expected Broader Impact and Applications	8
1.3.1 Driving Safety	8
1.3.2 Training/Learning	10
1.3.3 Telemedicine	12
1.4 Dissertation Overview	14
CHAPTER 2 BACKGROUND	15
2.1 MAUI: A Multimodal Affective User Interface	15
2.1.1 User Text Input	16
2.1.2 Ongoing Video and Captured Image	17

2.1.3	Feedback to User	17
2.1.4	Avatar	18
2.2	Emotion Elicitation	20
2.2.1	Emotion Elicitation with Movie Clips	20
2.2.2	Other Emotion Elicitation Techniques	21
2.3	Emotion Recognition from Physiological Signals	22
2.4	A Previous Study on Measuring Drivers' Stress	30
2.5	User Modeling	34
2.5.1	User Modeling of Affect	34
2.5.2	Other User Modeling Techniques	37
CHAPTER 3	EMOTION ELICITATION EXPERIMENTS	39
3.1	Pilot Panel Study	39
3.2	Emotional Signals Data Generation and Collection	43
3.3	Pattern Recognition of Physiological Signals with Machine Learning	45
3.3.1	k -Nearest Neighbor Algorithm	47
3.3.2	Discriminant Function Analysis	48
3.3.3	Marquardt Backpropagation	50
3.4	Further Evaluation with Feature Extraction	53

CHAPTER 4 DRIVING SIMULATOR EXPERIMENT IN VIRTUAL REALITY	58
4.1 Driving Simulator Experiment Scenario	61
4.2 Driving Simulator Experiment Setup	65
4.3 Emotion Recognition with Machine Learning	67
4.3.1 Feature Extraction	67
4.3.2 Emotion Recognition Accuracy with KNN, MBP, and Resilient Back-propagation Algorithms	69
CHAPTER 5 USER MODELING AND INTELLIGENT USER INTERFACE ADAPTATION	73
5.1 Bayesian Belief Networks	75
5.2 User Modeling with Bayesian Belief Networks	76
5.3 Interface Actions chosen by the Driver Model	80
CHAPTER 6 VISUALIZATION OF THE RESULTS	86
CHAPTER 7 FUTURE WORK	97
7.1 Real-Life and Real-Time Applications	97
7.2 More Modalities	98
7.3 Improved Algorithms for Emotion Recognition	99

7.4	User Modeling for Different Applications	99
7.5	Different Wearable Computers	100
CHAPTER 8 CONCLUSIONS		101
APPENDIX A ABOUT THE STUDENT		104
A.1	Short Biography	105
A.2	Publications	106
A.2.1	Journal Publications	106
A.2.2	International Refereed Conferences	107
A.2.3	Articles in Progress	108
APPENDIX B FORMS AND QUESTIONNAIRES USED IN THE EXPER- IMENTS		109
B.1	Forms and Questionnaires for Emotion Elicitation with Movie Clips Experiment	110
B.1.1	IRB Consent Form	110
B.1.2	Pre-Study Questionnaire	113
B.1.3	In-Study Questionnaire	115
B.2	Forms and Questionnaires for Driving Simulator Experiment in Virtual Reality	118
B.2.1	IRB Consent Form	118
B.2.2	Pre-Study Questionnaire	122

B.2.3 Post-Study Questionnaire	124
LIST OF REFERENCES	127

LIST OF TABLES

2.1	Previous Research on Emotion Recognition from Physiological Signals	29
2.2	Sensors Used to Measure the Physiological Signals	31
3.1	Movies from Gross and Levenson ([GL95])	41
3.2	Hit Rates and Average Intensities for Movies with $\geq 90\%$ Agreement	42
3.3	Movie Scenes Selected for the Study	42
3.4	Hit Rates and Average Intensities	45
4.1	Emotion Classification Accuracy with KNN for Each Emotion	70
4.2	Emotion Classification Accuracy with MBP for Each Emotion	70
4.3	Emotion Classification Accuracy with RBP for Each Emotion	71
4.4	Emotion Classification Accuracy with KNN, MBP and RBP for all emotions	72
4.5	Consequences Related to Emotion Recognition	72
5.1	Optimal Actions Chosen by the Model for Each Combination of Events . . .	82

LIST OF FIGURES

1.1	Human Multimodal Affect Expression matched with Multimedia Computer Sensing	5
1.2	a. BodyMedia Sensewear b. Polar Chest Strap	6
1.3	Graph Created by InnerWear Research Software.	7
2.1	MAUI:Multimodal Affective User Interface.	16
3.1	Emotion Recognition Results with KNN Algorithm	48
3.2	Emotion Recognition Results with DFA Algorithm	50
3.3	Neural Network Architecture	52
3.4	Emotion Recognition Results with MBP Algorithm	53
3.5	Emotion Recognition with More Feature Extraction	54
3.6	Emotion Recognition Results with KNN Algorithm with More Feature Ex- traction	55
3.7	Emotion Recognition Results with DFA Algorithm with More Feature Ex- traction	56

3.8	Emotion Recognition Results with MBP Algorithm with More Feature Ex- traction	56
4.1	The Car Operating in Virtual Reality	59
4.2	The Virtual Reality Environment	60
4.3	Control Room of the Driving Simulator	60
4.4	Child Walking to the Middle of the Road	62
4.5	Big Garbage Truck Making a 3-point Turn	63
4.6	Driver Waiting for the Pedestrians to Cross the Road	64
4.7	Elicitation and Recognition of Driving-related emotions from Physiological Signals.	68
4.8	Neural Network Architecture for Recognizing Driving-related emotions . . .	69
5.1	Integrating User Modeling in Emotion Recognition and User Interaction. . .	74
5.2	Bayesian Belief Network Representation of Model of the Driver.	77
5.3	Bayesian Belief Network Representation of Case 1	82
5.4	Bayesian Belief Network Representation of Case 2	83
5.5	Bayesian Belief Network Representation of Case 3	84
5.6	Bayesian Belief Network Representation of Case 4	85
6.1	Avatar Mirroring User's Anger and Sadness Respectively.	87

6.2	MAUI Before Recognizing the User.	88
6.3	MAUI After Recognizing the User.	89
6.4	Various Anthropomorphic Avatars for Different User Preferences.	90
6.5	Showing video as a Multimodal Stimulus and Avatar mirroring user's Neutral State	91
6.6	Avatar Mirroring User's Happiness	92
6.7	Avatar Mirroring User's Sadness	93
6.8	Avatar Mirroring User's Anger	94
6.9	MAUI as a research development tool to recognize emotions from pre-collected physiological data.	95
6.10	Avatar mirroring the emotion recognized from the physiological data.	96
7.1	BodyMedia Wireless Transmitter	98
7.2	Procomp+	100

CHAPTER 1

INTRODUCTION

In the recent years there has been an increasing attempt to develop computer systems and interfaces that recognize the affective states of users, learn their preferences and personality, and adapt to these accordingly [SFK02], [BL02], [HM02], [Con03], [LNL03]. Studies conducted in this direction also gave birth to a new field in Computer Science: Affective Computing [Pic97].

The main motivation of all these studies is that humans are social beings that emote and are affected by their emotions and that machine perception needs to be able to capture such phenomenon in order to enhance our everyday digital tools. Previous studies propose that people emote while they are interacting with computers [RN96]; moreover drivers emote while driving in their cars [Jam00] and patients do emote, too [Dam94]. But, is this reason enough for us to attempt to create *Affective Interfaces* with an *Intelligent Agent* that recognize the user's emotional states and respond accordingly?

People do not only emote, but they also are affected by their emotional states. Emotions influence various cognitive processes in humans. These include: perception and organization

of memory [Bow81], categorization and preference [Zaj84], goal generation, evaluation, and decision-making [Dam94], strategic planning [Led92], focus and attention [DT92], motivation and performance [CLN00], intention [Fri86], communication [Bir70], [EF75], [Cho91], and learning [Gol95].

Besides experiencing and being affected by emotions in real-life environments, humans also emote while they are in a *virtual environment* (VE, henceforth). The performance of the user while interacting within the VE; and the level (or quality) of *presence* experienced by the user in the VE are affected by the emotional states of the user carry the main importance.

One of the first authors on presence was Marvin Minsky [Min80], who coined the term *telepresence*, referring to the human operator's sense of being at a remote real environment in a teleoperation. In more recent work on presence, both virtual and real environments were taken into consideration. Ijsselstein *et al.* [IRF00] defined presence as the user's perception of 'being there' in a mediated environment, while Lombard and Ditton's [LD97] definition of presence was "the perceptual illusion of nonmediation". Lombard and Ditton [LD97] further conceptualized presence along six dimensions: social richness, realism, transportation, immersion, social actor within medium, and medium as a social actor. *Co-presence* was defined by Casanueva and Blake [CB01] as the user's sense that i) there are other participants existing in the VE and that ii) s/he is having interaction with real people.

Research studies on presence show that users are responding socially and emotionally to the systems, the characters in the VEs, or to the robots that they are interacting with. For example, emotions affect the user's perception of a VE [IJs02], [LD97], and the interaction

with the VE, in turn, affects the user's emotions [DKF00], [Kal00]. In addition, Freeman and Avons [FA00] reported that participants felt more present when they were emotionally involved in what was going on.

The strong interface between emotion and cognition and the effects of emotion on humans performances in VEs make it desirable, if not necessary, to create intelligent computer systems that understand users' emotional states, learn their preferences and personality, and respond accordingly. Section 1.1 describes this research study's approach to this problem.

1.1 Problem Statement

The field of Human-Computer Interaction (HCI) within Computer Science is defined as "a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them" [HBC92]. Maybury [May01] defines optimal human-computer interactions as ones where users are able to perform complex tasks more quickly and accurately, thus improving user satisfaction. The main objective of this dissertation is to create intelligent systems that i) achieve the goals of human-computer interaction defined by Maybury [May01] and ii) enhance social presence through algorithms for real-time emotion recognition and adaptation. The systems that are desired to create will be applicable to driving safety, training, and telemedicine.

The main focus of this dissertation is recognizing emotions in humans and developing Affective Intelligent User Interfaces that can adapt to these emotions for optimal Human-Computer Interaction. Objectives of this study include:

- Designing experiments to elicit emotions from participants, collect and measure their physiological signals and record their facial expressions. These experiments include: an in-lab experiment to elicit anger, fear, sadness, surprise, frustration, and amusement via multimodal stimuli and a driving simulator experiment in order to elicit driving related emotions such as anger and panic.
- Analyzing measured physiological signals by implementing machine learning algorithms to find unique patterns of physiological signals for each emotion.
- Using these patterns to recognize user's emotions and develop intelligent affective user interfaces that give appropriate feedback to the user about her/his emotional state.
- Creating models of users and integrating them with the interfaces developed, in order to adapt to the user's emotional state, her personality and to her preferences in the current context and application.

Figure 1.1 shows the overall architecture of this approach [LN02]. The system's input is both mental and physiological components of a particular emotion experienced by the user. Physiological components are to be identified and collected from observing the user by using receiving sensors with different modalities: *Visual* (Facial Expressions), *Kinesthetic*

(Autonomic Nervous System [ANS] Arousal and Motor Activities), and *Auditory* (Vocal Intonation) (V, K, A). The system is also intended to receive input from *Linguistic* tools (L) in the form of linguistic terms for emotion concepts, which describe the subjective experience associated with a particular emotion.

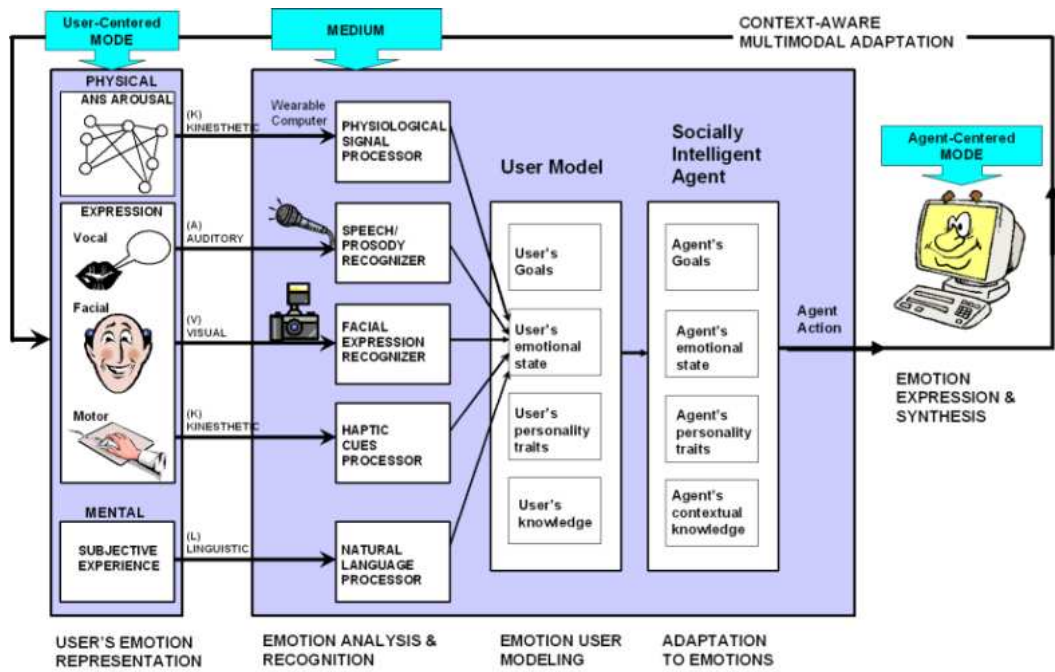


Figure 1.1: Human Multimodal Affect Expression matched with Multimedia Computer Sensing

The input is interpreted by implementing various pattern recognition algorithms such as Artificial Neural Networks. The output of the system is given in the form of a synthesis for the most likely emotion concept corresponding to the sensory observations. This synthesis

constitutes a descriptive feedback to the user about her current emotional state, including suggestions as to what next action might be possible to change state.

In this dissertation, the piece where emotion recognition is performed from ANS Arousal using various Machine Learning Techniques and appropriate feedback and response are given based on the user's current state and preferences is implemented.

1.2 Non-invasive Multimodal Sensing Devices

BodyMedia SenseWear Armband and Polar chest strap that works in compliance with the armband will be used as multimodal sensing devices to collect ANS arousal data from participants of the emotion elicitation experiments: These equipment are shown in Figure 1.2.



Figure 1.2: a. BodyMedia Sensewear b. Polar Chest Strap

BodyMedia SenseWear Armband (Figure 1.2a) is a versatile and reliable wearable body monitor created by BodyMedia, Inc. It is worn on the upper arm and includes a galvanic

skin response sensor, skin temperature sensor, 2-axis accelerometer, heat-flux sensor, and a near-body ambient temperature sensor. The system also includes a chest strap which works in compliance with the armband for heart rate monitoring. SenseWear Armband is capable of collecting, storing, processing, and presenting physiological signals such as *Galvanic Skin Response* (GSR), *Heart Rate*, *Temperature*, *Movement*, and *Heat Flow*. After collecting the signals the SenseWear Armband is connected to the InnerWear Research Software that is also developed by BodyMedia, Inc. either with a dock station or wirelessly, and transfers the data. The data can either be stored in XML files for further interpretation with pattern recognition algorithms or the software itself can process the data and present it using graphs. Figure 1.2 shows a graph created by the InnerWear Research Software.

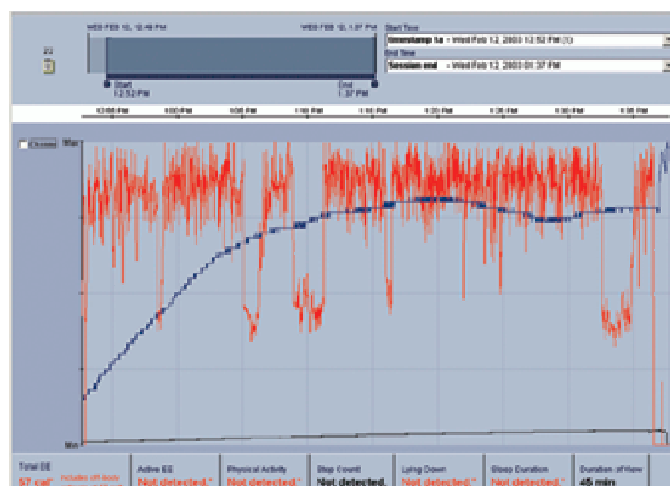


Figure 1.3: Graph Created by InnerWear Research Software.

Until now, all the studies performed on measuring physiological signals for emotion recognition (some of them are described in Section 2.3) used sensing devices that were connected to the body at one end, and connected to the computer at the other end. The most appealing feature of the SenseWear Armband is that it is wireless, hence enabling the user to wear the non-invasive armband and to be immersed in real-life situation. Some of the organizations that use BodyMedia Sensewear for body monitoring are: NASA and The National Bio-Computation Center, Sandia National Laboratories, Virginia Tech, and University of Pittsburgh Medical Center. However, we are the first research group who use the armband for emotion recognition purposes.

1.3 Expected Broader Impact and Applications

This section discusses the possible applications where emotions play an important role and it is usable and desirable to develop affective interfaces with an intelligent agent that can recognize and adapt to users' emotional state in the current application and context.

1.3.1 Driving Safety

Aggressive driving in the United States results in 425,000 deaths and 35 million injuries per decade and approximately costs \$250 billion per year [Jam00]. The inability to manage

one's emotions while driving is often identified as one of the major causes for accidents. As mentioned by James, "anger intensifies aggressiveness and judgment becomes impaired accordingly, . . . angry aggression accelerates and explodes into impulsive and sometimes violent action and depending on the aggressive action, one suffers deep regret, embarrassment, financial loss, depression, injury, or death" (p.111).

Anger is one of the emotions that negatively affect one's driving. When drivers become angry, they start feeling self-righteous about events and anger impairs their normal thinking and judgment, their perception is altered, thus leading to the misinterpretation of events [Jam00]. For example, when a driver misses a green light she may get impatient to pass at the next green light even if speeding for it might put her in danger. Other states that lead to negative effects are *frustration*, *anxiety*, *fear*, and *stress* [FS02]. In order to be a safer driver on the highways, a person needs to be better aware of her emotions and possesses the ability to manage them effectively [Jam00].

Once drivers are aware of their emotional states it becomes easier for them to respond to the situation in a safe manner, but drivers can often lack in awareness. For example, some drivers often lack the ability to calm themselves down when they are angry or frustrated. Another example is, sleepiness, which is one of the most dangerous states to be in while driving, yet when people find they are sleepy, they often force themselves to continue driving instead of stopping to rest [Jam00].

James [Jam00] and Larson [LR99] discussed techniques for drivers to manage their anger including relaxation techniques to reduce physical arousal and mental reappraisal of the

situation. The aim of creating affective intelligent user interfaces for driving safety application is to enhance *co-presence* in the driving environment. For example, when the system recognizes the anger or rage of a driver it might suggest the driver to perform a breathing exercise [LR99]. Similarly, when the system recognizes driver's sleepiness, it might change the radio station for a different tune or roll down the window for fresh air. Having a natural communication between this system and the driver, and taking the precautions mentioned above according to drivers' personal preferences will increase the drivers' feeling that there exists a real person in the car with them to assist them while they drive.

1.3.2 Training/Learning

Learning is one of the cognitive processes that are affected by user's emotional state. *Frustration* for example leads to a reduction in the ability to learn [LW89]. Rozell and Gardner's [RG00] study points out that when people have negative attitudes towards computers, their self-efficacy toward using them decreases, which then reduces their chances of performing computer-related tasks (when compared to those with positive attitudes towards computers). This research also emphasized that individuals with more positive affect exert more effort on computer-related tasks.

Another emotion that influences learning is *anxiety*. In training situations, anxiety is presumed to interfere with the ability to focus cognitive attention on the task at hand because that attention is preoccupied with thoughts of past negative experiences with similar

tasks, in similar situations [Mar94], [WB95]. It follows that learning may be impaired when trainees are experiencing high levels of anxiety during training. Indeed, with a sample of university employees in a microcomputer class, Martocchio [Mar94] found that anxiety was negatively related to scores on a multiple choice knowledge test at the end of training. In addition, individuals who had more positive expectations prior to training had significantly less anxiety than individuals who had negative expectations of training.

Anxiety also appears to influence reactions to training. For example, with a sample of British junior managers enrolled in a self-paced management course, Warr and Bunce [WB95] found that task anxiety was positively related to difficulty reactions in training. Individuals who experienced high task anxiety perceived training to be more difficult than individuals who experienced low task anxiety. In this study, interpersonal and task anxiety were assessed prior to training. Task anxiety was significantly higher than interpersonal anxiety and only task anxiety was associated with difficulty reactions. Finally, in their meta-analytic path analysis, Colquitt *et al.* [CLN00] reported that anxiety was negatively related to motivation to learn, pre-training self-efficacy, post-training self-efficacy, learning, and training performance.

In summary, the most consistent findings are that anxiety is negatively related to self-efficacy, motivation, learning, and training performance. In addition, social anxiety may influence training outcomes when trainees are taught new tasks as a team. Furthermore, facilitating a mastery orientation towards the task may help to reduce the anxiety (*e.g.*,

attitude change) experienced during training and allow trainees to focus their cognitions on the task at hand, resulting in better learning [Mar94].

Affective intelligent user interfaces will enhance *presence* and *co-presence* in the learning environment. For example when the system recognizes that the learner is anxious, in response, it might provide encouragement in order to reduce anxiety and allow the individual to focus more attention on the task. Similarly, when the system recognizes the learner as being frustrated or bored it might adjust the pace of the training accordingly so that the optimal level of arousal for learning is achieved. Finally, when the system recognizes that a person is confused it might clarify the information just presented. All these adaptation techniques will improve the learner's sense of being in a real classroom environment where a live instructor would typically recognize these same emotions and respond accordingly.

1.3.3 Telemedicine

Tele-Home Health Care (Tele-HCC) has been performed in United States since the early 1990's. Tele-HHC provides communication between medical professionals and patients in cases where hands-on care is not required, but regular monitoring is necessary. For example, tele-HHC interventions are currently used to collect vital sign data remotely (*e.g.*, ECG, blood pressure, oxygen saturation, heart rates, and breath sounds), verify compliance with medicine and/or diet regimes, and assess mental or emotional status [ARC96], [CKC96], [DC00], [War97]. With increasing use of Tele-HHC, it is important that the care giver and

care recipient communicate along the affective channel to allow for better assessment and responsiveness. However, formulating an assessment may be particularly difficult in tele-HHC settings where patients are treated and monitored remotely by medical professionals using multiple media devices with social and emotional cues filtered out. Social presence during patient-physician communication is indeed essential. Furthermore, the rising use of Tele-HHC signifies a need for efforts aimed at enhancing such presence. Not only may appropriate emotional state assessment be a key indicator of the patient's mental or physical health status, but the power of emotions themselves over the recovery process has also been documented [Dam94].

