In *Proceedings of 11th International Conference on Human Computer Interaction*, Las Vegas, NV, USA, July 2005

Affective Intelligent Car Interfaces with Emotion Recognition

Christine L. Lisetti

Department of Multimedia Communications,
Institut Eurecom
Sophia-Antipolis, France
christine.lisetti@eurecom.fr

Fatma Nasoz

School of Computer Science,
University of Central Florida
Orlando, FL
fatma@cs.ucf.edu

Abstract

In this paper, we uncover a new potential application for multi-media technologies: affective intelligent car interfaces for enhanced driving safety. We also describe the experiment we conducted in order to map certain physiological signals (galvanic skin response, heart beat, and temperature) to certain driving-related emotions and states (Frustration/Anger, Panic/Fear, and Boredom/Sleepiness). We demonstrate the results we obtained and describe how we use these results to facilitate a more natural Human-Computer Interaction in our Multimodal Affective Car Interface for the drivers of the future cars.

1 Introduction and Motivation

Humans are social beings that emote and their cognition is affected by their emotions. Emotions influence various cognitive processes in humans, including perception and organization of memory (Bower, 1981), categorization and preference (Zajonc, 1984), goal generation, evaluation, and decision-making (Damasio, 1994), strategic planning (Ledoux, 1992), focus and attention (Derryberry & Tucker, 1992), motivation and performance (Colquitt, LePine, & Noe, 2000), intention (Frijda, 1986), communication (Birdwhistle, 1970; Ekman & Friesen 1975; Chovil 1991), and learning (Goleman, 1995). Previous studies also suggest that people emote while they are interacting with computers (Reeves & Nass, 1996). Given the involvement of emotions in human-computer interaction and the strong interface between emotions and cognition, machine perception needs to be able to capture such phenomenon and respond accordingly in order to enhance our everyday digital tools.

An important everyday activity for people is driving, and yet again research suggests that drivers do emote while they are driving in their cars and their driving is affected by their emotions (James, 2000). Aggressive driving in the United States results in 425,000 deaths and 35 million injuries per decade and it approximately costs $250 billion per year (James, 2000). The inability to manage one's emotions while driving is often identified as one of the major causes for accidents. *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 (James, 2000). Other states that lead to negative effects are *frustration*, *anxiety*, *fear*, and *stress* (Fuller & Santos, 2002). In order to be a safer driver on the highways, a person needs to be better aware of her emotions and possess the ability to manage them effectively (James, 2000).

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 (James, 2000).

James (James, 2000) and Larson (Larson, 1999) discussed techniques for drivers to manage their anger including relaxation techniques to reduce physical arousal and mental reappraisal of the situation. Our aim in creating an affective intelligent car interface is to enhance driving safety by facilitating a natural human-computer interaction with the driver and help her to be better aware of her emotional state during driving. For example, when the intelligent system recognizes the anger or rage of a driver it might suggest the driver to perform a breathing exercise (Larson, 1999). 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 the system and the driver, and taking the precautions mentioned above automatically depending on the drivers' personal preferences would increase the drivers' feeling that there exists a real person in the car with them to assist them while they drive.

Figure 1, which was originally proposed by Lisetti (Lisetti, 1999), shows the overall architecture of the system that would recognize the driver's current affective state and respond accordingly (Bianchi & Lisetti, 2002; Lisetti & Nasoz, 2002). Affective state of a driver can be assessed by interpreting both the mental and the physiological components of the particular emotion experienced by the driver. Physiological components can be identified and collected from observing the driver 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).