Affective intelligent user interfaces will enhance *telepresence* and *presence* in telemedicine environments. For example, when the health-care provider and the telemedicine patient are communicating, the avatar (discussed in Section 2.1.4) mimicking the facial expressions of each user at both sites (Lisetti et al., 2003) will increase the sense of telepresence for these users. Furthermore, during this interaction, when the system accurately recognizes depression or sadness from telemedicine patients and forwards this information to the health-care providers monitoring them, they will be better equipped and ready to respond. This will improve the patients' health and satisfaction, and increase the patients' sense of being monitored in a real medical environment.

These three applications suggest that by enhancing the social and emotional cues in human-computer interactions, users may benefit from improved health and learning as well as increased satisfaction and driving safety. Furthermore, technological interactions that

allow for social presence may increase human acceptance of such systems beyond that of typical cold technologies.

1.4 Dissertation Overview

This dissertation is organized as follows: Background information is given in Chapter 2. Literature research on various research areas that this project involves and combines: our in-house research tool Multimodal Affective User Interface (MAUI) (Section 2.1), emotion elicitation techniques (Section 2.2), emotion recognition from physiological signals (Section 2.3), a specific research on recognizing driver's stress (Section 2.4), and user modeling (Section 2.5) is provided. Chapter 3 describes the experiment conducted to measure and analyze physiological data in order to map them to their corresponding emotions that were elicited with various multi-modal stimuli. Similarly, Chapter 4 discussed the experiments that were conducted to measure and analyze physiological signals while eliciting driving related emotions in a Virtual Driving Environment. How to build models of users and how to adapt the intelligent system accordingly is discussed in Chapter 5. Chapter 6 presents how the complete system is visualized. Chapter 7 discusses the possible ways of improving this research. Finally, Chapter 8 concludes these studies, results, and findings.

CHAPTER 2

BACKGROUND

This research brings together a variety of research methods, which are described next, and will be threaded together in the next chapters for a single goal. The following subsections give brief background information on our in-house research tool: Multimodal Affective User Interface (MAUI), emotion elicitation techniques, emotion recognition from physiology, measuring driver's stress, and user modeling.

2.1 MAUI: A Multimodal Affective User Interface

Figure 2.1 shows the generic Affective Intelligent User Interface designed by Lisetti [Lis99], [BL02], [LN02], and used as our in-house research tool. The following subsections give detailed information about the various functions of the system, such as recording videos and taking snapshots of users' facial expressions for facial expression recognition purposes, giving feedback to users about their emotional states, and interacting with users via facial expres-

sions and vocal intonation.



Figure 2.1: MAUI:Multimodal Affective User Interface.

2.1.1 User Text Input

The text field in the lower right hand corner of Figure 2.1 is for users to communicate with the system via text. Users enter here the information about their own emotional states in their own words. Users' input will be used by natural language understanding algorithms for more accurate recognition of emotions.

2.1.2 Ongoing Video and Captured Image

The first image in the lower left hand corner of Figure 2.1 displays the ongoing video of the user, which is recorded by a camera connected to the user's computer. This video captured during interaction is saved in order to compare the system's interpretation of changes in user's physiological arousal and the changes in her facial expressions over time. The second image displays the still image of user captured at specific times for facial expression recognition purposes.

2.1.3 Feedback to User

In the upper left hand corner of Figure 2.1, the system displays, in a text format, its interpretation of the user's current emotional state (*i.e.*, happy, sad, frustrated, angry, afraid, *etc.*) by indicating the emotion components (*i.e.*, valence, intensity, causal chain, *etc.*) associated with the emotion, facial expression reading, and physiological signals. As mentioned previously, the information about the user's affective state that feeds the system is gathered by physiological measurements of arousal. These data are then interpreted through pattern recognition algorithms, which identify the user's current emotion.

2.1.4 Avatar

The upper right section of the MAUI displays an anthropomorphic avatar that can express emotions via its face. Earlier studies have emphasized that facial expressions are universally expressed and recognized by humans [Ekm89]. In addition, the human face is considered an independent channel of communication that helps to coordinate conversations in human-human interactions [TN93]. In human-computer interactions, research suggests that having an avatar as part of an interface helps to increase human performance. For example, Walker *et al.* [WSS94] reported that subjects in an interview simulation spent more time, made fewer mistakes, and wrote more comments when interacting with an avatar than subjects being interviewed with a text-based interface. In another study, Takeuchi and Nagao [TN93] gave participants ten minutes to ask a series of questions regarding functions and prices of computer products. Individuals interacting with the avatar successfully completed the interaction more often than individuals interacting with a text-based program. Similarly, a study conducted by Koda [Kod96] searched the effect of an agent with face in a virtual poker game environment. The results showed that having a face was found to be likeable, engaging, and comfortable by participants of the study. Moreover, participants rated the intelligence, likeability, and engagingness of Human Face higher than Dog's Face or Smiley. Baylor's study [Bay00] discussed the increase in "motivational qualities" of learners by achieving human resemblance in agents of a learning environment.

Including expressive avatars also has the potential to increase perceptions of presence. For example, Casanueva and Blake [CB01] found that the level of subjective co-presence was higher for subjects with characters displaying gestures and facial expressions than it was for subjects with avatars communicating with a static neutral expression and no gestures. Slater et al. [SHS00] had three separate pairs of actors working with a director to rehearse a short play in remote environments. Each pair of actors rehearsed the same play, each actor and the director were in separate remote environments, and avatar characters were used to represent the facial expressions and gestures of other members in each group. The male avatar was created originally with the Distributed Virtual Environment (DIVE) system and edited to improve the face and hair, while the female avatar was adapted from an H-Anim compliant avatar Nancy by adding facial animation. The participants rehearsed the play four times in the remote environment and then came together to rehearse in a real environment. After rehearsing one time in a real environment, the play was presented to a real audience. The actors' self-reports indicated that their sense of presence and co-presence of the other actor increased over the four remote rehearsal trials. In addition, the evaluations of the directors suggested that it was not possible that the quality of the live performances were a result of only one live rehearsal.

2.2 Emotion Elicitation

Several studies on emotion elicitation have been conducted using various emotion elicitation techniques. One very commonly used technique to elicit emotions is to use movie clips discussed next and the other techniques are provided in Section 2.2.2.

2.2.1 Emotion Elicitation with Movie Clips

Results of Gross and Levenson's [GL95] work were used to guide the design of the study described in Chapter 3. Based on five years of research, the authors reported their findings of the most effective films eliciting discrete emotions. Over 250 movie scenes were in the initial pool. Seventy eight of these films were subject to further evaluation based on the following three requirements: (1) the length of the scene needed to be relatively short, (2) the scene needed to be understood without explanation, and (3) the scene needed to elicit a single emotion. The 78 films were viewed by 31 groups of undergraduates: each film was viewed by a minimum of 25 subjects and each group viewed approximately 10 film clips each. Between each scene, the subjects rated the degree they experienced 16 emotions on an eight-point scale and were then given 20 seconds to "clear their minds of all thoughts, feelings, and memories" (p. 90 [GL95]).

The two movie scenes resulting in the highest subject agreement for eliciting discrete emotions were selected and presented in a table. Agreement rates for these films were:

amusement (84% and 93%), disgust (85% and 80%), sadness (94% and 76%), surprise (75% and 67%), fear (71% and 60%), contentment (58% and 43%), and anger (22% and 42%). Amusement, disgust, and sadness were most successful in producing the target emotion that was intended to be elicited. The more difficult emotions to elicit were anger, contentment, and fear. Upon further analyses, the authors reported that contentment films elicited high degrees of happiness; anger films were affiliated with a host of other emotions, including disgust; and fear films were confounded with tension and interest. The authors concluded that "with films, it appears that there is a natural tendency for anger to co-occur with other negative emotions . . . we are becoming increasingly convinced that elicitation of discrete anger with brief films is going to be extremely difficult, if not impossible" and "perhaps the co-occurrence of fear, tension, and interest is a natural one" (p. 104). However, if the goal is to elicit one emotion more intensely than others, films are a viable choice.

2.2.2 Other Emotion Elicitation Techniques

Using movie clips is not the only technique used for emotion elicitation from participants. Several other emotion elicitation techniques were used while conducting psychological experiments. These techniques include:

- Making the participant imagine scenarios where she experiences the specified emotions [Vra93], [SP96], [PHV01].

- Assigning tasks or problems at different levels of difficulty to the participant [WCP86], [PS96].
- Showing the participant emotionally loaded pictures [CVD97].
- Making the participant interact with slow game interfaces [SFK02], [MPM03].
- Instructing the participant to make facial expressions [ELF83], [LHE92].
- Assigning driving tasks to the participant [Hea00].

2.3 Emotion Recognition from Physiological Signals

There are several studies conducted on understanding the connection between emotions and physiological arousal. Manual analyses have been successfully used for this purpose [ELF83], [GL97]. However, interpreting the data with statistical methods and algorithms is beneficial in terms of actually being able to map them to specific emotions. Studies have demonstrated that algorithms can be very successfully implemented for recognition of emotions from physiological signals.

Collet *et al.* [CVD97] showed neutral and emotionally loaded pictures to participants in order to elicit happiness, surprise, anger, fear, sadness, and disgust. The physiological signals measured were: Skin conductance (SC), skin potential (SP), skin resistance (SR), skin blood flow (SBF), skin temperature (ST), and Instantaneous respiratory frequency

(IRF). Statistical comparison of data signals was performed pair-wise, where 6 emotions formed 15 pairs. Out of these 15 emotion-pairs, electrodermal responses (SR, SC, and SP) distinguished 13 pairs, and similarly combination of thermo-circulatory variables (SBF and ST) and Respiration could distinguish 14 emotion pairs successfully.

Picard *et al.* [PHV01] showed pictures eliciting happiness, sadness, anger, fear, disgust, surprise, neutrality, platonic love, and romantic love. The physiological signals measured were GSR, heartbeat, respiration, and electrocardiogram. The algorithms used to analyze the data were Sequential Forward Floating Selection (SFFS), Fisher Projection, and a hybrid of these two. The best classification achievement was gained by the hybrid method, which resulted in 81% overall accuracy.

Table 2.1 in summarizes results of studies investigating the relationship between emotions and physiological arousal using other statistical procedures such as ANOVA and Hidden Markov Models. All these studies succeeded in finding a pattern of physiological signals for each of the emotions elicited. In summary, the results of these studies suggest that physiological patterns can successfully be identified using statistical procedures.

Year	Emotion Elicitation Method	Emotions Elicited	Subjects	Signals Measured	Data Analyze Technique	Results
1974 [SFG74]	Personalized Imagery	Happiness and sadness, and anger	20 people in 1st study, 12 people in 2nd study	Facial EMG	Manual analysis	EMG reliably discriminated between all four conditions when no overt facial differences were apparent
1983 [ELF83]	Facial action task, relived emotion task.	Anger, fear, sadness, disgust, and happiness	12 professional actors and 4 scientists	Finger temperature, heart rate, and skin conductance	Manual analysis	Anger, fear and sadness produce a larger increase in heart rate than disgust. Anger produces a larger increase in finger temperature than fear. Anger and fear produces larger hear rate than happiness and fear and disgust produces larger skin conductance than happiness.
1986 [LO86]	Vocal tone, slide of facial expressions, electric shock	Happiness and fear	60 undergraduate students (23 female and 37) male	Skin conductance (galvanic skin response)	ANOVA	Fear produced a higher level of tonic arousal and larger phasic skin conductance
1986 [VCL86]	Imagining and silently repeating fearful and neutral sentences	Neutral and fear	64 introductory psychology students	Heart rate, self report	ANOVA Newman-Keuls pair-wise comparison	Heart rate acceleration was more during fear imagery than neutral imagery or silent repetition of neutral sentences or fearful sentences.

1986 [WCP86]	Easy, moderately and extremely difficult memory task	Difficult Problem Solving	64 undergraduate females from Stony Brook	Heart rate, systolic and diastolic blood pressure	ANOVA	Both SBP and goal attractiveness were non-monotonically related to expected task difficulty.
1989 [Smi89]	Personalized imagery	Pleasant emotional experiences (low-effort vs. high effort, and self-agency vs. other-agency)	96 Stanford University Undergraduates (48 female, 48 male)	Facial EMG, heart rate, skin conductance, and self-report	ANOVA and Regression	Eyebrow frown and smile are associated with evaluations along pleasantness dimension, HR measure offered strong support between anticipated effort and arousal. Skin conductance offers further support for that but not as strong as HR.
1990 [LEF90]	Contracting facial muscles into facial expressions	Anger and fear.	12 actors (6 female, 6 male) and 4 researchers (1 female, 3 male)	Finger temperature.	Manual analysis.	Anger increases temperature, fear decreases temperature.
1992 [LHE92]	Contracting facial muscles into prototypical configurations of emotions	Happiness, sadness, disgust, fear, and anger	46 Minangkabau men	Heart rate, finger temperature, finger pulse transmission and finger pulse amplitude, and respiratory period and respiratory depth	MANOVA	HR- Anger, fear, and sadness were associated significantly larger than disgust. Happiness was intermediate.

1993 [Vra93]	Imagery	Disgust, anger, pleasure, and joy.	50 people (25 male, 25 female)	Self-reports, Heart rate, skin conductance, facial EMG	ANOVA	Acceleration of heart rate was greater during disgust, joy, and anger imageries than during pleasant imagery. Disgust could be discriminated from anger using facial EMG.
1993 [WD93]	Difficult Task Solving	Difficult Task Solving	58 undergraduate students of an introductory psychology course	Cardiovascular activity (heart rate and blood pressure)	ANOVA and ANCOVA	Systolic and diastolic blood pressure responses were greater in the difficult standard condition than in the easy standard condition for the subjects who received high-ability feedback, however it was the opposite for the subjects who received low-ability feedback.
1996 [PS96]	Difficult Problem Solving	Difficult Problem Solving	32 university undergraduates (16 male, 16 female)	Skin conductance, Self-report Objective task performance	ANOVA, MANOVA Correlation/regressionals analyses	Within trials, skin conductance increased at the beginning of the trial, but decreased by the end of the trial. Correlation/regressionals for the most difficult condition.
1996 [SP96]	Imagery script development	Neutral, fear, joy, action, sadness, and anger	27 right-handed male between ages 21-35	Heart rate, skin conductance, finger temperature, blood pressure, electro-oculogram, facial electromyograms	Discriminant Function Analyses, ANOVA	99% correct classification was obtained. This indicates that emotion-specific response patterns for fear and anger are accurately differentiable from each other and from neutral.

1997 [CVD97]	Neutrally and emotionally loaded slides (pictures)	Happiness, surprise, anger, fear, sadness, and disgust	30 people (16 female and 14 male)	Skin conductance, skin potential, skin resistance, skin blood flow, skin temperature, and instantaneous respiratory frequency	Friedman variance analyses	Electrodermal responses distinguished 13 emotion-pairs out of 15. Skin resistance and skin conductance ohmic perturbation duration indices separated 10 emotion-pairs. However, conductance amplitude could distinguish 7 emotion-pairs.
1997 [GL97]	Film showing.	Amusement, neutral, and sadness.	180 females	Skin conductance, inter-beat interval, pulse transit times and respiratory activation.	Manual analysis.	Inter-beat interval increased for all three states, but for neutral it was less than the amusement and sadness. Skin conductance increased after the amusement film, decreased after the neutral film and stayed same after the sadness film.
1999 [ADL99]	Subjects were instructed to make facial expressions	Happiness, sadness, anger, fear, disgust, surprise.	6 people (3 females and 3 males).	Heart rate, general somatic activity, GSR and temperature.	Discriminant function analyses	66% accuracy in classifying emotions.
2000 [PSA00]	Unpleasant and neutral film clips	Fear, disgust, anger, surprise, and happiness.	46 undergraduate students (31 female, 15 male)	Self-report, electrocardiogram, heart rate, T-wave amplitude, respiratory sinus arrhythmia, and skin conductance	ANOVA, Greenhouse-Geisser Correction. Post hoc means comparisons and simple effects analyses.	Films containing violent threats, increased sympathetic activation, whereas the surgery film increased the electrodermal activation, decreased the heart rate and increased the T-wave.

2001 [TKK01]	11 auditory mixed with standard and target sounds.	Surprise	20 healthy controls (as a control group) and 13 psychotic patients	GSR	Principal Component Analysis clustered by centroid method.	78% accuracy for all, 100% accuracy for patients.
2001 [CAA01]	Arithmetic tasks, video games. Showing faces, expressing specific emotions.	Attention, concentration. Emotions of happiness, sadness, anger, fear, disgust, surprise and neutral.	10 to 20 college students.	GSR, heart rate and skin temperature.	Manual analysis.	No recognition found, some observations only.
2001 [PHV01]	Personal imagery	Happiness, sadness, anger, fear, disgust, surprise, neutral, pleasant, romantic love.	A healthy graduate student with two years of acting experience.	GSR, heart rate, ECG and respiration.	Sequential floating forward search (SFSS), Fisher Projection (FP) and hybrid (SFSS and FP).	81% for by hybrid SFSS and Fisher method with 40 features 54% rate with 24 features.

2002 [SFK02]	A slow computer game interface.	Frustration	36 undergraduate and graduate students	Skin conductivity and blood volume pressure	Hidden Markov Models	Pattern recognition worked significantly better than random guessing while discriminating between regimes of likely frustration from regimes of much less likely frustration.
-----------------	---------------------------------	-------------	--	---	----------------------	---

Table 2.1: Previous Research on Emotion Recognition from Physiological Signals

2.4 A Previous Study on Measuring Drivers' Stress

Jennifer Healey's research from Massachusetts Institute of Technology (MIT) Media Lab [Hea00] was focused on recognizing stress levels of drivers by measuring and analyzing their physiological signals. The study answered the questions about how affective models of users should be developed for computer systems and how computers should respond to the emotional states of users appropriately. The results showed that people don't just create preference lists, but they use affective expression to communicate and to show their satisfaction or dissatisfaction.

Before the driving experiment was conducted a preliminary emotion elicitation experiment was designed where eight states (anger, hate, grief, love, romantic love, joy, reverence, and no emotion *i.e.* neutrality) were elicited from the participants. These eight emotions were Clynes' [Cly77] emotion set for basic emotions. This set of emotions was chosen to be elicited in the experiment because each emotion in this set was found to produce a unique set of finger pressure patterns [Cly77]. While the participants were experiencing these emotions the changes in their physiological responses were measured. Table 2.2 shows the equipment used to measure various physiological signals:

The experiment was conducted over 32 days in a single subject-multiple sessions setup. However only twenty sets (days) of complete data were obtained at the end of the experiment. Guided imagery technique (*i.e.*, the participant imagines that she is experiencing the emotion by picturing herself in a certain given scenario) was used to generate the emotions listed

Table 2.2: Sensors Used to Measure the Physiological Signals

Physiological Signal	Sensor
Skin Conductance	Skin Conductance Sensor
Heart Activity	Blood Volume Pressure (BVP) Sensor Electrocardiograph (EKG)
Respiration	Respiration Sensor
Muscle Activity	Electromyogram (EMG)
Finger Pressure	Sentograph

above. The participant attempted to feel and express eight emotions for a varying period of three to five minutes (with random variations). At the end of the experiment, the participant reported the arousal level and the valence of each emotion she experienced as follows:

- **Neutral:** Low arousal, neutral valence
- **Anger:** Highest arousal of all eight emotions, constantly negative valence
- **Hate:** Low arousal and negative valence.
- **Grief:** High arousal and negative valence
- **Platonic love:** Low arousal and positive valence
- **Romantic love:** High arousal and positive valence
- **Joy:** High arousal and positive valence
- **Reverence:** Low arousal and neutral valence

Eleven features were extracted from raw EMG, GSR, BVP, and Respiration measurements by calculating the mean, the normalized mean, the normalized first difference mean,

and the first forward distance mean of the physiological signals. Eleven dimensional feature space of 160 emotions was projected into a two dimensional space by using Fisher projection. Leave-one-out cross validation was used for emotion classification. The results showed that it was hard to discriminate all eight emotion states. However when the emotions were grouped as being 1) anger or peaceful, 2) high arousal or low arousal, and 3) positive valence or negative valence, they could be classified successfully as follows:

- Anger: 100%, Peaceful: 98%
- High arousal: 80%, Low arousal: 88%
- Positive: 82%, Negative: 50%

Another preliminary experiment conducted was the daily monitoring of a participant while she was performing normal activities. The aim of this experiment was to determine how feasible the ambulatory (*i.e.* not stationary; movable) affect detection would be. EMG, BVP, respiration, and skin conductance sensors were used to measure physiological signals while the participant went through her daily activities and made annotation at certain times. The anecdotal results showed that emotion recognition is hard in ambulatory environments because of the motion artifacts and difficulty of capturing and coding physical and emotional events. For example EMG indicated more muscle activity in the morning than later in the day. It might be due to an emotional episode, but it was most likely due to a motor activity: the participant was carrying the wearable computer on the left shoulder in the morning, and

then she carried it on her lap later in the day. These findings influenced the design of the final driving experiment.

Because of the results of the experiments described above, it was found that it is difficult to perform emotion recognition during natural situations. So, the scope of the driving experiment was limited to recognition of levels of only one emotional state: emotional stress.

In the beginning of the experiment, participants drove in and exited a parking garage, then they drove in a city and on a highway, and returned to the same parking garage at the end. The experiment used three subjects who repeated the experiment multiple times and six subjects who drove only once. Videos of the participants were recorded during the experiments and self-reports were obtained at the end of each session. Task design and questionnaire responses were used to recognize the driver's stress separately. The results obtained from these two methods were as follows:

- **Task design analysis** could recognize driver stress level as being rest (*e.g.* resting in the parking garage), city (*e.g.* driving in Boston streets), or highway (*e.g.* two lane merge on the highway) with 96% accuracy.
- **Questionnaire analysis** could categorize four stress classes as being lowest, low, higher, or highest with 88.6% accuracy.

Finally, video recordings were annotated on a second by second basis by two independent researchers for validation purposes. This annotation was used to find a correlation between

stress metric created from the video and variables from the sensors. The results showed that physiological signals closely followed the stress metric provided by the video coders.

The results of these three methods coincided in classifying the driver's stress and showed that stress levels could be recognized by measuring the physiological signals and analyzing them by pattern recognition algorithms. Similarities and differences between this discussed approach and this dissertation's approach are discussed in Chapter 4

2.5 User Modeling

In this section relevant background on User Modeling, a subfield of Computer Science that enables computer systems to record relevant user information is given. A User Model is defined as "the description and knowledge of the user maintained by the system" [NS89]. An adaptive system should modify the user model as the individual user changes, because each user differs from other users and the same specific user may change while s/he is interacting with the system [NS89].

2.5.1 User Modeling of Affect

In the recent years, there has been a significant increase in the number of attempts to build user models that include affect at some level in the user model. Carofiglio and de

Rosis' ongoing project [CR03] discussed the effect of emotions on an argumentation in a dialog, and modeled how emotions were activated and how the argumentation was adapted accordingly using Belief Networks. In the model, the user interacting with the system's *agent* is represented as the *receiver*. For example when a positive emotion "hope" is activated, the intensity of this emotion is influenced by the receiver's belief that a specific event will occur to self in the future, the event that the belief is desirable, and the belief that this situation favors achieving the agent's goal. The model represents both logical and emotional reasoning in the same structure. The model evaluated every candidate argument applying simulative reasoning by focusing on the current state of the dialog and by guessing the effect of the candidate on the receiver of argument. The difficulties encountered in some cases when classifying the argumentation as being 'rational' or 'emotional' increased the author's belief in the importance of including emotional factors while creating natural argumentation systems.

Affect and Belief Adaptive Interface System (ABAIS), created as a rule-based system by Hudlicka and McNeese [HM02], assessed pilots' affective states and active beliefs and took adaptive precautions to compensate for their negative affects. The architecture of this system was built on i) sensing the user's affective state and beliefs (User State Assessment Module), ii) inferring the potential effects of them on the user's performance (Impact Prediction Module), iii) selecting a strategy to compensate, and finally (Strategy Selection Module) iv) implementing this strategy (Graphical User Interface[GUI]/Decision Support System[DSS] Adaptation Module). Sensing the user's affective state and beliefs is defined as

the most critical of the whole structure and it receives various data about the current user to identify his current affective state (e.g. high anxiety) and his beliefs (e.g. hostile aircraft approaching). The potential effects (generic effects and the task-specific effects) (e.g. task neglect) of the user's emotions and beliefs are inferred in the Impact Prediction Module using a rule-based reasoning. Rule-based reasoning is also used when selecting a counter strategy (e.g. present reminder of neglected tasks) to compensate the negative effects on the user's performance and finally the selected strategy is implemented in the GUI/DSS Module by using a rule-based reasoning to consider the individual pilot preferences.

Conati's [Con03] probabilistic user model, which was based on Dynamic Decision Networks (DDN), was built to represent the emotional state of users while interacting with an educational game by also including causes and effects of the emotion, as well as user's personality and goals. Probabilistic dependencies between causes, effects, and emotional states and their temporal evolution are represented by DDNs. DDNs were used due to their capability of modeling uncertain knowledge and environments to change over time. Assessing users' emotional states is a very important component in both Hudlicka and McNeese's [HM02] and Conati's [Con03] models; however, none of these studies actually perform emotion recognition. Objective of the work described in this dissertation is to perform both emotion recognition and appropriate interface adaptation and combine them in the same system.

Guinn and Hubal [GH03] created a system, in which the user's spoken language was tagged and the response was given according to the politeness, anger, complexity, confusion,

levels of the user's input. This system is currently used as a law enforcement trainer for managing encounters with the mentally ill and also as a telephone survey interviewer trainer.

The model described in this dissertation combines both the user's model of emotions and personality, and the emotion recognition process that enable us to actually activate the adaptation strategies as discussed in Chapter 5.

2.5.2 Other User Modeling Techniques

Conventional user models are built on what the user knows or does not know about the specific context, what her/his skills and goals are, and her/his self-report about what s/he likes or dislikes.

In Millan and Perez-de-la-Cruz's [MC02] student model, the knowledge level of the student was represented with various variables in an integrated Bayesian Network. Variables and nodes were used to represent student's declarative knowledge, skills, and abilities and to collect information about the student's knowledge state.

In Barker *et al.*'s ([BJB02]) student model, five different learner characteristics were used to build the model. These were 1) language level and language support, 2) cognitive style differentiation, 3) task level and task type differentiation 4) question level and question type differentiation, and 5) help and support system differentiation. These characteristics were modeled in a simple way: assigning initial values to each characteristic in order to define its

level and using these values to present a different version of an application to the learner. The values of characteristics were re-assigned as a result of learner's interaction with the application using a rule-based adaptation.

Billsus and Pazzani [BP00] introduced a system called Daily Learner, for adaptive news access, and discussed how user modeling was built by implementing machine learning techniques. The system was designed to regularly download news stories from the Internet and store them, let the user access these news stories, store the user's feedback, and recommend stories to the user according to her/his personal interests. Short term and long term interests of the user were modeled with the Nearest Neighbor Algorithm and Naive Bayesian Classifier respectively.

Although the user models that are discussed in this section do not model affective states of the users *per se*, they are important in terms of implementing various techniques to model different characteristics of the users and adapt to these.

CHAPTER 3

EMOTION ELICITATION EXPERIMENTS

This chapter describes the experiments used in the preliminary study and presents its results and findings.

3.1 Pilot Panel Study

Before collecting physiological data, a pilot panel study with movie scenes resulting in high subject agreement from Gross and Levenson's [GL95] work was conducted. Because some of their movies were not obtainable and anger and fear movie scenes evidenced low subject agreement, alternative clips were also investigated. The purpose of the panel study was to determine the movies that may result in high subject agreement for the subsequent study, which will be described shortly. The following sections describe the panel study and results.

Sample: The sample included 14 undergraduate and graduate students from the psychology and computer science departments of University of Central Florida. There were 7 females and 7 males: 10 Caucasians, 1 Hispanic American, 1 African American, and 2

Asians. Their ages ranged from 18-35. Specific ages were not requested, therefore a mean age was not calculated.

Movie clips: Emotions were elicited using scenes from 21 movies. Seven movies were included in the analysis based on the findings of Gross and Levenson [GL95] (see Table 3.1). An additional 14 movie clips were found by the authors. The final sample included 4 movies targeted to elicit anger (Eye for an Eye, Schindler's List, American History, and My Bodyguard), 3 movie clips to elicit sadness (Powder, Bambi, and The Champ), 4 to elicit amusement (Beverly Hillbillies, When Harry Met Sally, Drop Dead Fred, and The Great Dictator), 1 to elicit disgust (Fear Factor), 5 to elicit fear (Jeepers Creepers, Speed, The Shining, Hannibal, and Silence of the Lambs), and 4 to elicit surprise (Jurassic Park, The Hitcher, Capricorn One, and a homemade clip called Grandma).

Procedure: The 14 subjects participated as a group simultaneously. Once consent forms were completed, the participants were given questionnaires and asked to answer the demographic items before beginning the study. The subjects were then informed that they would be watching scenes from various movies geared to elicit emotions. They were also told that between each movie, they would be prompted to answer questions about the emotions they felt as a result of the scene. Lastly, they were asked to respond according to the emotions they experienced and not the emotions displayed by the actors. A computerized slide show played the scenes and, after each of the 21 clips, a slide was presented asking the participants to answer the survey items for the prior scene.

Table 3.1: Movies from Gross and Levenson ([GL95])

Emotion	Movie	N	Agreement	Mean Intensity
Sadness	Bambi	72	76%	5.35
	The Champ	52	94%	5.71
Amusement	When Harry Met Sally	72	93%	5.54
Fear	The Shining	59	71%	4.08
	Silence of the Lambs	72	60%	4.24
Anger	My Bodyguard	72	42%	5.22
Surprise	Capricorn One	63	75%	5.05

Measures: The questionnaire included three demographic questions: age ranges (18-25, 26-35, 36-45, 46-55, or 56+), gender, and ethnicity. For each scene, 4 questions were asked. The first question asked, what emotion they experienced from the video clip they viewed, and provided 7 options (anger, frustration, amusement, fear, surprise, sadness, and other). If the participant checked "other" they were asked to specify which emotion they felt. The second question asked the participants to rate the intensity of the emotion they felt on a 6 point scale. The third question asked participants if they had experienced any other emotions at the same intensity or higher, and if so, to specify what that emotion was. The final question asked participants if they had seen the movie in the past.

Results: The goal of the pilot study was to find the movie scenes that resulted in (a) 90% agreement or higher on the target emotion and (b) 3.5 or higher average intensity. Table 3 lists the hit rates and average intensities for the clips with $\geq 90\%$ agreement. There was not a movie with a high level of agreement for anger. With intensity in mind, Gross and Levenson's [GL95] clips were most successful at eliciting the emotions in the investigation, except for anger. In their study, the movie with the highest hit rate for anger was My

Table 3.2: Hit Rates and Average Intensities for Movies with $\geq 90\%$ Agreement

Emotion	Movie	Agreement	Mean Intensity	SD
Sadness	Powder	93%	3.46	1.03
	Bambi	100%	4.00	1.66
	The Champ	100%	4.36	1.60
Amusement	Beverly Hillbillies	93%	3.69	1.13
	When Harry Met Sally	100%	5.00	0.96
	Drop Dead Fred	100%	4.00	1.21
	Great Dictator	100%	3.07	1.14
Fear	Shining	93%	3.62	0.96
Surprise	Capricorn One	100%	4.79	1.25

Table 3.3: Movie Scenes Selected for the Study

Emotion	Movie	Scene
Sadness	The Champ	death of the Champ
Anger	Schindler's List	woman engineer being shot
Amusement	Drop Dead Fred	restaurant scene
Fear	Shining	boy playing in hallway
Surprise	Capricorn One	agents burst through the door

Bodyguard (42%). In the pilot study, the hit rate was 29% with a higher hit rate for frustration (36%). However, because anger is an emotion of interest for future research with driving simulators, the movie with the highest hit rate, Schindler's List (hit rate was 36%, average intensity was 5.00) was included. In addition, for amusement, the movie Drop Dead Fred was chosen to replace When Harry Met Sally due to the embarrassment experienced by some of the subjects when viewing the latter. The final set of movie scenes chosen for the study is presented in Table 3.3.

3.2 Emotional Signals Data Generation and Collection

Sample: The final sample included 29 undergraduate students enrolled in a computer science course. There were 3 females and 26 males: 21 Caucasians, 1 African American, 1 Asian American and 6 individuals who did not report their ethnicity. Their ages ranged from 18-40 (19 individuals were 18-25 and 10 were 26-40). Specific ages were not requested; therefore, a mean age was not calculated.

Procedure: One to three subjects participated in the study during each session. After signing consent forms, non-invasive SenseWear armband (see Figure 1.2a) was placed on each subject's right arm to collect galvanic skin response (GSR), heart rate, and temperature data. As the participants waited for the armband to detect their physiological signals, they were asked to complete a pre-study questionnaire. Once the armband signaled that it was ready, the subjects were instructed on how to place the chest strap. After the chest straps were activated, the in-study questionnaire was placed face down in front of the subjects and the participants were told the following: (i) to find a comfortable sitting position and try not to move around until answering a questionnaire item, (ii) that the slide show would instruct them to answer specific items on the questionnaire, (iii) to please not look ahead at the questions, and (iv) that someone would sit behind them at the beginning of the study to activate the armband.

A 45-minute computerized slide show was then activated. The study began with a slide asking the subjects to relax, breathe through their nose, and listen to soothing music. Slides

of natural scenes were presented, including pictures of the ocean, mountains, trees, sunset, and butterflies. These slides were presented for 6 seconds each. After 2.5 minutes, the first movie clip played (sadness). Once the clip was over, the next slide asked the participants to answer the questions relevant to the scene they watched. This slide stayed on screen for 45 seconds. Starting again with the slide asking the subjects to relax while listening to soothing music, this process continued for the anger, fear, surprise, frustration, and amusement clips. The frustration segment of the slide show asked the participants to answer analytical math problems without paper and pencil. The movie scenes and frustration exercise lasted from 70 to 231 seconds each. After the slide show ended, the participants were asked to remove their chest straps first and then the armbands. The in-study questionnaires were collected and the subjects were asked if they had any questions or comments.

Measures: The pre-questionnaire included three demographic questions: age ranges (18-25, 26-35, 36-45, 46-55, or 56+), gender, and ethnicity. The in-study questionnaire included 3 questions for each emotion. The first asked the participants if they experienced sadness (or the relevant emotion) during this section of the experiment, and required a yes or no response. The second question asked the participants to rate the intensity of the emotion they felt on a 6-point scale. The third question asked participants if they had experienced any other emotions at the same intensity or higher, and if so, to specify what that emotion was. Finally, the physiological data gathered included heart rate, skin temperature, and GSR.

Self-report: Table 3.4 reports subject agreement and average intensities for each movie scene and the math problems. A two sample binomial test of equal proportions was conducted to determine whether the agreement rates for the panel study differed from the results obtained with this sample. Participants in the panel study agreed significantly more to the target emotion for the sadness and fear films. On the other hand, the subjects in this sample agreed more for the anger film. This lack of reliability in subject agreement across studies may be due to small sample sizes.

Table 3.4: Hit Rates and Average Intensities

Emotion	Movie	N	Agreement	Mean Intensity	SD
Sadness	The Champ	27	56%	3.53	1.06
Anger	Schindler's List	24	75%	3.94	1.30
Fear	Shining	23	65%	3.58	1.61
Surprise	Capricorn One	21	90%	2.73	1.28
Frustration	Math Problems	22	73%	3.69	1.35
Amusement	Drop Dead Fred	23	100%	4.26	1.10

3.3 Pattern Recognition of Physiological Signals with Machine Learning

After determining the time slots corresponding to the point in the film where the intended emotion was most likely to be felt, the procedures described above resulted in the following set of physiological records: 24 for anger, 23 for fear, 27 for sadness, 23 for amusement, 22 for frustration, and 21 for surprise. Data was stored in a three dimensional array of real

numbers. The three dimensions are (1) the subjects who participated in the experiment, (2) the emotion classes (sadness, anger, surprise, fear, frustration, and amusement), and (3) the data signal types (GSR, temperature, and heart rate).

Each slot of the array consists of the normalized average value of one specific data signal belonging to one specific participant while s/he was experiencing one specific emotion, (*e.g.* a slot contains the normalized average skin temperature value of participant #1 while s/he was experiencing anger). The data was normalized for each emotion in order to calculate how much the physiological responses changed as the participants go from a relaxed state to the state of experiencing a particular emotion. Normalization is also important for minimizing the individual differences of participants in terms of the physiological responses they give while experiencing a specific emotion.

The values of each data type were normalized by using the average value of corresponding data collected during the relaxation period for the same participant. For example, equation 1 shows how the GSR values were normalized:

$$normalized_GSR = \left(\frac{original_GSR - relaxation_GSR}{relaxation_GSR} \right) \quad (1)$$

After storing the normalized data in the three dimensional array, three algorithms were implemented to analyze it. These three algorithms described in the following sections are: k-Nearest Neighbor Algorithm (KNN), Discriminant Function Analysis (DFA), and Marquardt Backpropagation (MBP).

3.3.1 k -Nearest Neighbor Algorithm

k -Nearest Neighbor Algorithm (KNN) [Mit97] uses two data sets: (1) training data set (to learn the patterns) and (2) test data set (to verify the validity of learned patterns). In this case, training data set contains instances of GSR, skin temperature, and heart rate values and the corresponding emotion class. Test data set is similar to the training data set, except that it does not have the emotion information. In order to classify an instance of a test data into an emotion, KNN calculates the distance between the test data and each instance of training data set: Let an arbitrary instance x be described by the feature vector $\langle a_1(x), a_2(x), \dots, a_n(x) \rangle$, where $a_r(x)$ is the r^{th} feature of instance x . The distance between instances x_i and x_j is defined as ,

$$d(x_i, x_j) = \sqrt{\sum_n (a_r(x_i) - a_r(x_j))^2} \quad (2)$$

The algorithm then finds the k closest training instances to the test instance. The emotion with the highest frequency among k emotions associated with these k training instances is the emotion mapped to the test data.

The KNN Algorithm was the first to be implemented for this research's emotion recognition purposes. KNN was chosen to be implemented to test the feasibility of performing pattern recognition on physiological signals that are associated with emotions.

As shown in Figure 3.3.1, with KNN algorithm the recognition accuracy obtained was: 67% for sadness, 67% for anger, 67% for surprise, 87% for fear, 72% for frustration, and finally 70% for amusement.

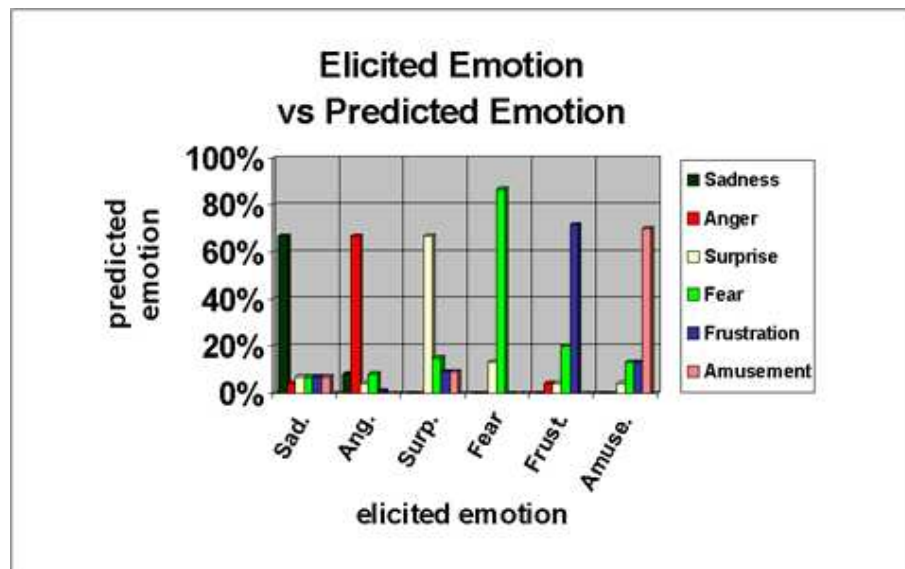


Figure 3.1: Emotion Recognition Results with KNN Algorithm

3.3.2 Discriminant Function Analysis

The second algorithm was developed using Discriminant Function Analysis (DFA) [Nic99], which is a statistical method to classify data signals by using linear discriminant functions. Discriminant function analysis is used to find a set of linear combinations of the variables, whose values are as close as possible within groups and as far apart as possible between

groups. These linear combinations are called discriminant functions. Thus, a discriminant function is a linear combination of the discriminating variables. In this study's implication of discriminant analysis, the groups are the emotion classes and the discriminant variables are the data signals GSR, skin temperature, and heart rate. The number of discriminant functions needed was determined by finding the minimum of (1) number of groups and (2) number discriminant variables. There were six groups (emotions: sadness, anger, surprise, fear, frustration, and amusement) and three (data signal types: GSR, skin temperature, and heart rate). Therefore, number of functions needed was three.

Let x_{data} be the average value of a specific data signal. The functions used to solve the coefficients was,

$$f_i(x_{gsr}, x_{temp}, x_{hr}) = u_0 + u_1 * x_{gsr}, u_2 * x_{temp}, u_3 * x_{hr} \quad (3)$$

The objective of DFA is to calculate the values of the coefficients u_0 , u_1 , u_2 , and u_3 in order to obtain the linear combination. In order to solve for these coefficients, we applied the generalized Eigenvalue decomposition to the between-group and within-group covariance matrices. The vectors gained as a result of this decomposition were used to derive coefficients of the discriminant functions. The coefficients of each function were derived in order to get a maximized difference between the outputs of group means and a minimized difference within the group means. Every function was orthogonal to each other. DFA was chosen to

be implemented since it was successfully used in several physiological signal analyses studies [FI83], [SP96] before.

As can be seen in Figure 3.2, the results of the DFA algorithm demonstrated a similar pattern of accuracy across emotions to that of the KNN algorithm. The DFA algorithm successfully recognized sadness (78%), anger (72%), surprise (71%), fear (83%), frustration (68%), and amusement (74%).

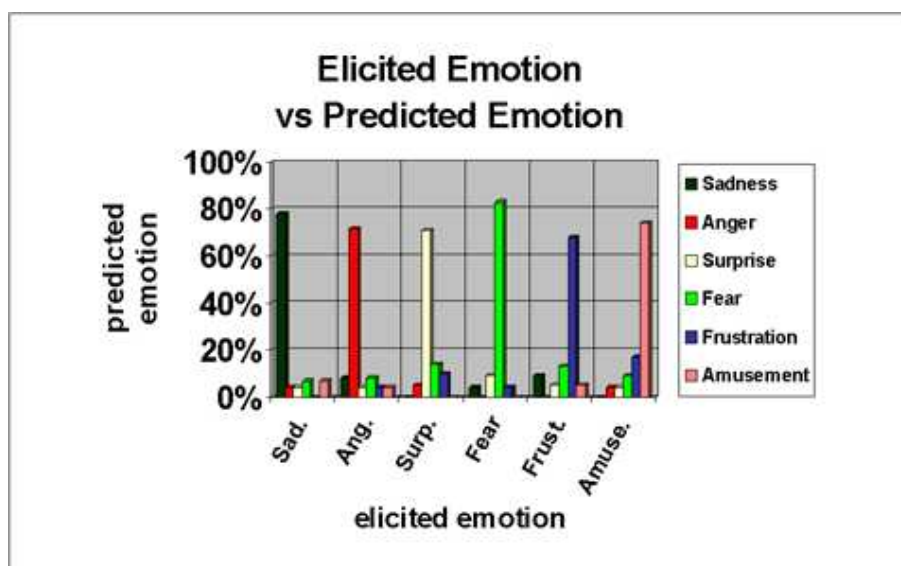


Figure 3.2: Emotion Recognition Results with DFA Algorithm

3.3.3 Marquardt Backpropagation

The third algorithm used was a derivation of a back-propagation algorithm with Marquardt-Levenberg modification called Marquardt Backpropagation (MBP) [HM94]. In this tech-

nique, first the Jacobian matrix, which contains first derivatives of the network errors with respect to the weights and biases, is computed. Then the gradient vector (is computed as a product of the Jacobian matrix $J(\underline{x})$ and the vector of errors $e(\underline{x})$ and the Hessian approximation is computed as the product of the Jacobian matrix $J(\underline{x})$ and the transpose of the Jacobian matrix $JT(\underline{x})$ [HM94].

Then the Marquardt-Levenberg Modification to the Gauss-Newton method is given by the following equation:

$$\Delta \underline{x} = [J^T(\underline{x}) * J(\underline{x}) + \mu * I]^{-1} J^T(\underline{x}) e(\underline{x}) \quad (4)$$

When μ is 0 or a small value, then this is Gauss-Newton's method that is using the Hessian approximation. When μ is a large value, then this equation is a gradient descent with a small step size ($1/\mu$). The aim is to make μ converge to 0 as fast as possible, and this is achieved by decreasing μ when there is a decrease in the error function and increasing it when there is no decrease in the error function. And the algorithm converges when gradient value reaches below a previously determined value [HM94]. This algorithm was chosen to be implemented due to its fast converging nature.