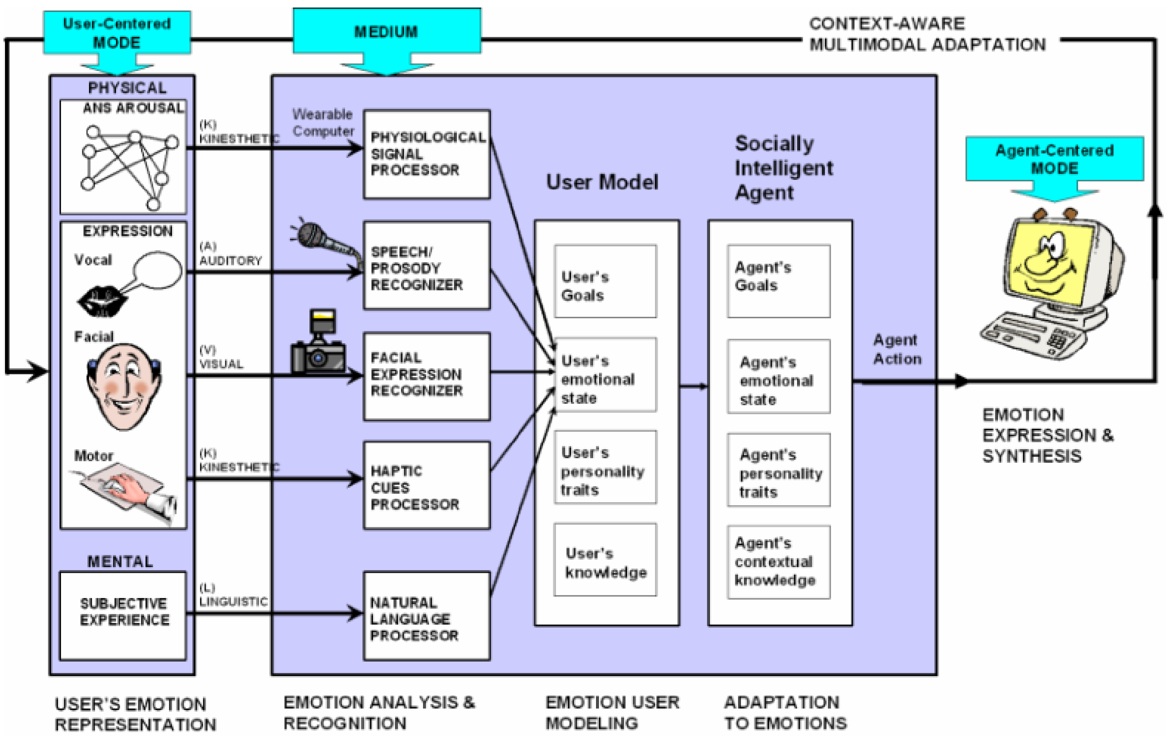

Figure 1 Human Multimodal Affect Expression matched with Multimedia Computer Sensing

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.

Currently we are focusing our research on recognizing the affective states of the drivers by collecting and analyzing their physiological signals. The next section discusses the previous studies on physiological emotion recognition and the preliminary experiments designed and conducted by us to find a relationship between certain physiological signals and emotions.

2 Related Research and Our Preliminary Experiments

2.1 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 (Ekman, Levenson, & Friesen, 1983; Gross & Levenson, 1997). 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.* (Collet, Vernet-Maury, Delhomme, & Dittmar, 1997) 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.* (Picard, Healey, & Vyzas, 2001) used personalized imagery and emotionally loaded pictures to elicit 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.

Healey's research (Healey, 2000) was focused on recognizing stress levels of drivers by measuring and analyzing their physiological signals (skin conductance, heart activity, respiration, and muscle activity). During the experiment participants of this study drove in a parking garage, in a city, and on a highway. Results showed that the drivers' stress could be recognized as being *rest* (*i.e.* resting in the parking garage), *city* (*i.e.* driving in the Boston streets), and *highway* (*i.e.* two lane merge on the highway) with 96% accuracy.

2.2 Our Preliminary Emotion Elicitation and Recognition Experiments

In our emotion elicitation experiment we used movie clips and difficult mathematical questions to elicit six emotions: *sadness, anger, surprise, fear, frustration,* and *amusement* and a non-invasive wireless wearable computer – BodyMedia SenseWear Armband (Figure 2) – to collect the physiological signals of our participants: *galvanic skin response, heart rate,* and *temperature*.

Figure 2 BodyMedia SenseWear Armband

Mathematical questions were used to elicit frustration and movie clips were used to elicit the other five emotions. Movie clips were chosen by conducting a pilot study that was guided by the previous research of Gross and

Levenson (Gross & Levenson, 1995). Movie scenes resulting in high subject agreement in eliciting the target emotions at the end of our pilot study were chosen to elicit specific emotions.

After choosing the multimodal stimuli for emotion elicitation, movie clips and mathematical questions were presented to the participants in a power point slide show. The participants' physiological signals were collected while they were watching the slide show and their self reports were collected between each emotion elicitation session.

Collected physiological signals were normalized in order to minimize the individual differences among participants and four features (minimum, maximum, mean, and variance) were extracted from physiological data for each physiological data type (resulting in a total of 12 features). Three supervised learning algorithms were implemented to analyse these 12 features: k-Nearest Neighbor (KNN) (Mitchell, 1997), Discriminant Function Analyses (DFA) (Nicol & Pexman, 1999), and Marquardt Backpropagation (MBP) (Hagan & Menhaj, 1994) Algorithms. Overall, those three algorithms, KNN, DFA, and MBP could recognize emotions with 72.3%, 75.0%, and 84.1% accuracy respectively.