Figure 3.3 shows the implemented Neural Network architecture with three input and six output nodes.

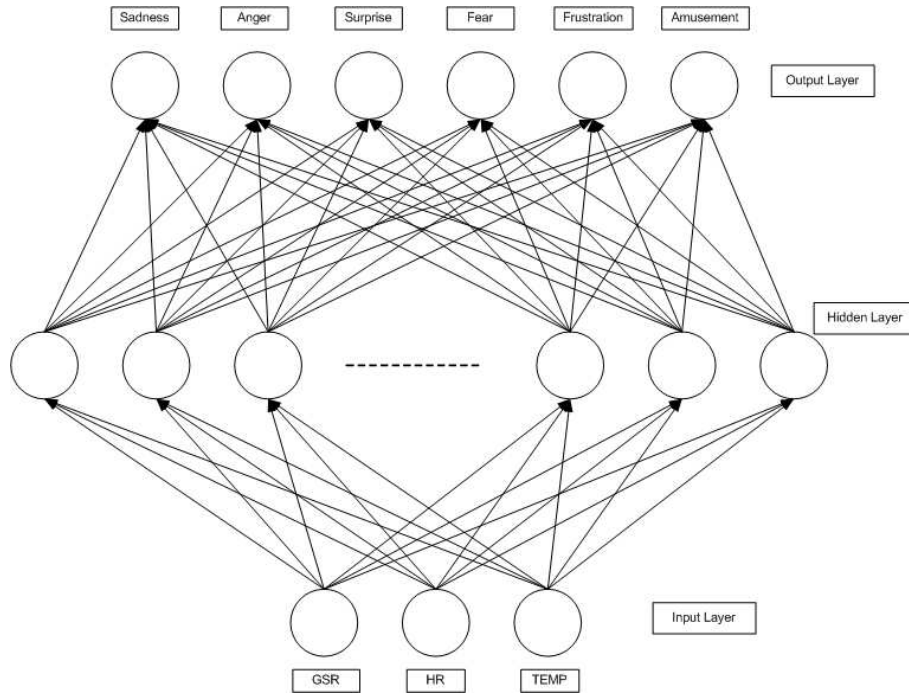


Figure 3.3: Neural Network Architecture

As shown in Figure 3.4 the recognition accuracy gained with MBP algorithm was: 92% for sadness, 88% for anger, 70% for surprise, 87% for fear, 82% for frustration, and 83% for amusement.

Overall, the DFA algorithm was better than the KNN algorithm for sadness, anger, surprise, and amusement. On the other hand, KNN performed better for frustration and fear. MBP Algorithm performed better than both DFA and KNN for all emotion classes except for surprise.

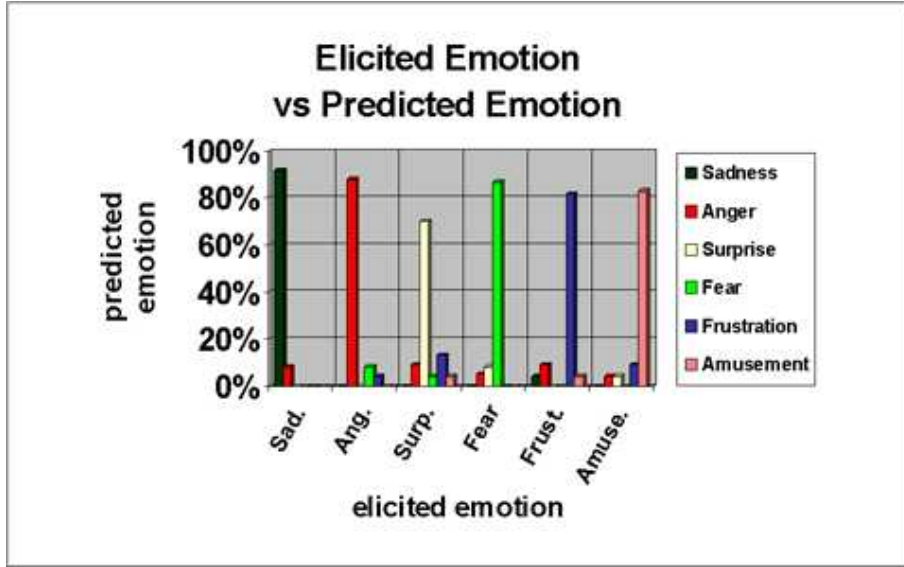


Figure 3.4: Emotion Recognition Results with MBP Algorithm

3.4 Further Evaluation with Feature Extraction

Data collected during the experiment discussed in this chapter was also analyzed by applying a different feature extraction technique. After data signals were normalized in the same way stated in Section 3.3, features were extracted from the normalized data. Four features were extracted for each data signal type: *minimum*, *maximum*, *mean*, and *variance* of the normalized data. Data was stored in a three dimensional array of real numbers: (1) the subjects who participated in the experiment, (2) the emotion classes (sadness, anger, surprise, fear, frustration, and amusement) and (3) extracted features of data signal types (minimum, maximum, mean, and variance of GSR, temperature, and heart rate).

Each slot of the array consists of one specific feature of a specific data signal type, belonging to one specific participant while s/he was experiencing one specific emotion. (e.g. a slot contains the mean of normalized skin temperature value of, say, participant #1 while s/he was experiencing anger, while another slot, for example, contains the variance of normalized GSR value of participant #5 while s/he was experiencing sadness). As mentioned, four features were extracted for each data type and then they were analyzed with the three supervised learning algorithms discussed in Section 3.3. Figure 4.7 shows how emotion recognition is performed by extracting more features.

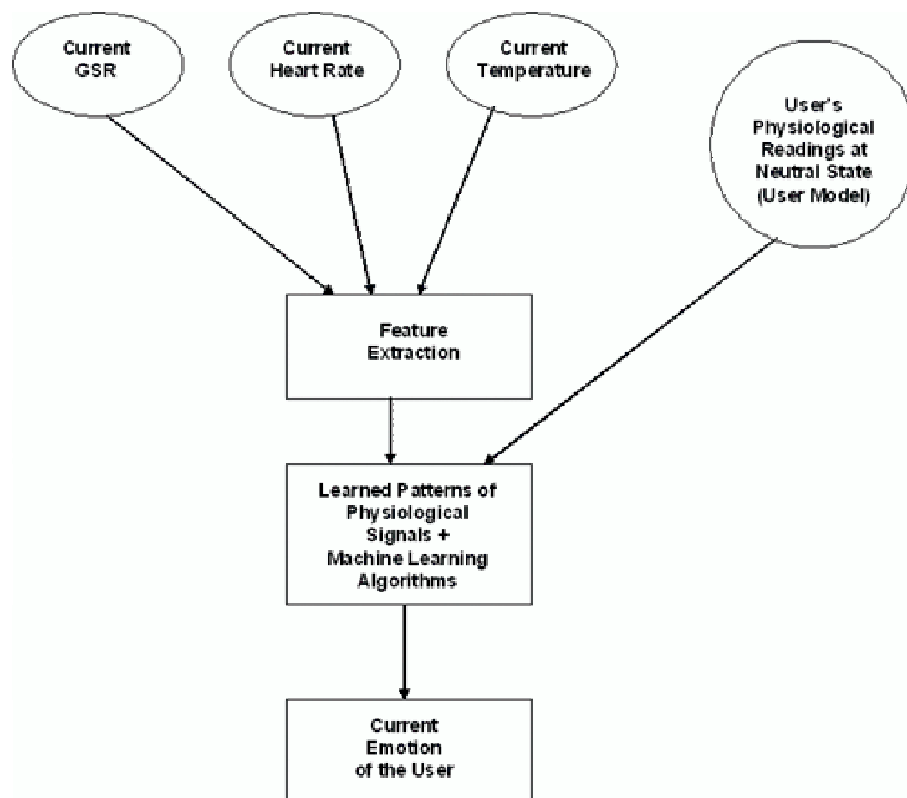


Figure 3.5: Emotion Recognition with More Feature Extraction

Figure 3.6 shows the emotion recognition accuracy rates with KNN algorithm for each of the 6 emotions. KNN could classify sadness with 70.4%, anger with 70.8%, surprise with 73.9%, fear with 80.9%, frustration with 78.3%, and amusement with 69.6% accuracy.

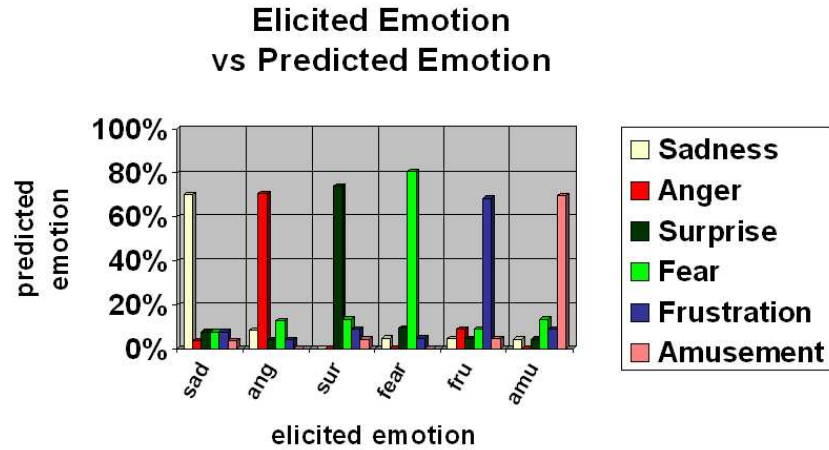


Figure 3.6: Emotion Recognition Results with KNN Algorithm with More Feature Extraction

As can be seen Figure 3.7, the DFA algorithm's recognition accuracy was 77.8% for sadness, 70.8% for anger, 69.6% for surprise, 80.9% for fear, 72.7% for frustration, and 78.3% for amusement.

The recognition accuracy gained with MBP algorithm is shown in Figure 3.8, which was 88.9% for sadness, 91.7% for anger, 73.9% for surprise, 85.6% for fear, 77.3% for frustration, and finally 87.0% for amusement.

Multimodal experiment results showed that emotions can be distinguished from each other and that they can be recognized by collecting and interpreting physiological signals

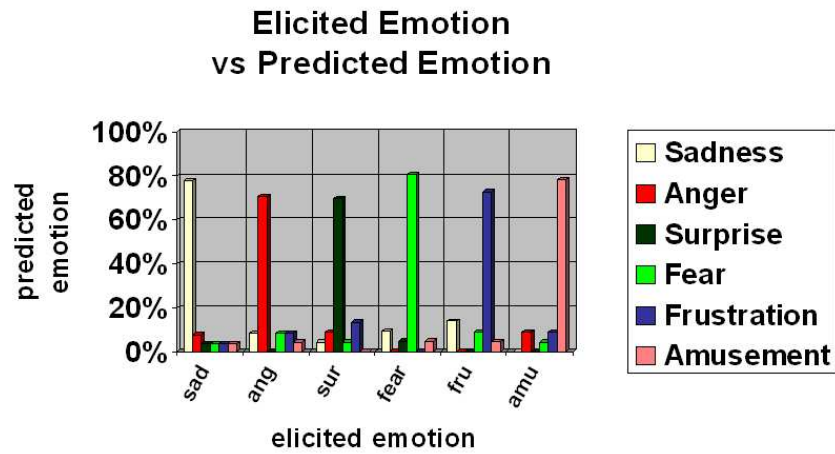


Figure 3.7: Emotion Recognition Results with DFA Algorithm with More Feature Extraction

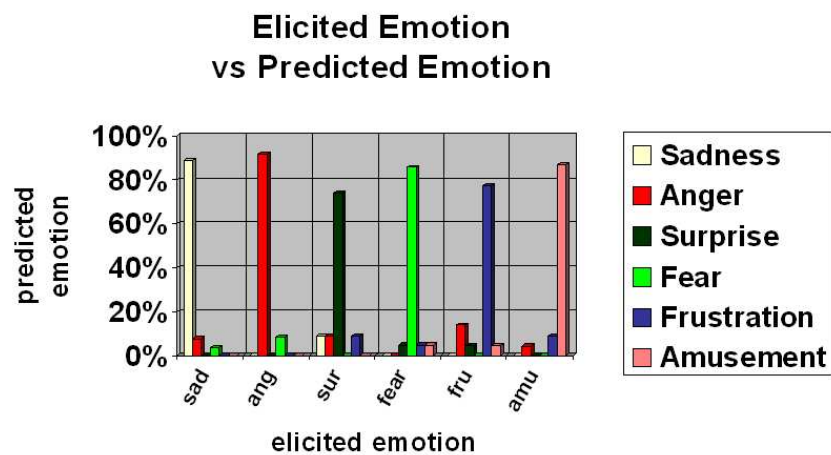


Figure 3.8: Emotion Recognition Results with MBP Algorithm with More Feature Extraction

of the participants. Different physiological signals were important in terms of recognizing different emotions. These results show a relationship between galvanic skin response and frustration. When a participant was frustrated her GSR increased. The difference in GSR

values of the frustrated participants was higher than the differences in both heart rate and temperature values. Similarly, heart rate was more related to anger and fear. Heart rate value of a feared participant increased, whereas it decreased when the participant was angry.

Overall, three algorithms: KNN, DFA, and MBP could recognize emotions with 72.3%, 75.0%, and 84.1% accuracy respectively. In the previous interpretation discussed in Section 3.3, the overall recognition accuracy was 71% with KNN, 74% with DFA, and 83% with MBP. The results of this later study showed that implementing a feature extraction technique slightly improved the performance of all three algorithms.

Recognition accuracy for some emotions was higher with the pattern recognition algorithms than the agreement of the subjects on the same emotions. For example, fear could be recognized with 80.9% accuracy by KNN and DFA and with 85.6% accuracy by MBP, although the subject agreement on fear was 65%. This might be understood from Feldman Barrett et al.'s study [BGC01]: the results of this study indicate that individuals vary in their ability to identify the specific emotions they experience. For example, some individuals are able to indicate whether they are experiencing a negative or a positive emotion, but they cannot identify the specific emotion.

CHAPTER 4

DRIVING SIMULATOR EXPERIMENT IN VIRTUAL REALITY

Designing, conducting, and analyzing the results of the Emotion Elicitation Experiment discussed in Chapter 3 showed that emotions experienced by people can be recognized by analyzing their physiological signals. Designing and conducting an experiment that focuses on eliciting emotions that are related to a specific application (*i.e.* Driving Safety) was the next step.

Driving Simulator Experiment in Virtual Reality was designed to elicit driving related emotions and states - *panic/fear*, *frustration/anger*, and *boredom/sleepiness* - in order to measure the participants' physiological signals while they were experiencing these emotions. Driving Simulator (Figure 4.1 - located in the new Engineering Building of University of Central Florida [UCF]) operating in Virtual Reality (Figure 4.2) was used as the Driving Environment. This driving simulator is created using Virtual Reality technologies and operated by Center for Advanced Transportation Systems Simulation (CATSS) at UCF. Figure 4.3 shows the control room of the driving simulator.



Figure 4.1: The Car Operating in Virtual Reality

A scenario consisting series of events was designed and implemeted to be run on the simulator to elicit panic/fear, frustration/anger, and boredom from the participants of the study. Physiological body signals (GSR, temperature, and heart rate) of the participants were collected via non-invasive wearable computer BodyMedia SenseWear Armband. At the same time an ongoing video of each driver was recorded for annotation and future facial expression recognition purposes. These measurements and recordings were analyzed to find unique patterns of physiological signals for driving-related emotions.



Figure 4.2: The Virtual Reality Environment



Figure 4.3: Control Room of the Driving Simulator

4.1 Driving Simulator Experiment Scenario

In order to elicit driving-related emotions from the participants a scenario that contained a series of traffic events was created. The events were ordered in a way that they would first elicit panic/fear, then frustration/anger, and finally boredom/sleepiness. Baselines were inserted before and after eliciting each emotion.

Below are the events that were created in the scenario to elicit each specific emotion:

- **Panic/fear:** While driving downhill in an accident scene, a child suddenly walks to the middle of the road and stops there and the driver hits him unavoidably. Even the driver tries to avoid hitting the child, this is prevented by disabling the simulator car's breaks and also by placing barricades to both sides of the road so that the driver cannot change lanes (Figure 4.4).
- **Frustration/anger:** After hitting the child unavoidably down the hill, the driver is directed to the city where frustration/anger stimuli were created in. Frustration/anger was elicited through a series of events since only one event would not be enough to elicit the target emotions from some of the participants.

First the driver has to stop in the middle of the road and wait for a couple of men who are carrying a big glass and at the same time talking to another man they meet on the road, thus blocking the road.



Figure 4.4: Child Walking to the Middle of the Road

After passing the glass carrying men, the driver is instructed to turn right at the next intersection, however the way of the driver is blocked by a car that spends excessive amount of time at the lights to make a right turn.

When the driver finally turns right, after traveling around 20 feet, the road is again blocked with a big garbage truck that is trying to make a 3-point turn and park (Figure 4.5). Also, there is taxi behind the participant's car that honks its horn constantly in order to annoy the driver.

After passing by the garbage truck that parked, the driver is instructed to turn left at traffic lights. At this point there is a white car waiting in front of the participant's car to turn left (Figure 4.6). However, as soon as the lights turn to green for the driver,



Figure 4.5: Big Garbage Truck Making a 3-point Turn

several pedestrians start passing across the road and the lights turn to red again before the driver gets a chance to turn left.

After passing the pedestrians and starting to drive in a narrow road, a bus driver rives right on to the participant's car as they are going to collide, but turns his wheel at the last moment, and finally insults the driver verbally while passing by him.

- **Boredom/Sleepiness:** After leaving the city where the frustrating events happened, the participants drive in a straight, long road where no event happens.
- **Baseline:** Baseline contains an eventless and enjoyable drive between the emotion eliciting events.



Figure 4.6: Driver Waiting for the Pedestrians to Cross the Road

One of the biggest differences between Healey’s work [Hea00] and this dissertation is the driving environments where the experiments were conducted in. In Healey’s experiments real-life traffic is used as opposed to a simulator in a Virtual Reality. A Virtual Reality environment provides a totally controlled environment and the advantages of this controlled environment over the unpredictable real-life traffic environment are:

- Every participant experiences the exact same scenario, which makes it possible to do comparisons between the participants and derive general results.
- Distracters such as noise and motion that influence the physiological signals are kept equal within each scenario including the baselines, which makes it possible to capture the changes in responses that are only due to the changes in emotional states of the participants.

4.2 Driving Simulator Experiment Setup

Sample: The sample included 41 undergraduate and graduate students enrolled in UCF. There were 5 females and 36 males and their ages ranged from 18 to 55. Specific ages were not requested; therefore, a mean age was not calculated.

Procedure: One subject participated in the study during each session. After signing the consent forms and filling out the pre-study questionnaires, non-invasive BodyMedia SenseWear Armband (Figure 1.2) was placed on the participants' left arm (to collect galvanic skin response and temperature values). After the armband was activated Polar chest strap that works in compliance with the armband was placed on the participants' chest (to collect heart rate values). Once the chest strap signaled that it started communicating with the armband the participants were told the following: (i) They would be driving a Saturn automatic transmission car in a Virtual Reality environment. (ii) They are expected to obey the regular traffic rules such as not driving over the speed limit and stopping at red lights and stop signs. (iii) The red and yellow arrows on the simulator screen would show them which way to turn (iv) The car has motion and as a result it might cause motion sickness. In case that happens they should stop the car and not continue the experiment.

After the participants took their places in the driving seat of the simulator car, they were told the following: (i) to fasten their seat belts, (ii) to start the car by turning on the ignition key, and (iii) to put the gear in 'D' (Drive) and start driving. The driving simulator scenario discussed in section 4.1 was activated once they turned the ignition key on. While

the participants were driving the car, the videos of their faces were recorded with a digital camcorder that was mounted on the dash of the simulator car. These videos are saved for future facial expression recognition studies. The scenario lasted for 12-16 minutes depending on the driving speed of each participant. The simulator warned the participants vocally when the scenario was over. After they put the gear in park, stopped and left the car, chest straps and armbands were removed and the data collected in the armbands were downloaded to a computer. Finally the participants were asked to fill out the post-study questionnaire. After the post-study questionnaires were collected the participants were thanked for their time and for joining the study and they were asked if they had any questions.

Measures: The pre-questionnaire included demographic questions about profession, gender, age range, participants' driver's license history, and driving frequency of the participants. The post study questionnaire included seven questions (3 on the emotions experienced, 1 on the realisticness of the simulator, and 3 on the participants' experiences in real-life traffic). Each of first three questions asked whether the participants experienced the elicited target emotion, the intensity of this emotion on a 6-point scale (6 being highest) if they experienced it, and whether there was another emotion they experienced. The fourth question asked how realistic the participants found the driving simulator on a 6-point scale (6 being highest). Finally, last three questions asked the participants how often they get frustrated or angry, how often they get panicked or fearful, and how often they get bored while driving on a 6-point scale (1 being never, 6 being always).

4.3 Emotion Recognition with Machine Learning

The physiological signals that were measured during the Driving Simulator Experiment were analyzed using k -Nearest Neighbor (KNN) and Marquardt-Backpropagation (MBP) discussed in Section 3.3 and Resilient Backpropagation (RBP) Algorithms to find unique patterns of physiological signals that match driving-related emotions.

4.3.1 Feature Extraction

After determining the time slots corresponding to the point in the driving scenario where the intended emotion was most likely to be experienced, the experiment described above resulted in the following set of physiological records: 29 for panic/fear, 30 for frustration/anger, and 27 for boredom/sleepiness. Data was stored and normalized in the same way as discussed in section 3.3 .

The same features stated in Section 3.4 were extracted from the collected and normalized data (minimum, maximum, mean, and standard deviation) for each physiological signal type (GSR, temperature, and heart rate). These features were given as input to the pattern recognition algorithms. Figure 4.7 shows in detail the interaction of the driver within the car and how elicitation and recognition of driving-related emotions were performed.

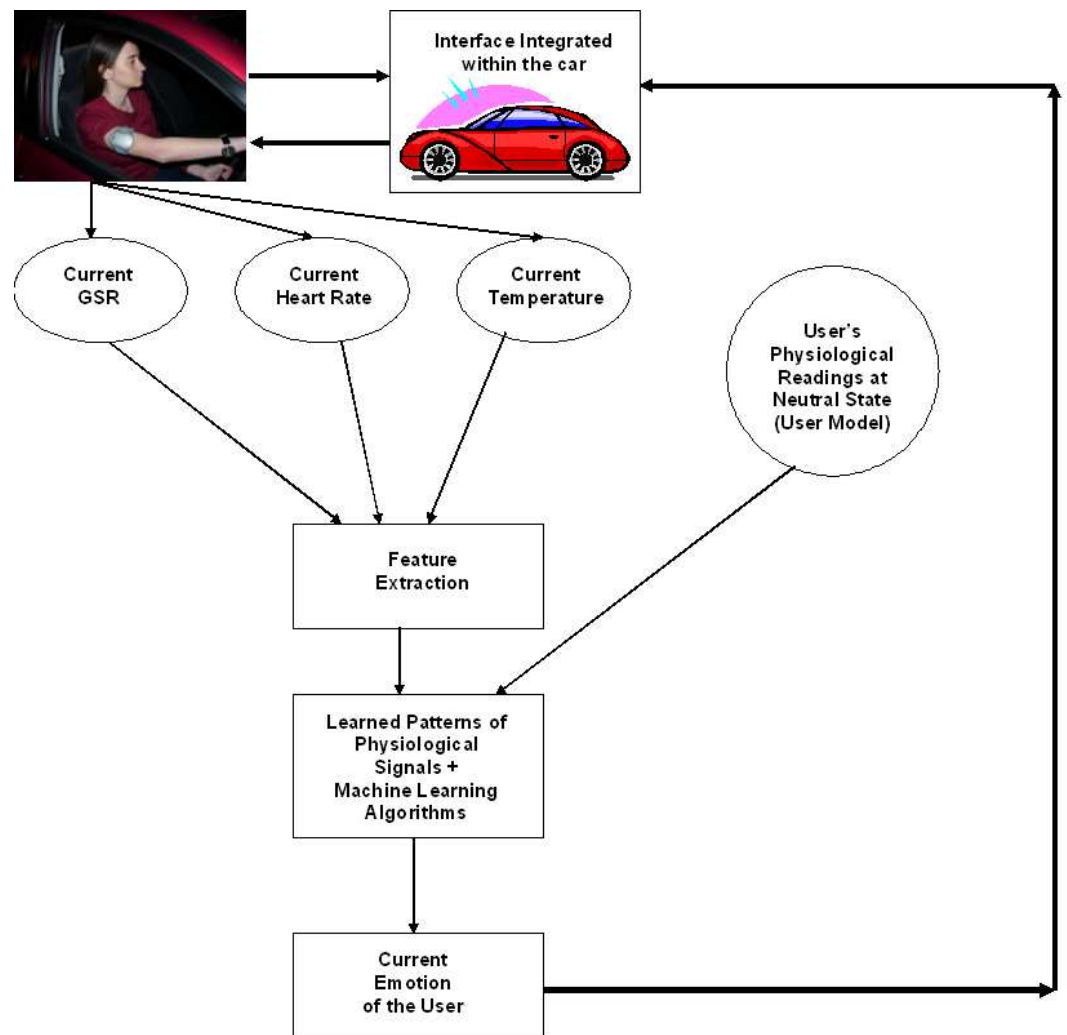


Figure 4.7: Elicitation and Recognition of Driving-related emotions from Physiological Signals.

4.3.2 Emotion Recognition Accuracy with KNN, MBP, and Resilient Backpropagation Algorithms

4.3.2.1 k-Nearest Neighbor and Marquardt Backpropagation Algorithms

The data was first analyzed with KNN and MBP algorithms discussed in Sections 3.3.1 and 3.3.3 respectively. Figure 4.8 shows the neural network structure used with the Marquardt Backpropagation Algorithm. The network is consisted of an input layer with 12 nodes, a hidden layer with 17 nodes, and an output layer with 3 nodes.

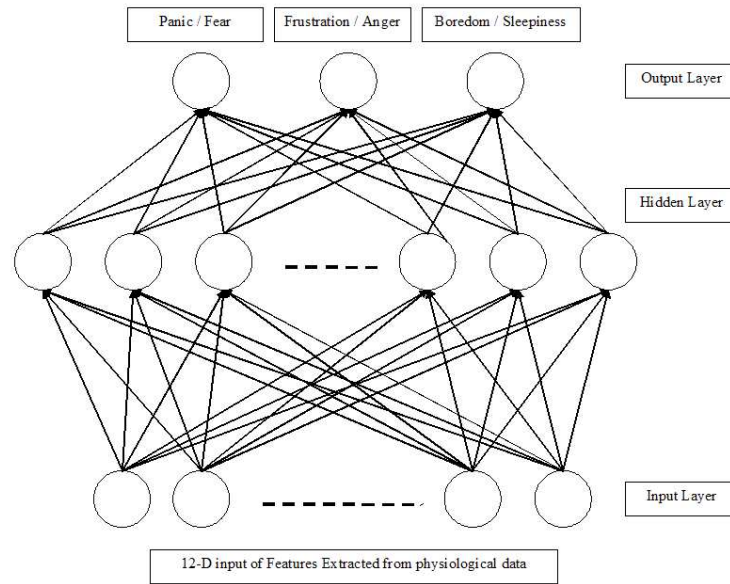


Figure 4.8: Neural Network Architecture for Recognizing Driving-related emotions

Tables 4.1 and 4.2 report the classification accuracy of each emotion set with KNN and MBP respectively.

Table 4.1: Emotion Classification Accuracy with KNN for Each Emotion

Emotion	Correctly Classified Instances	Total Instances	Accuracy
<i>Panic/Fear</i>	24	29	82.8%
<i>Frustration/Anger</i>	22	30	73.3%
<i>Boredom</i>	11	27	40.7%

Table 4.2: Emotion Classification Accuracy with MBP for Each Emotion

Emotion	Correctly Classified Instances	Total Instances	Accuracy
<i>Panic/Fear</i>	25	29	86.2%
<i>Frustration/Anger</i>	20	30	66.7%
<i>Boredom</i>	21	27	77.8%

4.3.2.2 Resilient Backpropagation Algorithm

As can be seen from the table 4.2, the MBP algorithm was not as successful as it was in recognizing the six emotions elicited in Emotion Elicitation Experiment with Movie Clips (Chapter 3). For this reason Resilient Backpropagation (RBP) Algorithm, which is another Neural Network Algorithm was implemented.

Resilient Backpropagation Algorithm [RB93] is a derivation of a backpropagation algorithm, where the magnitude of the derivative of the performance function has no effect on the weight update and only the sign of the derivative is used to determine the direction of the weight update. When the derivative of the performance function has the same sign for two consecutive iterations, the update value for each weight is increased and when the

derivative of the performance function changes sign from the previous iteration, the update value is decreased. No change is made when the derivative is equal to 0.

Each weight and bias value X is adjusted according to the following formula:

$$dX = \Delta X * \text{sign}(gX)$$

where the elements of ΔX are all initialized to the initial Δ and gX is the gradient value. The elements of ΔX are modified after each iteration. If gX has the same sign with the previous iteration then corresponding ΔX is incremented and if gX changes sign from the previous iteration, then the corresponding ΔX is decremented.

Table 4.3 shows the emotion recognition accuracy for emotion by RBP algorithm and Table 4.4 reports the emotion classification accuracy of k-Nearest Neighbor, Marquardt Backpropagation and Resilient Backpropagation Algorithms for all emotions.

Table 4.3: Emotion Classification Accuracy with RBP for Each Emotion

Emotion	Correctly Classified Instances	Total Instances	Accuracy
<i>Panic/Fear</i>	26	29	89.7%
<i>Frustration/Anger</i>	29	30	96.7%
<i>Boredom</i>	24	27	88.9%

An important issue while evaluating the performance of the algorithms in real-life applications is the rate of false negative results (i.e. system does not recognize the negative emotional state of the user) and false positive results (i.e. system recognizes a negative emo-

Table 4.4: Emotion Classification Accuracy with KNN, MBP and RBP for all emotions

Algorithm	Correctly Classified Instances	Total Instances	Accuracy
KNN	57	86	66.3%
MBP	66	86	76.7%
RBP	79	86	91.9%

tional state of the user although she is not experiencing this state) obtained by interpreting the physiological signals of the user. Table 4.5 summarizes the results that can be obtained while performing emotion recognition.

Table 4.5: Consequences Related to Emotion Recognition

Emotion	Emotion Recognized	Consequence
Experienced	Yes	Accurate Emotion Recognition
Experienced	No	False Negative
Not Experienced	Yes	False Positive
Not Experienced	No	Accurate No-Emotion Recognition

Due to the nature of emotion recognition problem, it is impossible to prevent all false negatives and false positives, however the rate of false negatives and false positives can be decreased by implementing various techniques. One of these techniques is combining different pattern recognition algorithms for higher recognition accuracy. Another useful technique to increase recognition accuracy might be integrating different modalities that the emotions can be recognized from such as physiology, facial expressions, and vocal intonation.

CHAPTER 5

USER MODELING AND INTELLIGENT USER INTERFACE ADAPTATION

Each user's responses are different from the other users' while interacting with an intelligent system [NS89]. This makes it necessary to build user models that will enable the system to record relevant user information in order to interact with its users appropriately.

Conventional user models are built on what the user knows or does not know about the specific context, what her/his skills and goals are, and her/his self-report about what s/he likes or dislikes. The applications of this traditional user modeling include student modeling [BJB02], [MC02], [CMS00], [Sel94], news access [BP00], e-commerce [FC00], and health-care [WFN02].

However, none of these conventional models includes a very important component of human intelligence: *Affect* and *Emotions* as discussed in Chapter 1. After recognizing the user's emotions successfully with the pattern recognition algorithms discussed in Chapters 3 and 4, and giving feedback to her about her emotional state, the next step is adapting MAUI to the user's emotional state by also considering the current context and application and user

dependent specifics, such as user's personality traits. Adapting to each user differently is a very important component of intelligent systems, since every user's interaction is different with these systems. Bayesian Belief Networks (BBN) [Pea88] formalization was employed to create these user models in order to enable MAUI to adapt itself to the applications discussed in Section 1.3. Figure 5.1 presents the complete user interaction and emotion recognition system integrated with Affective User Modeling for Driving Safety application.

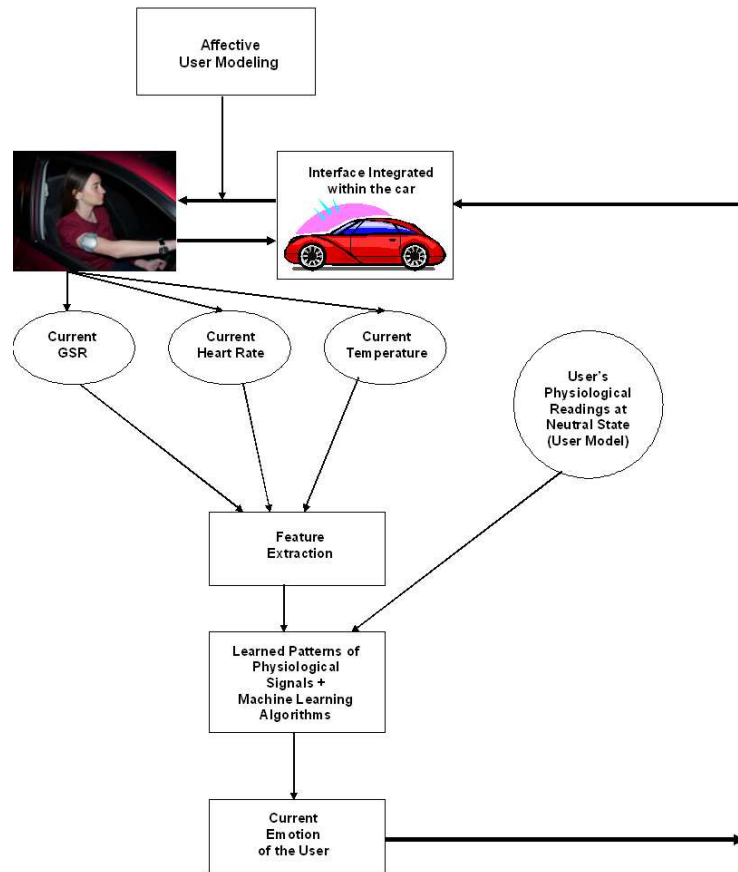


Figure 5.1: Integrating User Modeling in Emotion Recognition and User Interaction.

5.1 Bayesian Belief Networks

Bayesian Belief Networks (BBN) (also known as Bayesian Network or probabilistic causal network) were used to build the user modeling in this system. Bayesian Belief Networks [Pea88] are directed acyclic graphs (DGA), where each node represents a random discrete variable or uncertain quantity that can take two or more possible values. The directed arcs between the nodes represent the direct causal dependencies among these random variables. The conditional probabilities that are assigned to these arcs determine the strength of the dependency between two variables.

A Bayesian Belief Network can be defined by specifying:

1. Set of random variables: $\{X_1, X_2, X_3, \dots, X_n\}$.
2. Set of arcs among these random variables. The arcs should be directed and the graph should be acyclic. If there is an arc from X_1 to X_2 , X_1 is called as the *parent* of X_2 and X_2 is called as the *child* of X_1 .
3. Probability of each random variable that is dependent on the combination of its parents.

For a random variable X_i , the set of its parents is represented as $par(X_i)$, and the conditional probability of X_i is defined as:

$$P(X_i|par(X_i))$$

.

4. If a node has no parents unconditional probabilities are used.

Unlike the traditional rule-based expert systems, BBNs are able to represent and reason with uncertain knowledge. They can update a belief in a particular case when a new evidence is provided. This is Bayesian Belief Networks was chosen to create the user model in MAUI.

5.2 User Modeling with Bayesian Belief Networks

The Bayesian Belief Networks that record the related user information (preferences, personality, affective information) in the system is coded in Visual C++ by using and adpting Norsys' [Cor] Netica C Application Programmer Interface. Figure 5.2 shows the Bayesian Belief Network built for modeling the user in the driving environment.

As shown in 5.2, user model for the driving environment is built as a decision support system. There are various parameters that would affect the optimal action that should be chosen by the adaptive interface. These parameters (i.e. nodes of the Belief Network) are:

- Recognized emotion of the driver(*e.g.* Anger, Panic, *etc.*)
- Driver's personality traits (*e.g.* extravert, open, *etc.*)
- Driver's previous responses when interacting with the interface (i.e. driver's satisfaction)
- Driver's age

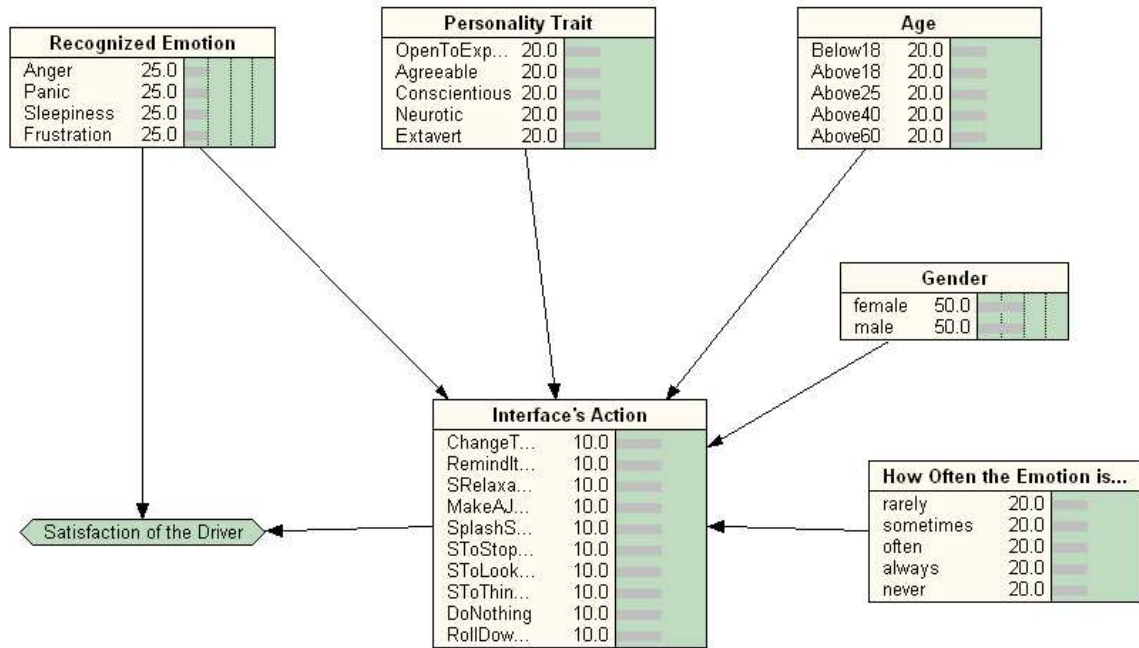


Figure 5.2: Bayesian Belief Network Representation of Model of the Driver.

- Driver's gender
- Possible actions that can be taken by the interface

These variables were chosen to be included in the model since previous studies suggest that people's driving is influenced by them. For the model, possible emotions and states that a driver can experience are chosen as: Anger, Frustration, Panic, Boredom, and Sleepiness and their influence on one's driving is discussed in Section 1.3.1.

Personality traits of the driver were included in the user model, because previous studies suggest that personality differences result in different emotional responses and physiological arousal to the same stimuli [Kah73], and the preferences of a person are affected by her

personality [NL00]. Questionnaires can be used in order to successfully identify a driver's personality. Five-Factor-Model was chosen to determine the personality traits [CM92]. Following are the personality traits based on the Five-Factor-Model:

- *Neuroticism* (High neuroticism leads to violent and negative emotions and interferes with the ability to handle problems)
- *Extraversion* (High extravert people work in people oriented jobs, while low extravert people mostly work in task oriented jobs)
- *Openness to experience* (Open people are more liberal in their values)
- *Agreeableness* (High agreeable people are skeptical and mistrustful)
- *Conscientiousness* (High conscientious people are hard-working and energetic) [CM92].

These personality traits also influence the way people are driving. Celler et al.'s study [CNY00] show that Agreeableness have slight negative correlation with the number of driving tickets and Arthur and Graziano's [AG96] study showed that people with low Conscientiousness level have higher risk of being in a traffic accident.

Age and gender also have effect on people's driving [Tra03]. Younger drivers are at a greater level of crash involvement (with a distinguished difference between 18-19-years-olds and 25-years-olds), more likely to take risks, they tend to show increased level social deviance, display the highest driving violation rates, and associate a lower level of risk perception, whereas older drivers tend to show a greater frequency of drowsy driving and

tend to more likely suffer from visual impairments that affect their driving [Tra03]. When it comes to gender differences, men are more likely to have accidents because of rule violations, and they make up majority of the aggressive drivers. Women on the other hand, are more likely to involve in crashes caused by perceptual or judgmental errors and they have lower driving confidence [Tra03].

The node *Satisfaction* shown in the model represents the satisfaction of the driver after an action is chosen and taken by the intelligent interface according to the emotion experienced by the driver. For example, if the driver was angry and the interface chose to suggest him try to be calm and control himself, and if this made the driver even more angry, this information is recorded in the satisfaction node for future references.

As it can be seen from the model, the dependencies between the variables are defined as probabilistic values. This network enables the adaptive interface to take an optimal action depending on the user's emotional state, personality, *etc.* These actions might include but not limited to:

- Remind the driver it feels good to be civil [Jam00]
- Change the radio station [Jam00]
- Suggest the driver to do Larson Driver Relaxation Exercise [LR99]
- Suggest the driver to say something positive out, smile and look pleasant [Jam00]
- Remind the driver what her loved ones would feel if something happened to her [Jam00]

- Roll down the window
- Splash some water on her/his face
- Suggest to stop the car and have some rest

Using Bayesian Belief Networks was chosen to model the users in driving environment because of the BBN's ability to represent uncertain knowledge. In the model described above there are five nodes (events) that the adaptive interface's action is dependent. Each of these events can occur in several different ways (for example recognized emotion might be anger, boredom, or panic or user's personality trait one of the five described above), which leads to hundreds of different possible combinations of events thus hundreds of different possibilities for choosing the optimum interface action. This model will be complete when an expert or experts provide knowledge in the form of causal structures among the variables.

5.3 Interface Actions chosen by the Driver Model

There are one thousand (1000) combinations of different events that can influence the adaptive interface's optimal action that will be determined by the Bayesian Belief Network. Those thousand combinations are calculated by multiplying the different events related to each variable in the Network. There are 4 possible events for recognized emotion (Anger, Panic, Sleepiness, Frustration), 5 possible events for driver's personality traits (Neuroticism, Extraversion, Openness to experience, Agreeableness, Conscientiousness), 5 possible events

for driver's age (below 18, 18 to 25, 25 to 40, 40 to 60, and over 60), 2 possible events for driver's gender (female and male), and 5 possible events for the frequency of experiencing the specific emotion by the driver (never, rarely, sometimes, often, and always).

$$\#of combinations = 4 * 5 * 5 * 2 * 5 = 1000$$

As stated in section 5.2, the probabilistic dependencies in this network should be determined by experts from the field of Psychology and Driving Safety with an extensive amount of effort. However, in order to evaluate the applicability of the Driver Model problem to Bayesian Belief Networks, same network discussed in section 5.2 was created with a smaller number of events for each variable (two events for each variable): Anger and sleepiness for recognized emotion/state, Conscientiousness and Neuroticism for personality trait, below 25 and above 25 for driver's age, female and male for driver's gender, and rarely and often for frequency of experiencing the specific emotion by the driver. Since there are only two possible events for each random variable, the total number of possible combinations of events is 32. Figure 5.3, 5.4, 5.5, and 5.6 shows the results of running this smaller model with various combinations of events.

The optimal interface actions chosen by the Bayesian Belief Network in each case is summarized in Table 5.1

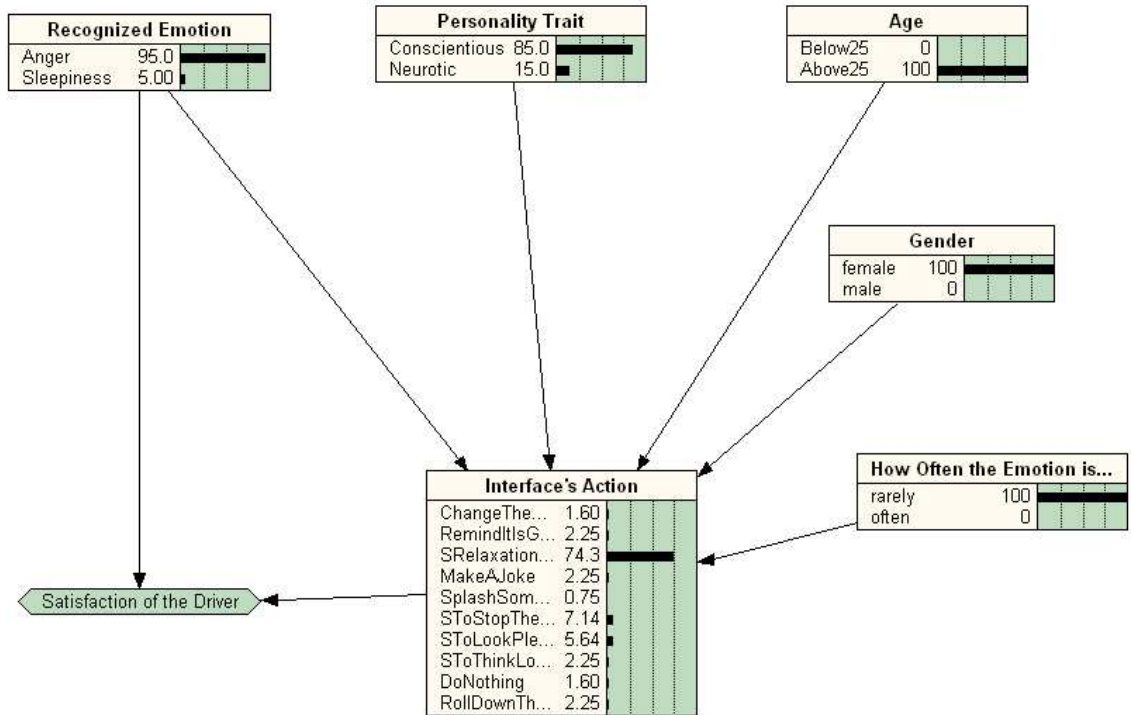


Figure 5.3: Bayesian Belief Network Representation of Case 1

Table 5.1: Optimal Actions Chosen by the Model for Each Combination of Events

Emotion	Personality Trait	Age	Gender	Frequency	Action
Anger (95%)	Conscientious(85%)	Above 25	Female	Rarely	Suggest a Relaxation Technique(74.3%)
Anger (90%)	Neurotic (90%)	Below 25	Male	Rarely	Make a Joke (58.1%)
Sleepiness (88%)	Conscientious (95%)	Below 25	Female	Often	Suggest to Stop the Car and Take a Rest(69.7%)
Sleepiness (88%)	Neurotic (85%)	Below 25	Female	Often	Change the Radio Station (65.1%)

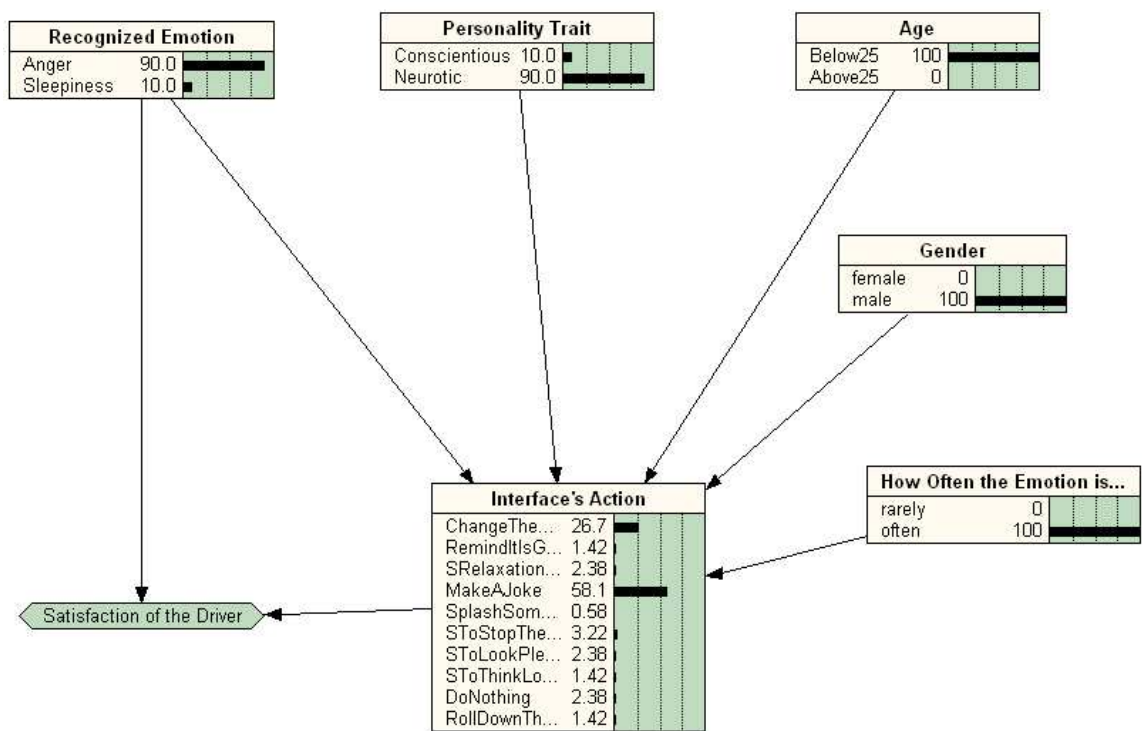


Figure 5.4: Bayesian Belief Network Representation of Case 2

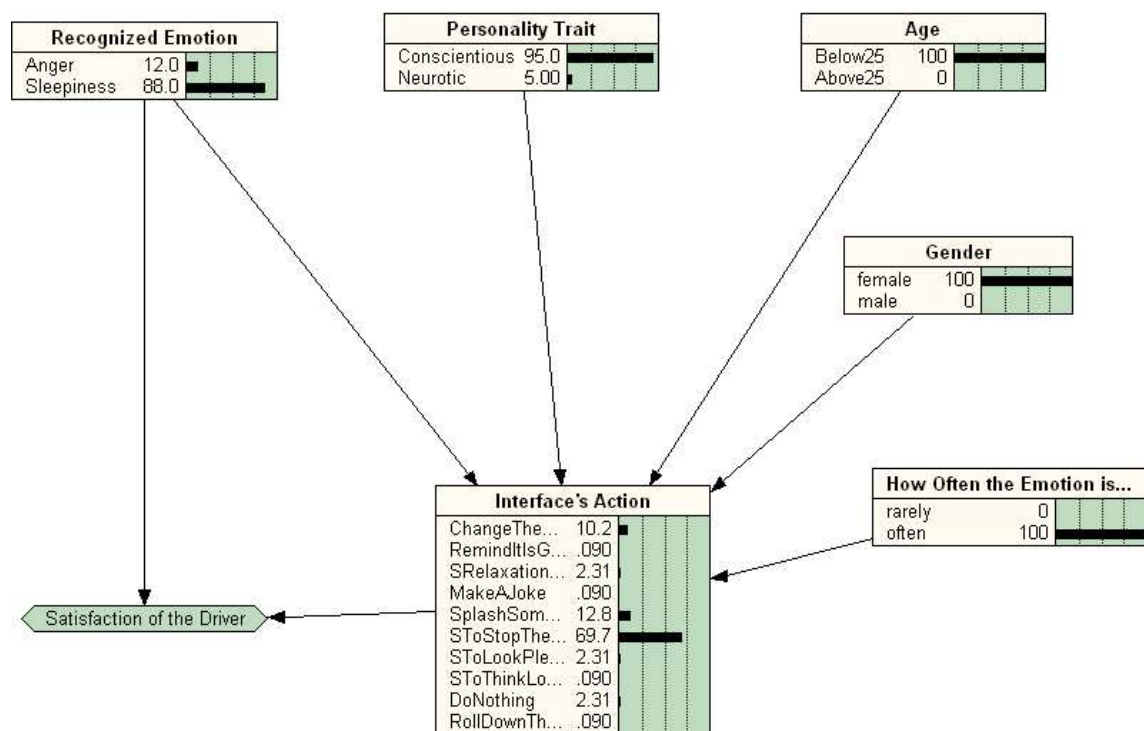


Figure 5.5: Bayesian Belief Network Representation of Case 3

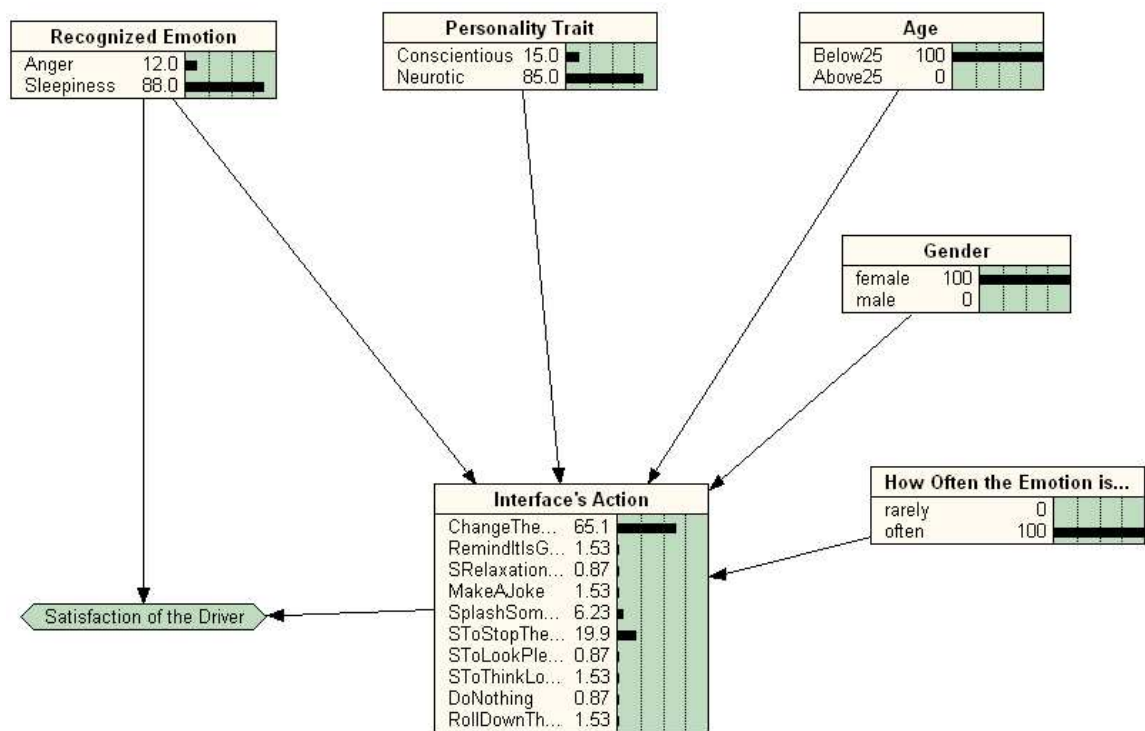


Figure 5.6: Bayesian Belief Network Representation of Case 4

CHAPTER 6

VISUALIZATION OF THE RESULTS

The experiment results obtained in the preliminary work (Chapter 3) and the results obtained from Driving Simulator Experiment (Chapter 4) enabled us improve the design of the Multimodal Affective User Interface discussed in Section 2.1. Analysis of data collected in these experiments found unique patterns of physiological signals that are associated with emotions. These patterns enabled us to recognize the emotional states of the MAUI users in real-life scenarios. Once these states are recognized, this information is used for facilitating a natural communication between MAUI and its users as shown in Figure 5.1 in Chapter 5.

The "Feedback to User" part of MAUI that was discussed in Section 2.1.3 is used to inform the user about her emotional state. This part of the Intelligent User Interface gives detailed information about the components of the emotion experienced by the user such as valence (*i.e.* negativity or positivity of the emotion), facial expression associated with the emotion (*e.g.* angry, sad, neutral), causal chain (*i.e.* causes and effects of the emotion that is being experienced), *etc.* It also allows the user to make changes to some of these components.

Another form of feedback given to users is by using the "avatar" discussed in Section 2.1.4. The avatar in the MAUI system was created using Haptik PeoplePutty software (Haptik, Inc.) and has the ability to make context relevant facial expressions. The various ways the avatar can be used are: mirroring users' emotions as a method to confirm their emotional states (see Figure 6.1); responding with socially appropriate facial expressions as users display their emotional states (i.e., avatar displays empathy when the user is frustrated by a task); assisting users in understanding their own emotional states by prompting them with simple questions and comparing the various components of the states they believe they are experiencing with the system's output; or displaying the facial expressions of real individuals in a text-based chat session in order to enhance communication.



Figure 6.1: Avatar Mirroring User's Anger and Sadness Respectively.

In addition, in order to address individual differences in user preferences, the MAUI system provides a choice of avatar displays including different ages, genders, ethnic backgrounds, skin colors, voices, hair, make-up, accessories, and backgrounds.

The current Multimodal Affective User Interface is capable of:

- Recognizing the current user from her/his face and adapting MAUI according to the preferences of the user (e.g. changing the avatar and the background) (Figure 6.2 and Figure 6.3)

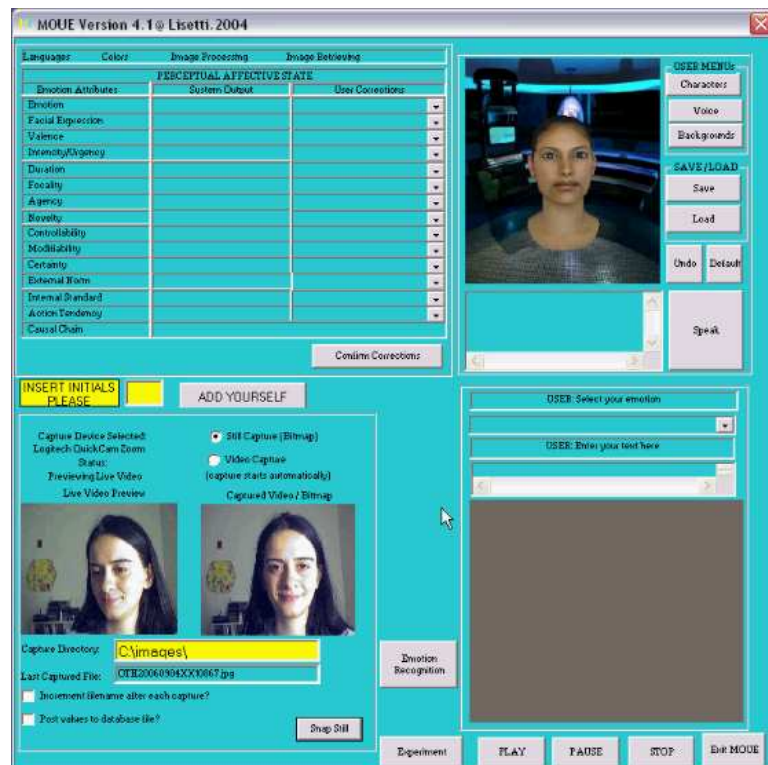


Figure 6.2: MAUI Before Recognizing the User.

- Providing a choice of avatar displays including different ages, genders, ethnic backgrounds, skin colors, voices, hair, make-up, accessories, and backgrounds in order to address individual differences in user preferences (Figure 6.4)

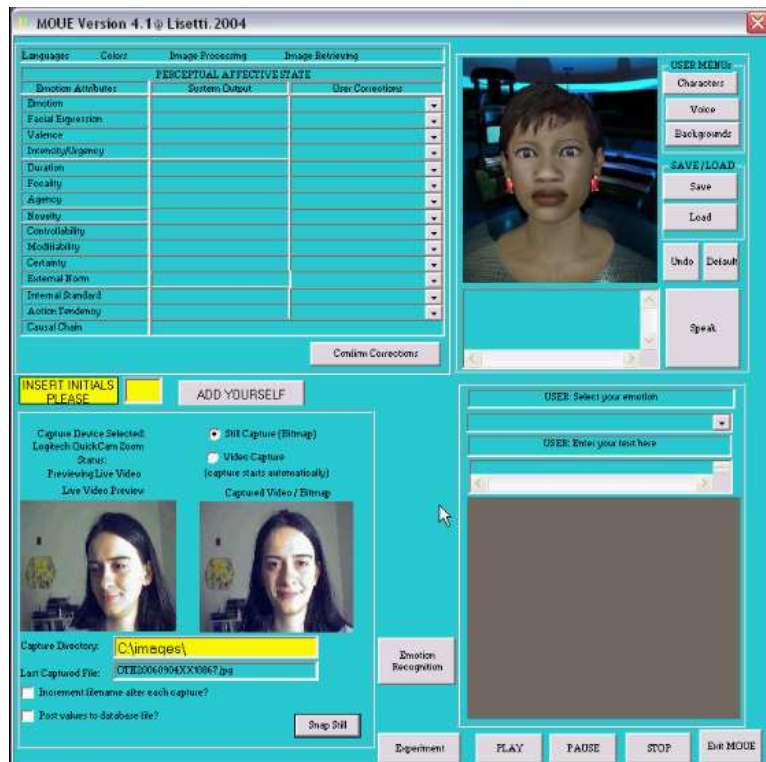


Figure 6.3: MAUI After Recognizing the User.

- A window for playing video on the lower right hand side of MAUI. This window enables to show the emotion eliciting stimuli, in this case a slide show, but could also be another multimodal stimulus. (Figure 6.5)
- Enabling the option of recording and re-playing the videos described above, as well as the recorded facial video of the participant. This allows us to observe the changes in various modalities (physiological signals, facial expressions) simultaneously.
- Mirroring the current emotional state of the user with avatar while s/he is watching the emotion elicitation movie clips (Figure 6.6 Figure 6.7 Figure 6.8).



Figure 6.4: Various Anthropomorphic Avatars for Different User Preferences.

- A separate dialog window that allows analyzing the previously collected physiological signals data stored in an XML file using the algorithms discussed in sections 3.3 and 4.3 to classify them into corresponding emotions (Figure 6.9).
- Avatar mirroring the recognized emotions from the physiological data mentioned in the previous item (Figure 6.10).
- Capturing images of users while they are showing emotions through their facial expressions and dumping these pictures to a web-based database called E-FERET.

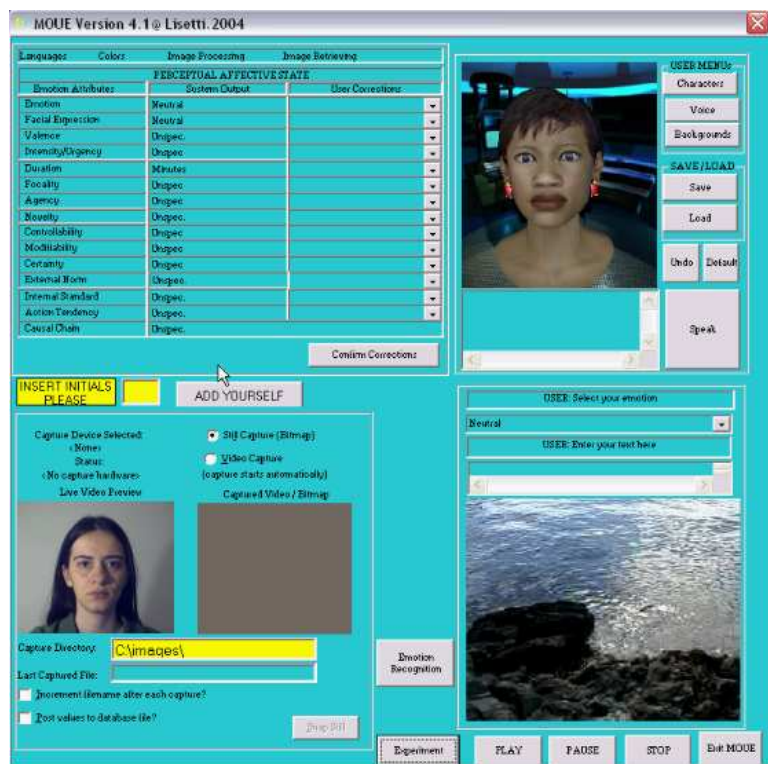


Figure 6.5: Showing video as a Multimodal Stimulus and Avatar mirroring user's Neutral State

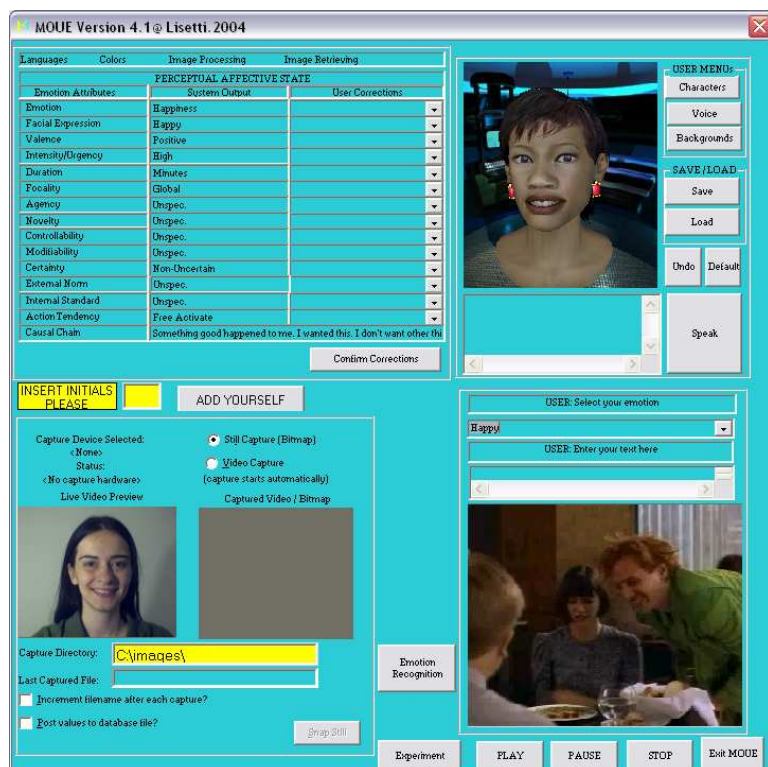


Figure 6.6: Avatar Mirroring User's Happiness

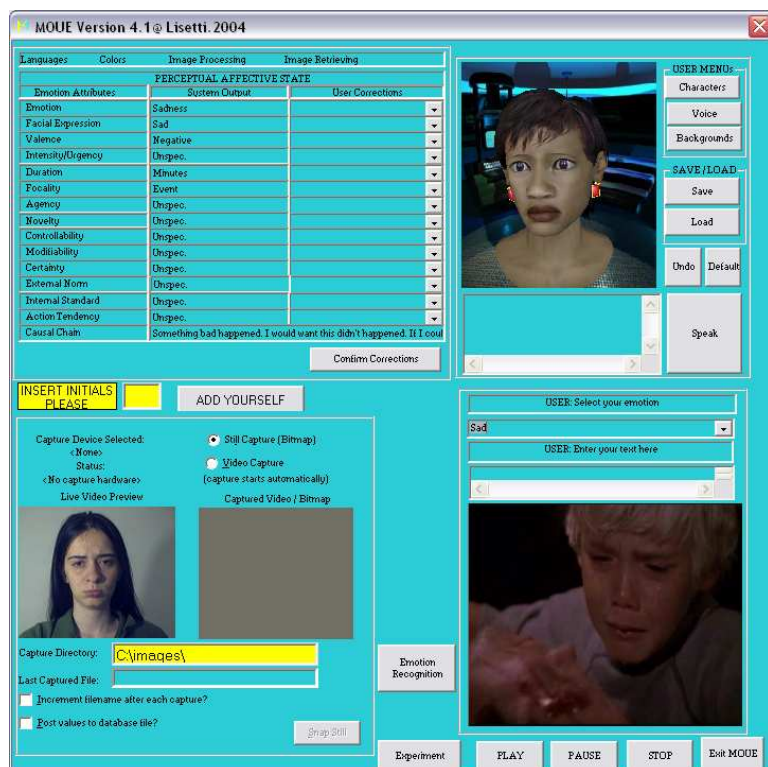


Figure 6.7: Avatar Mirroring User's Sadness



Figure 6.8: Avatar Mirroring User's Anger

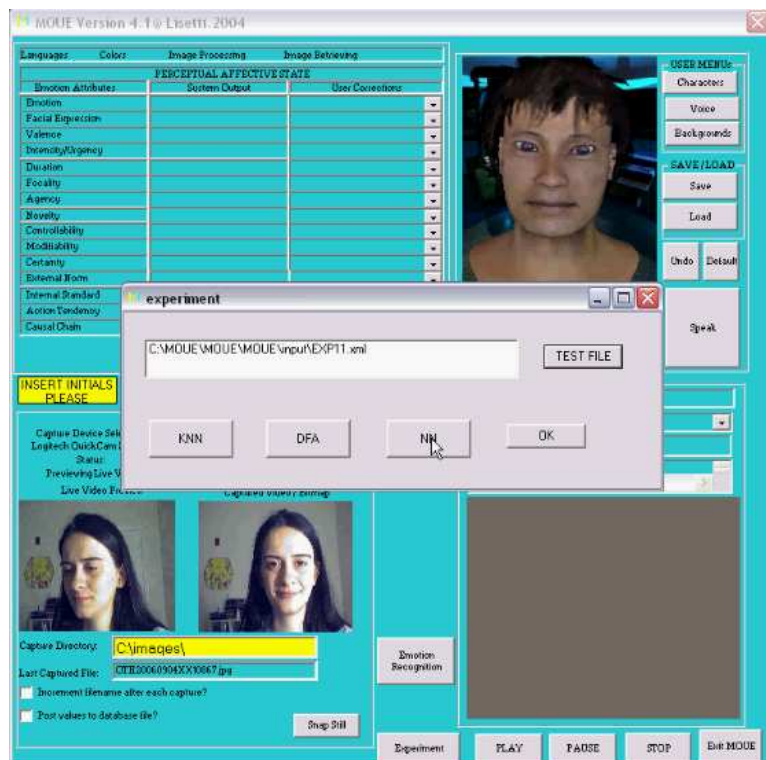


Figure 6.9: MAUI as a research development tool to recognize emotions from pre-collected physiological data.

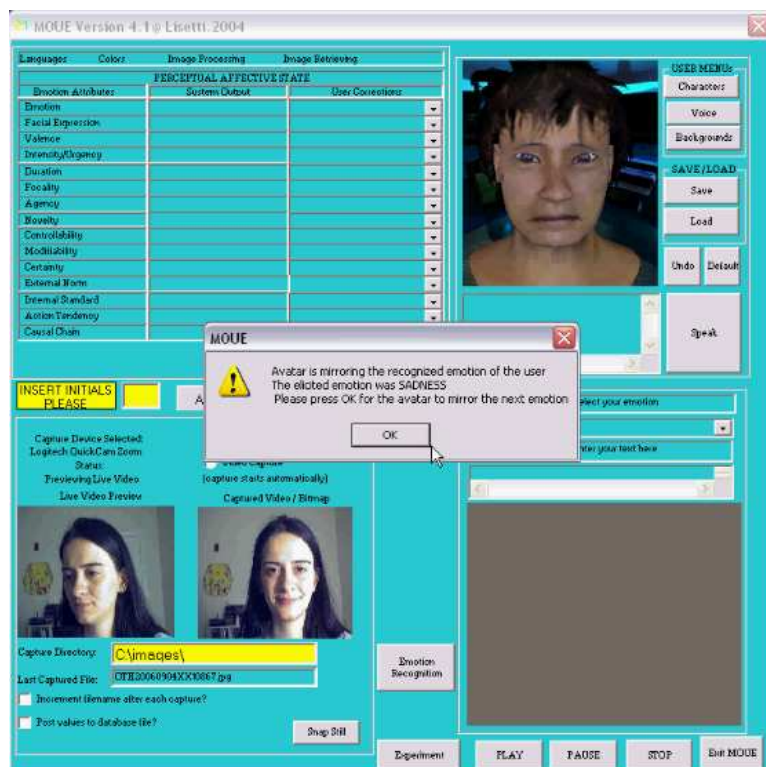


Figure 6.10: Avatar mirroring the emotion recognized from the physiological data.

CHAPTER 7

FUTURE WORK

This chapter discusses the possible improvements and extensions that can be made to the studies described in this dissertation.

7.1 Real-Life and Real-Time Applications

For this study, experiments in controlled environments were conducted in order to map physiological signals to corresponding emotions. These controlled environments were selected so that the changes in physiological signals were caused by changes in emotional states only because there are a lot of factors that influence the physiology in humans such as the exercises they are doing or the fluids they are drinking.

An important next step is analyzing physiological data measured in real-life situations. In order to handle the real-life situations the system should have an extensive knowledge of the environment, such as what the user is currently doing and it also should build a model of user's physiological data at the user's emotionally neutral state in the current environment.

Also, in order to recognize emotions in real-time, state-of-the-art wearable wireless computers that can transmit data to the system simultaneously should be used. BodyMedia's Wireless Transmitter shown in Figure 7.1 can be used for this real-time data transmission purpose.



Figure 7.1: BodyMedia Wireless Transmitter

7.2 More Modalities

Another approach to improve this study would be integrating emotion recognition from physiological signals with emotion recognition from other modalities such as facial expression and

vocal intonation recognition and natural language understanding to increase the recognition accuracy.

With this approach, even if one of the modalities fail to recognize the current emotional state of the user accurately or it doesn't categorize it in any emotion class, one or more of the other modalities can be more influential in categorizing the emotion and can recognize it correctly.

7.3 Improved Algorithms for Emotion Recognition

The algorithms implemented in this dissertation can be improved to analyze the data better and increase the recognition accuracy. In order to facilitate an improvement in data analysis more experiments can be designed to collect more data, different feature extraction techniques can be performed on this collected data, and different pattern recognition algorithms could be combined while analyzing the data.

7.4 User Modeling for Different Applications

The user model in this dissertation was built for the adaptive intelligent system to be able to adapt to the current emotional state of the user for a Driving Safety application. Creating adaptive intelligent systems for other various applications such as learning/training and

telemedicine is also another important near future research objective. In the user model created for a driver some of the variables that would influence the optimal action chosen by the interface were personality, age, and gender. When a student user model is created for a learning application, more parameters can be added to the ones mentioned above, such as goals and the knowledge of the student.

7.5 Different Wearable Computers

As another extension to this study, the experiments conducted for this dissertation would be re-conducted in order to measure and collect the physiological signals of the participants with different wearable computers. One of these wearable computers is Procomp+ shown in Figure 7.2. This would make it possible to compare the performances of various wearable computers.



Figure 7.2: Procomp+

CHAPTER 8

CONCLUSIONS

An Adaptive Intelligent System was designed and built to recognize affective states of its users from the physiological signals in order to give the appropriate respond and adapt accordingly. Experiments were designed and conducted to collect various physiological signals (Galvanic Skin Response, Heart Rate, and Temperature) while the participants are experiencing the elicited emotions. Non-invasive state-of-the-art wearable computers were used to collect the physiological data from the participants. Pattern recognition algorithms were implemented to analyze the collected data and to recognize the affective patterns in them. A User Model was developed to record the affective information about the user in order to enable the intelligent system adapt to the user and the situation according to the user dependent parameters such as personality traits, age, gender, etc.

The first experiment designed in a highly controlled environment where six emotions (Sadness, Anger, Surprise, Fear, Frustration, and Amusement) were elicited from the participants via Multi-modal Stimuli including emotional movie clips, nature pictures, music, and difficult mathematical and analytical problems. The physiological signals of the par-

ticipants were collected with non-invasive BodyMedia SenseWear Armband and Polar chest strap that works in compliance with the armband. Three different algorithms k -Nearest Neighbor [KNN], Discriminant Function Analysis [DFA], and Marquardt-Backpropagation [MBP] Algorithms) were implemented to recognize the patterns in the physiological signals of the participants measured while they were experiencing those six emotions. Overall, three algorithms: KNN, DFA, and MBP could classify those emotions with 72.3%, 75.0%, and 84.1% accuracy respectively.

In order to relate physiological signals to driving-related emotions, a driving experiment was designed and conducted in a highly controlled virtual reality environment. A driving scenario consisting various traffic events was created to elicit Panic/Fear, Frustration/Anger, and Amusement from the participants. BodyMedia SenseWear Armband and Polar chest strap were used to measure Galvanic Skin Response, Heart Rate, and Skin Temperature. KNN and Resilient Backpropagation [RBP] Algorithms were used to classify the physiological signals into corresponding emotions. Overall, KNN could classify these three emotions with 66.3%, MBP could classify them with 76.7% and RBP could classify them with 91.9% accuracy.

Once it was found that emotions can be recognized by analyzing the physiological signals, Adaptive Intelligent User Interface was designed and created that can recognize its user emotions by its pattern recognition algorithms and respond to these emotions by the Bayesian Belief Network formalization employed in it. The Bayesian Belief Networks were used to

build the user modeling that will enable appropriate adaptation of the interface to the user's experienced emotion, personality trait, age, *etc.*

APPENDIX A

ABOUT THE STUDENT

A.1 Short Biography

Fatma Nasoz is a PhD candidate in the Computer Science department of University of Central Florida, Orlando since August 2001. She earned her MSc. degree in Computer Science from the University of Central and her BSc. degree in Computer Engineering from Bogazici University in Turkey, in 2003 and 2000 respectively. She was awarded the Center for Advanced Transportation System Simulation (CATSS) Scholarship in 2002 to model emotions of drivers for increased safety.

Her research area is Affective Computing and she is a member of Affective Social Computing Laboratory (ASCL) at the University of Central Florida supervised by Dr Christine Lisetti. She specifically focuses on creating Adaptive Intelligent User Interfaces with emotion recognition abilities that adapt and respond to the users' current emotional state by also modeling their preferences and personality. Her research involves eliciting emotions in a variety of contexts, using non-invasive wearable computers to collect the participants' physiological signals, mapping these signals to affective states, and building adaptive interfaces to adapt appropriately to the current sensed data and context. She is a member of the American Association for Artificial Intelligence and of the Association for Computing Machinery and she has published multiple scientific articles.

A.2 Publications

A.2.1 Journal Publications

1. **C. L. Lisetti, F. Nasoz (In press)** Using Non-invasive Wearable Computers to Recognize Human Emotions from Physiological Signals. *European Journal on Applied Signal Processing - Special Issue on Multimodal Human-Computer Interaction* (August 2004).
2. **F. Nasoz, K. Alvarez, C. L. Lisetti, and N. Finkelstein. (2003)** Emotion Recognition from Physiological Signals for Presence Technologies. *International Journal of Cognition, Technology, and Work - Special Issue on Presence*, Vol. 6(1).
3. **C. L. Lisetti, F. Nasoz, C. LeRouge, O. Ozyer, and K. Alvarez (2003)** Developing Multimodal Intelligent Affective Interfaces for Tele-Home Health Care. *International Journal of Human-Computer Studies - Special Issue on Applications of Affective Computing in Human-Computer Interaction*, Vol. 59 (1-2):245-255.

A.2.2 International Refereed Conferences

1. **C. L. Lisetti, F. Nasoz (2005, Invited to)** Multimodal non-invasive emotion recognition for enhanced Human-Computer Interaction. *11th International Conference on Human-Computer Interaction*, July, Las Vegas, Nevada USA.
2. **C. L. Lisetti, F. Nasoz, N. Yu. (2003)** Visualizing User's Emotions for Enhanced Human Computer Interaction. In *Proceedings of the International Conference on Visual Languages and Computing 2003*, (Miami, FL, September 2003).
3. **F. Nasoz, C. L. Lisetti, K. Alvarez, and N. Finkelstein. (2003)** Emotion Recognition from Physiological Signals for User Modeling of Affect. In *Proceedings of User Modeling (UM) 03 Workshop "Assessing and Adapting to User Attitudes and Affect: Why, When and How?"*
4. **F. Nasoz, O. Ozyer, C. L. Lisetti, and N. Finkelstein. (2002)** Multimodal Affective Driver Interfaces for Future Cars. In *Proceedings of ACM Multimedia International Conference*, (Juan les Pins, France, December 2002).
5. **C. L. Lisetti and F. Nasoz (2002)** MAUI: A Multimodal Affective User Interface. In *Proceedings of ACM Multimedia International Conference*, (Juan les Pins, France, December 2002).
6. **C. L. Lisetti, C. Lerouge, and F. Nasoz (2002)** The MAUI Project: Building Multimodal Affective User Interfaces for Everyone. In *Proceedings of the AAAI 18th*

National Conference Workshop on Automation as Caregiver: The Role of Intelligent Technology in Elder Care, Menlo Park, CA: AAAI Press, (Alberta, Canada, July, 2002).

7. **C. L. Lisetti and F. Nasoz (2002)** Application of the Multilevel Process Theory of Emotion to User-Modeling. In *Proceedings of the 15th International Florida Artificial Intelligence Research Society Conference (FLAIRS'02)*, Menlo Park, CA: AAAI Press, (Pensacola, FL, May 2002).

A.2.3 Articles in Progress

1. **F. Nasoz and C. L. Lisetti** Enhancing Driving Safety via Non-invasive Emotion Recognition and Intelligent Interface Adaptation. *International Journal of Human-Computer Studies*.
2. **F. Nasoz and C. L. Lisetti** User Modeling for Multimodal Intelligent Driving Safety Interfaces. *International Journal of User Modeling and User-Adapted Interaction*

APPENDIX B

FORMS AND QUESTIONNAIRES USED IN THE

EXPERIMENTS

B.1 Forms and Questionnaires for Emotion Elicitation with Movie Clips Experiment

B.1.1 IRB Consent Form

Consent Form

11/15/2003

The primary investigators of this study are Fatma Nasoz and Kaye Alvarez, graduate students at UCF working under the supervision of faculty member, Dr. Christine L. Lisetti. You are being asked to participate in a BodyMedia and/or Procomp+ experiment designed to measure participants' physiological signals (or body signals) while they are performing specific activities such as exercising, playing a computer game, driving a simulation truck, watching movie clips or movies, problem solving, looking at pictures, and/or listening to music. These activities are designed to elicit a variety of emotions such as sadness, frustration, and amusement. As a result, you may briefly experience discomfort; however, your experiences should not be overwhelming. This research project was designed solely for research purposes and no one except the research team will have access to any of your responses. All responses will be kept confidential. Your identity will be kept confidential using a numerical coding system. With your permission, this research session may be videotaped, and portions of the videos may be included in presentations.

Your participation in this project is voluntary. You do not have to answer any question(s) that you do not wish to answer. Please be advised that you may choose not to participate in

this research, and you may withdraw from the experiment at any time without consequence. Furthermore, if any of the activities make you uncomfortable, you may skip that activity and continue at the start of the next activity.

This experiment session will last approximately one hour. There are no anticipated risks associated with participation.

If you have any questions or comments about this research, please contact Fatma Nasoz or Dr. Christine Lisetti, College of Engineering and Computer Science, University of Central Florida. Questions or concerns about research participants' rights may be directed to the UCFIRB office, University of Central Florida Office of Research, Orlando Tech Center, 12443 Research Parkway, Suite 207, Orlando, FL 32826. The phone number is (407) 823-2901.

Sincerely,

Fatma Nasoz Kaye Alvarez

Participant's Signature: _____

Date: _____

_____I have read the procedure described above.

_____I voluntarily agree to participate in the BodyMedia Experiment and I have received a copy of this description.

_____I would like to receive a copy of the procedure described above

_____I would not like to receive a copy of the procedure described above

If you believe you have been injured during participation in this research project, you may file a claim against the State of Florida by filing a claim with the University of Central Florida's Insurance Coordinator, Purchasing Department, 4000 Central Florida Boulevard, Suite 360, Orlando, FL 32816, (407) 823-2661. University of Central Florida is an agency of the State of Florida and that the university's and state's liability for personal injury or property damage is extremely limited under Florida law. Accordingly, the university's and the state's ability to compensate you for any personal injury or property damage suffered during this research project is very limited.

Information regarding your rights as a research volunteer may be obtained from:

Chris Grayson

Institutional Review Board (IRB)

University of Central Florida (UCF)

12443 Research Parkway, Suite 207

Orlando, Florida 32826-3252

Telephone: (407) 823-2901

B.1.2 Pre-Study Questionnaire

PRE-STUDY QUESTIONNAIRE

Experiment No: _____

Participant Name: _____

Please answer the following questions.

1. Which of the following best describes your current primary job/activity?

_____ Student (please list major) _____

_____ UCF Professor (please list department) _____

_____ UCF Employee (other than professor) _____

_____ Other (please explain) _____

2. What is your gender?

_____ Female _____ Male

3. How old are you?

_____ 18-25

_____ 26-40

_____ 41-55

_____ over 55

4. What is your ethnicity?

_____ Caucasian

_____ African American

_____ Asian

——- Hispanic

Other —————

5. Do you wear glasses or contact lenses?

——- Yes Please describe your vision problem and correction method, for example, near-sighted, farsighted; bifocals, glasses, contact lenses. —————

——- No

6. Are you color blind?

——- Yes

——- No

B.1.3 In-Study Questionnaire

IN-STUDY QUESTIONNAIRE

Experiment No: _____

Participant Name: _____

Question #1. Did you experience SADNESS during this section of the experiment?

_____Yes

_____No

If you answered yes, please rate the intensity of your sadness (6 being the highest)

_____1 _____2 _____3 _____4 _____5 _____6

Were there any other emotions you experienced at the same intensity level or higher?

____Yes ____No

If yes, please name the emotion(s) you felt_____.

Question #2. Did you experience ANGER during this section of the experiment?

_____Yes

_____No

If you answered yes, please rate the intensity of your anger (6 being the highest)

_____1 _____2 _____3 _____4 _____5 _____6

Were there any other emotions you experienced at the same intensity level or higher?

____Yes ____No

If yes, please name the emotion(s) you felt_____.

Question #3. Did you experience SURPRISE during this section of the experiment? —

—Yes

—No

If you answered yes, please rate the intensity of your surprise (6 being the highest)

—1 —2 —3 —4 —5 —6

Were there any other emotions you experienced at the same intensity level or higher?

—Yes —No

If yes, please name the emotion(s) you felt_____.

Question #4. Did you experience FEAR during this section of the experiment?

—Yes

—No

If you answered yes, please rate the intensity of your fear (6 being the highest)

—1 —2 —3 —4 —5 —6

Were there any other emotions you experienced at the same intensity level or higher?

—Yes —No

If yes, please name the emotion(s) you felt_____.

Question #5. Did you experience FRUSTRATION during this section of the experiment?

—Yes

—No

If you answered yes, please rate the intensity of your frustration (6 being the highest)

—1 —2 —3 —4 —5 —6

Were there any other emotions you experienced at the same intensity level or higher?

—Yes —No

If yes, please name the emotion(s) you felt_____.

Question #6. Did you experience AMUSEMENT during this section of the experiment?

——Yes

——No

If you answered yes, please rate the intensity of your amusement (6 being the highest)

——1 ——2 ——3 ——4 ——5 ——6

Were there any other emotions you experienced at the same intensity level or higher?

—Yes —No

If yes, please name the emotion(s) you felt_____.

B.2 Forms and Questionnaires for Driving Simulator

Experiment in Virtual Reality

B.2.1 IRB Consent Form

Consent Form

03/01/2004

The principal investigator of this study is Fatma Nasoz, PhD candidate at UCF working under the supervision of faculty member, Dr. Christine L. Lisetti. You are being asked to participate in a BodyMedia and/or Procomp+ experiment designed to measure participants' physiological signals (or body signals) while they are performing specific activities such as exercising, playing a computer game, driving a truck and/or car simulator, watching movie clips or movies, problem solving, looking at pictures, and/or listening to music. These activities are designed to elicit a variety of emotions such as sadness, frustration, and amusement. You might watch movie clips and/or encounter traffic scenarios that might be upsetting. As a result, you may briefly experience discomfort; however, your experiences should not be overwhelming. This research project was designed solely for research purposes and no one except the research team will have access to any of your responses. All responses will be kept confidential. Your identity will be kept confidential using a numerical coding system. With your permission, this research session may be videotaped, and portions of the videos may be included in presentations. The video tapes will be stored in a locked filing cabinet and only the researcher(s) will have access to them.

Your participation in this project is voluntary. You do not have to answer any question(s) that you do not wish to answer. Please be advised that you may choose not to participate in this research, and you may withdraw from the experiment at any time without consequence. Furthermore, if any of the activities make you uncomfortable, you may skip that activity and continue at the start of the next activity.

This experiment session will last approximately one hour. There is minimal risk in participating due to the sensitive information that you will be viewing, which may cause emotional discomfort.

If you have any questions or comments about this research, please contact Fatma Nasoz or Dr. Christine Lisetti, College of Engineering and Computer Science, University of Central Florida. Questions or concerns about research participants' rights may be directed to the UCFIRB office, University of Central Florida Office of Research, Orlando Tech Center, 12443 Research Parkway, Suite 207, Orlando, FL 32826. The phone number is (407) 823-2901.

Sincerely,

Fatma Nasoz

Participant's Signature: _____

Date: _____

_____I have read the procedure described above. _____I voluntarily agree to participate in the BodyMedia Experiment and I have received a copy of this description.

_____I would like to receive a copy of the procedure described above

_____I would not like to receive a copy of the procedure described above

——I would like my video tape(s) to be used for this study and other possible research studies, and agree to let the researcher use my tape(s) for possible presentation purposes in the future.

——I would like my video tape(s) to be used for this study and other possible research studies, but not for presentations.

——I would like my video tape(s) to be used only for this study and then destroyed by the researcher at the end of this study.

If you believe you have been injured during participation in this research project, you may file a claim against the State of Florida by filing a claim with the University of Central Florida's Insurance Coordinator, Purchasing Department, 4000 Central Florida Boulevard, Suite 360, Orlando, FL 32816, (407) 823-2661. University of Central Florida is an agency of the State of Florida and that the university's and state's liability for personal injury or property damage is extremely limited under Florida law. Accordingly, the university's and the state's ability to compensate you for any personal injury or property damage suffered during this research project is very limited.

Information regarding your rights as a research volunteer may be obtained from:

Chris Grayson

Institutional Review Board (IRB)

University of Central Florida (UCF)

12443 Research Parkway, Suite 207

Orlando, Florida 32826-3252

Telephone: (407) 823-2901

B.2.2 Pre-Study Questionnaire

DRIVING SIMULATOR EXPERIMENT PRE-STUDY QUESTIONNAIRE

Experiment No:_____ Participant Name:_____ Date:_____

Please answer the following questions.

1. Which of the following best describes your current primary job/activity?

_____ Student (please list major)_____

_____ UCF Professor (please list department)_____

_____ UCF Employee (other than professor)_____

_____ Other (please explain)_____

2. What is your gender?

_____ Female _____ Male

3. How old are you?

_____ 18-25

_____ 26-40

_____ 41-55

_____ over 55

4. How long have you had your driver's license?_____

5. How often do you drive?

_____ Few times a day

_____ Few times a week

——Few times a year

——Never

6. Do you wear glasses or contact lenses?

——Yes Please describe your vision problem and correction method, for example, near-sighted, farsighted; bifocals, glasses, contact lenses.——

—— No

B.2.3 Post-Study Questionnaire

DRIVING SIMULATOR EXPERIMENT POST-STUDY QUESTIONNAIRE

Experiment No: _____ Participant Name: _____ Date: _____

Question #1. Did you experience PANIC and/or FEAR during section 1 of the experiment?

PANIC: _____ Yes _____ No

FEAR: _____ Yes _____ No

If you answered yes for PANIC, please rate the intensity of your panic (6 being the highest)

_____ 1 _____ 2 _____ 3 _____ 4 _____ 5 _____ 6

If you answered yes for FEAR, please rate the intensity of your fear (6 being the highest)

_____ 1 _____ 2 _____ 3 _____ 4 _____ 5 _____ 6

Were there any other emotions you experienced at the same intensity level or higher?

_____ Yes _____ No

If yes, please name the emotion(s) you felt _____

Question #2. Did you experience FRUSTRATION and/or FRUSTRATION during section 2 of the experiment?

FRUSTRATION: _____ Yes _____ No

FRUSTRATION: _____ Yes _____ No

If you answered yes for PANIC, please rate the intensity of your frustration (6 being the highest)

——1 ——2 ——3 ——4 ——5 ——6

If you answered yes for FEAR, please rate the intensity of your anger (6 being the highest)

——1 ——2 ——3 ——4 ——5 ——6

Were there any other emotions you experienced at the same intensity level or higher?

——Yes ——No

If yes, please name the emotion(s) you felt _____

Question #3. Did you experience BOREDOM during section 3 of the experiment?

——Yes ——No

If you answered yes, please rate the intensity of your boredom (6 being the highest)

——1 ——2 ——3 ——4 ——5 ——6

Were there any other emotions you experienced at the same intensity level or higher?

——Yes ——No

If yes, please name the emotion(s) you felt _____.

Question #4. How realistic did you find the driving simulator (level of being realistic increases from 1 to 6)

——1 ——2 ——3 ——4 ——5 ——6

Question #5. How often do you get frustrated or angry while driving? (frequency increases from 1 to 6, 1 being never, 6 being always)

——1 ——2 ——3 ——4 ——5 ——6

Question #6. How often do you get panicked or fearful while driving? (frequency increases from 1 to 6, 1 being never, 6 being always)

——1 ——2 ——3 ——4 ——5 ——6

Question #7. How often do you get bored while driving? (frequency increases from 1 to 6, 1 being never, 6 being always)

——1 ——2 ——3 ——4 ——5 ——6

LIST OF REFERENCES

- [ADL99] W. Ark, D. C. Dryer, and D. J. Lu. *Human-Computer Interaction: Ergonomics and User Interfaces*, pp. 818–823. Lawrance Erlboun Assoc., 1999.
- [AG96] W. Jr. Arthur and W. G. Graziano. “The five-factor model, constienciousness, and driving accident involvement.” *Journal of Personality*, **64**:593–618, 1996.
- [ARC96] A. Allen, L. Roman, R. Cox, and B. Cardwell. “Home health visits using a cable television network: User satisfaction.” *Journal of Telemedicine and Telecare*, **2**:92–94, 1996.
- [Bay00] A. L. Baylor. “Beyond butlers: Intelligent agents as mentors.” *Journal of Educational Computing and Research*, **22**:373–382, 2000.
- [BGC01] L. Feldman Barrett, J. J. Gross, T. Conner Christensen, and M. Benvenuto. “Knowing what you’re feeling and knowing what to do about it: Mapping the relation between emotion differentiation and emotion regulation.” *Cognition and Emotion*, **15**:713–724, 2001.
- [Bir70] R. Birdwhistle. *Kinesics and Context: Essays on Body Motion and Communication*. University of Pennsylvania Press, 1970.
- [BJB02] T. Barker, S. Jones, J. Britton, and D. Messer. “The Use of a Co-operative Student Model of Learner Characteristics to Configure a Multimedia Application.” *User Modeling and User-Adapted Interaction*, **12**:207–241, 2002.
- [BL02] N. Bianchi and C. L. Lisetti. “Modeling Multimodal Expression of User’s Affective Subjective Experience.” *User Modeling and User-Adapted Interaction*, **12**(1):49–84, 2002.
- [Bow81] G. Bower. “Mood and Memory.” *American Psychologist*, **36**(2):129–148, 1981.
- [BP00] D. Billsus and M. J. Pazzani. “User Modeling for Adaptive News Access.” *User Modeling and User-Adapted Interaction*, **10**:147–180, 2000.
- [CAA01] M. E. Crosby, B. Auernheimer, C. Aschwanden, and C. Ikehara. “Physiological Data Feedback for Application in Distance Education.” In *Proceedings of Workshop on Perceptual User Interfaces*, Orlando FL, USA, November 15-16 2001.

- [CB01] J. S. Casanueva and E. H. Blake. “The Effects of Avatars on Co-presence in a Collaborative Virtual Environment.” Technical report, Department of Computer Science, University of Cape Town, South Africa, 2001.
- [Cho91] N. Chovil. “Discourse-Oriented Facial Displays in Conversation.” *Research on Language and Social Interaction*, **25**:163–194, 1991.
- [CKC96] T. M. Crist, S. B. Kaufman, and K. R. Crampton. “Home Telemedicine: A Home Health Care Agency Strategy for Maximizing Resources.” *Home Health Care Manage Practice*, **8**:1–9, 1996.
- [CLN00] J.A. Colquitt, J.A. LePine, and R.A. Noe. “Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research.” *Journal of Applied Psychology*, **85**:678–707, 2000.
- [Cly77] D. M. Clynes. *Sentics: The Touch of the Emotions*. Anchor Press, Doubleday, 1977.
- [CM92] P.T. Costa and R. R. McCrae. *The Revised NEO Personality Inventory (NEO PI-R) professional manual*. Psychological Assessment Resources, Odessa, FL, 1992.
- [CMS00] A. Corbett, M. McLaughlin, and S. C. Scarpinato. “Modeling Student Knowledge: Cognitive Tutors in High School and College.” *User Modeling and User-Adapted Interaction*, **10**:81–108, 2000.
- [CNY00] D. F. Cellar, Z. C. Nelson, and C. M. Yorke. “The Five-Factor Model and Driving Behavior: Personality and Involvement in Vehicular Accidents.” *Psychological Results*, **86**:454–456, 2000.
- [Con03] C. Conati. “Probabilistic Assessment of User’s Emotions in Educational Games.” *Journal of Applied Artificial Intelligence*, 2003.
- [Cor] Narsys Software Corp. “www.natica.com.”.
- [CR03] V. Carofiglio and F. de Rosis. “Combining Logical with Emotional Reasoning in Natural Argumentation.” In *Proceedings of 9th International Conference on User Modeling - Assessing and Adapting to User Attitudes and Affect: Why, When, and How?*, pp. 9–15, Pittsburgh, PA, June 2003.
- [CVD97] C. Collet, E. Vernet-Maury, G. Delhomme, and A. Dittmar. “Autonomic Nervous System Response Patterns Specificity to Basic Emotions.” *Journal of the Autonomic Nervous System*, **62**(1-2):45–57, 1997.
- [Dam94] A. Damasio. *Descartes’ Error*. Avon Books, New-York, NY, 1994.

- [DC00] A. W. Darkins and M. A. Carey. *Telemedicine and Telehealth: Principles, Policies Performance and Pitfalls*. Springer Publishing Company, Inc., New York, NY, 2000.
- [DKF00] C. Dillon, E. Keogh, J. Freeman, and J. Davidoff. “Aroused and Immersed: The Psychophysiology of Presence.” In *Proceedings of 3rd International Workshop on Presence*, Delft University of Technology, Delft, The Netherlands, March 2000.
- [DT92] D. Derryberry and D. Tucker. “Neural Mechanisms of Emotion.” *Journal of Consulting and Clinical Psychology*, **60**(3):329–337, 1992.
- [EF75] P. Ekman and W. V. Friesen. *Unmasking the Face: A Guide to Recognizing Emotions from Facial Expressions*. Prentice Hall, Inc, New Jersey, 1975.
- [Ekm89] P. Ekman. *Handbook of Social Psychophysiology*, pp. 143–146. John Wiley, Chichester, 1989.
- [ELF83] P. Ekman, R. W. Levenson, and W. V. Friesen. “Autonomic Nervous System Activity Distinguishes Between Emotions.” *Science*, **221**(4616):1208–1210, 1983.
- [FA00] J. Freeman and S.E. Avons. “Focus Group Exploration of Presence through Advanced Broadcast Services.” In *Proceedings of the SPIE, Human Vision and Electronic Imaging V*, 2000.
- [FC00] J. Fink and A. Cobsa. “A Review and Analysis of Commercial User Modeling Servers for Personalization on the World Wide Web.” *User Modeling and User-Adapted Interaction*, **10**:209–249, 2000.
- [FI83] A. Fridlund and C.E. Izard. *Social Psychophysiology: A Sourcebook*, chapter Electromyographic Studies of Facial Expressions of Emotions and Patterns of Emotions, pp. 243–286. Guilford Press, New York, 1983.
- [Fri86] N. H. Frijda. *The Emotions*. Cambridge University Press, New York, 1986.
- [FS02] R. Fuller and J. A. Santos, editors. *Human Factors for Highway Engineers*. Elsevier Science, Oxford, UK, 2002.
- [GH03] C. Guinn and R. Hubal. “Extracting Emotional Information From the Text of Spoken Dialog.” In *Proceedings of 9th International Conference on User Modeling - Assessing and Adapting to User Attitudes and Affect: Why, When, and How?*, pp. 23–27, Pittsburgh, PA, June 2003.
- [GL95] J.J. Gross and R. W. Levenson. “Emotion elicitation using films.” *Cognition and Emotion*, **9**:87–108, 1995.

- [GL97] J.J. Gross and R. W. Levenson. "Hiding Feelings: The Acute Effects of Inhibiting Negative and Positive Emotions." *Journal of Abnormal Psychology*, **106**(1):95–103, 1997.
- [Gol95] D. Goleman. *Emotional Intelligence*. Bantam Books, New York, 1995.
- [HBC92] T. Hewett, R. Baecker, S. Card, J. Gasen, M. Mantei, G. Perlman, G. Strong, and W Verplank. *ACM SIGCHI Curricula for Human-computer Interaction*. Report of the ACM SIGCHI Curriculum Development Group, ACM, 1992.
- [Hea00] J. Healey. *Wearable and Automotive Systems for Affect Recognition from Physiology*. PhD thesis, Massachusetts Institute of Technology, May 2000.
- [HM94] M. T. Hagan and M. B. Menhaj. "Training Feedforward Networks with the Marquardt Algorithm." *IEEE Transactions on Neural Networks*, **5**(6):989–993, 1994.
- [HM02] E. Hudlicka and M. D. McNeese. "Assessment of User Affective and Belief States for Interface Adaptation: Application to an Air Force Pilot Task." *User Modeling and User-Adapted Interaction*, **12**:1–47, 2002.
- [IJs02] W.A. IJsselsteijn. "Elements of a multi-level theory of presence: Phenomenology, mental processing and neural correlates." In *Proceedings of PRESENCE*, pp. 245–259, Universidade Fernando Pessoa, Porto, Portugal, October 9-11 2002.
- [IRF00] W.A. IJsselsteijn, H. de Ridder, J. Freeman, and S.E. Avons. "Presence: Concept, determinants and measurement." In *Proceedings of the SPIE, Human Vision and Electronic Imaging V*, 2000.
- [Jam00] L. James. *Road Rage and Aggressive Driving*. Prometheus Books, Amherst, NY, 2000.
- [Kah73] D. Kahneman. *Attention and Effort*, chapter Arousal and attention, pp. 28–49. Prentice-Hall, Englewood Cliffs, N.J., 1973.
- [Kal00] R. S. Kalawsky. "The Validity of Presence as a Reliable Human Performance Metric in Immersive Environments." In *Proceedings of 3rd International Workshop on Presence*, Delft, Netherlands, 2000.
- [Kod96] T. Koda. "Agents with Faces: the Effects of Personification." In *Proceedings of Human-Computer Interaction'96*, London, 1996.
- [LD97] M. Lombard and T. Ditton. "At the Heart of It All: The Concept of Presence." *Journal of Computer Mediated Communication*, **3**(2), 1997.
- [Led92] J. Ledoux. "Brain Mechanisms of Emotion and Emotional Learning." *Current Opinion in Neurobiology*, **2**:191–197, 1992.

- [LEF90] R. W. Levenson, P. Ekman, and W. V. Friesen. "Voluntary Facial Action Generates Emotion-Specific Autonomic Nervous System Activity." *Psychophysiology*, **27**:363–384, 1990.
- [LHE92] R. W. Levenson, K. Heider, P. Ekman, and W. V. Friesen. "Emotion and Autonomic Nervous-System Activity in the Minangkabau of West Sumatra." *Journal of Personality and Social Psychology*, **62**(6):972–988, 1992.
- [Lis99] C. L. Lisetti. "A User Model of Emotion-Cognition." In *Proceedings of the UM'99 Workshop on Attitude, Personality, and Emotions in User-Adapted Interaction*, Banff, Canada, June 1999.
- [LN02] C. L. Lisetti and F. Nasoz. "MAUI: A Multimodal Affective User Interface." In *Proceedings of the ACM Multimedia International Conference*, Juan les Pins, France, December 2002.
- [LNL03] C. L. Lisetti, F. Nasoz, C. Lerouge, O. Ozyer, and Alvarez K. "Developing Multimodal Intelligent Affective Interfaces for Tele-Home Health Care." *International Journal of Human-Computer Studies Special Issue on Applications of Affective Computing in Human-Computer Interaction*, **59**(1-2):245–255, 2003.
- [LO86] J. T. Lanzetta and S. P. Orr. "Excitatory Strength of Expressive Faces - Effects of Happy and Fear Expressions and Context on the Extinction of a Conditioned Fear Response." *Journal of Personality and Social Psychology*, **50**(1):190–194, 1986.
- [LR99] J. Larson and C. Rodriguez. *Road Rage to Road-Wise*. Tom Doherty Associates, Inc., New york, NY, 1999.
- [LW89] V.E. Lewis and R.N. Williams. "Mood-congruent vs. mood-state-dependent learning: Implications for a view of emotion." *Special issue of the Journal of Social Behavior and Personality*, **4**:157–171, 1989.
- [Mar94] J.J. Martocchio. "Effects of conceptions of ability on anxiety, self-efficacy, and learning in training." *Journal of Applied Psychology*, **79**:819–825, 1994.
- [May01] M.T. Maybury. "Human Computer Interaction: State of the Art and Further Development in the International Context - North America (Invited talk)." In *International Status Conference on MTI Program*, Saarbruecken, Germany, October 26-27 2001.
- [MC02] E. Millan and J. L. Perez-de-la Cruz. "A Bayesian Diagnostic Algorithm for Student Modeling and its Evaluation." *User Modeling and User-Adapted Interaction*, **12**:281–330, 2002.
- [Min80] M. Minsky. "Telepresence." *Omni*, pp. 45–51, June 1980.

- [Mit97] T. M. Mitchell. *Machine Learning*. McGraw-Hill Companies Inc., 1997.
- [MPM03] J. Mori, H. Prendinger, S. Mayer, H. Dohi, and M. Ishizuka. “Using bio-signals to Track the Effects of a Character-based Interface.” In *Proceedings of 9th International Conference on User Modeling - Assessing and Adapting to User Attitudes and Affect: Why, When, and How?*, pp. 39–44, Pittsburgh, PA, June 2003.
- [Nic99] A. A. Nicol. *Presenting Your Findings: A Practical Guide for Creating Tables*. American Physiological Association, Washington, DC:, 1999.
- [NL00] C. Nass and K. M. Lee. “Does computer-generated speech manifest personality? An experimental test of similarity-attraction.” In *Proceedings of the CHI 2000 Conference*, The Hague, The Netherlands, April 1-6 2000.
- [NS89] A. F. Norcio and J. Stanley. “Adaptive Human-Computer Interfaces: A Literature Survey and Perspective.” *IEEE Transactions on Systems, Man, and Cybernetics*, **19**(2):399–408, 1989.
- [Pea88] J. Pearl. *Probabilistic Reasoning in Expert Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers, Inc., San Francisco, 1988.
- [PHV01] R. W. Picard, J. Healey, and E. Vyzas. “Toward Machine Emotional Intelligence Analysis of Affective Physiological State.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **23**(10):1175–1191, 2001.
- [Pic97] R. W. Picard. *Affective Computing*. MIT Press, 1997.
- [PS96] A. Pecchinenda and C. Smith. “Affective Significance of Skin Conductance Activity During Difficult Problem-solving Task.” *Cognition and Emotion*, **10**(5):481–503, 1996.
- [PSA00] D. Palomba, M. Sarlo, A. Angrilli, A. Mini, and L. Stegagno. “Cardiac Responses Associated with Affective Processing of Unpleasant Film Stimuli.” *International Journal of Psychophysiology*, **36**(1):45–57, 2000.
- [RB93] M. Riedmiller and H. Braun. “A direct adaptive method for faster backpropagation learning: The RPROP algorithm.” In *Proceedings of the IEEE International Conference on Neural Networks*, San Francisco, CA, 1993.
- [RG00] E.J. Rozell and W.L. Gardner. “Cognitive, motivation, and affective processes associated with computer-related performance: A path analysis.” *Computers in Human Behavior*, **17**:199–222, 2000.
- [RN96] B. Reeves and C. I. Nass. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge University Press, Cambridge, 1996.

- [Sel94] T. (1994) Selker. “A Teaching Agent that Learns.” *Communications of the ACM*, **37**(7):92–99, July 1994.
- [SFG74] G. E. Schwartz, P. L. Fair, P. S. Greenberg, Friedman M. J., and G. L. Klerman. “Facial Electromyography in Assessment of Emotion.” *Psychophysiology*, **11**(2):237–237, 1974.
- [SFK02] J. Scheirer, R. Fernandez, J. Klein, and R. W. Picard. “Frustrating the User on Purpose: A Step Toward Building an Affective Computer.” *Interacting with Computers*, **14**(2), 2002.
- [SHS00] M. Slater, J. Howell, A. Steed, D. P. Pertaub, M. Garau, and S. Springel. “Acting in Virtual Reality.” *ACM Collaborative Virtual Environments*, **16**:103–110, 2000.
- [Smi89] C. A. Smith. “Dimensions of Appraisal and Physiological-Response in Emotion.” *Journal of Personality and Social Psychology*, **56**(3):339–353, 1989.
- [SP96] R. Sinha and O. Parsons. “Multivariate Response Patterning of Fear and Anger.” *Cognition and Emotion*, **10**(2):173–198, 1996.
- [TKK01] M. P. Tarvainen, A. S. Koistinen, M. V. Korhonen, J. Partnen, and P. A. Karjalainen. “Analysis of Galvanic Skin Responses with Principal Components and Clustering Techniques.” *IEEE Transactions on Biomedical Engineering*, **48**(10):1071–1079, 2001.
- [TN93] A. Takeuchi and K. Nagao. “Communicative facial displays as a new conversational modality.” In *Proceedings of the INTERCHI’93 Conference on Human Factors in Computing Systems*, pp. 187–193, Amsterdam, The Netherlands, 1993.
- [Tra03] Department for Transport. “International review of the individual factors contributing to driving behaviour.” Technical report, Department for Transport, October 2003.
- [VCL86] S. C. Vrana, B. N. Cuthbert, and P. J. Lang. “Fear Imagery and Text Processing.” *Psychophysiology*, **23**(3):247–253, 1986.
- [Vra93] S. Vrana. “The Psychophysiology of Disgust: Differentiating Negative Emotional Contexts with Facial EMG.” *Psychophysiology*, **30**:279–286, 1993.
- [War97] I. Warner. “Telemedicine Applications for Home Health Care.” *Journal of Telemedicine and Telecare*, **3**:65–66, 1997.
- [WB95] P. Warr and D. Bunce. “Trainee characteristics and the outcomes of open learning.” *Personnel Psychology*, **48**:347–475, 1995.

- [WCP86] R. A. Wright, R. J. Contrada, and M. J. Patane. “Task-Difficulty, Cardiovascular-Response, and the Magnitude of Goal Valence.” *Journal of Personality and Social Psychology*, **51**(4):837–843, 1986.
- [WD93] R.A. Wright and J.C. Dill. “Blood Pressure responses and incentive appraisals as a function of perceived ability and objective task demand.” *Psychophysiology*, **30**(2):152–160, 1993.
- [WFN02] J. R. Warren, H. K. Frankel, and J. T. Noone. “Supporting special-purpose health care models via adaptive interfaces to the web.” *Interacting with Computers*, **14**:251–267, 2002.
- [WSS94] J.H. Walker, L. Sproull, and R. Subramani. “Using a Human Face in an Interface.” In *Proceedings of Human Factors in Computing Systems*, pp. 85–91, Reading, MA, CHI’94, 1994.
- [Zaj84] R. Zajonc. “On the Primacy of Affect.” *American Psychologist*, **39**:117–124, 1984.