3 Driving Simulator Experiment in Virtual Reality

Designing, conducting, and analyzing the results of the Emotion Elicitation Experiment discussed in the Section 2.2 showed that emotions experienced by people can be classified by analyzing their physiological signals. Designing and conducting an experiment that focuses on eliciting emotions that are related to a 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* n order to measure the participants' physiological signals while they were experiencing these emotions. Driving Simulator (Figure 3 - located in the new Engineering Building of University of Central Florida [UCF]) operating in Virtual Reality (Figure 4) 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 5 shows the control room of the simulator.

Figure 3 The Car Operating in Virtual Reality

Figure 4 Virtual Reality Environment

Figure 5 Control Room of the Driving Simulator

A scenario consisting series of events was designed and implemented to be run on the simulator to elicit panic/fear, frustration/anger, and boredom/sleepiness 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 (Figure 2). 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.

3.1 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 6).

Figure 6 Child Walking to the Middle of the Road

Frustration/anger: After hitting the child 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 to cross the street, who are carrying a big glass and at the same time talking to another man they meet on the road, thus blocking 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 7). 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 8). However, as soon as the lights turn to green for the driver, several pedestrians start passing across the road and the lights turn to red again before the driver gets a chance to turn left.

Figure 7 Garbage Truck Making a 3-point Turn

Figure 8 Waiting for the Pedestrians to Cross

After passing the pedestrians and starting to drive in a narrow road, a bus driver rides his bus right on to the participant's car like 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.

One of the biggest differences between Healey's work (Healey, 2000) and our research 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 as follows:

- 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.

3.2 Driving Simulator Experiment Setup

Sample: The sample included 41 undergraduate and graduate students enrolled in University of Central Florida (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 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 stopping at red lights and stop signs and not driving over the speed limit. (iii) The red and yellow arrows on the simulator screen would indicate 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 3.1 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 they experienced another emotion. 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 in real traffic on a 6-point scale (1 being never, 6 being always).

3.3 Emotion Recognition with Machine Learning

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

3.3.1 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. The number of data sets for each emotion is different from the total sample size because for some participants, collected data were not complete for every emotion. Data was stored and normalized and the same features stated in Section 2.2 were extracted from the 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.

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

The data was first analyzed with KNN and MBP algorithms. The neural network structure used with the Marquardt Backpropagation Algorithm was consisted of an input layer with 12 nodes, a hidden layer with 17 nodes, and an output layer with 3 nodes. Table 1 and Table 2 and report the classification accuracy of each emotion set with KNN and MBP respectively.

Table 1 Emotion Classification Accuracy with KNN for each Emotion

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

Table 2 Emotion Classification Accuracy with MBP for each Emotion

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

As can be seen from the table Table 2, the MBP algorithm was not as successful as it was in recognizing the six emotions elicited in Emotion Elicitation Experiment with Movie Clips (Section 2.2). For this reason Resilient Backpropagation (RBP) Algorithm (Riedmiller & Braun, 1993), which is another Neural Network Algorithm, was implemented.

Table 3 shows the emotion recognition accuracy for emotion by RBP algorithm and Table 4 reports the emotion classification accuracy of k-Nearest Neighbor, Marquardt Backpropagation and Resilient Backpropagation Algorithms for all emotions.

Table 3 Emotion Classification Accuracy with RBP for each Emotion

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

Table 4 Emotion Classification Accuracy with KNN, MBP, and RBP for all Emotions

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

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 emotional state of the user although she is not experiencing this state) obtained by interpreting the physiological signals of the user. Table 5 summarizes the results that can be obtained while performing emotion recognition.

Table 5 Consequences Related to Emotion Recognition

Emotion Experienced	Emotion Recognized	Consequence
YES	YES	Accurate Emotional State Recognition
YES	NO	False Negative
NO	YES	False Positive
NO	NO	Accurate Emotional State 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.

4 Conclusion and Future Work

In order to map physiological signals to driving-related emotions, a driving experiment was designed and conducted in a highly controlled virtual reality environment. A driving scenario that is consisted of 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.

All the experiments discussed in this article were conducted in totally controlled environments and during all those experiments, physiological data was analysed after the experiment was completed. An important next step to our research will be collecting physiological data during real-life situations and analysing them and performing emotion recognition in real-time. Another improvement to our study will be applying different feature extraction techniques and combining different pattern recognition algorithms for increased accuracy in emotion recognition.

References

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.

R. Birdwhistle. *Kinesics and Context: Essays on Body Motion and Communication*. University of Pennsylvania Press, 1970.

G. Bower. Mood and Memory. *American Psychologist*, 36(2):129-148, 1981.

N. Chovil. Discourse-Oriented Facial Displays in Conversation. *Research on Language and Social Interaction*, 25:163-194, 1991.

C. Collet, E. Vernet-Maury, G. Delhomme, and A. Dittmar. Autonomic Nervous System Response Patterns Specificity to Basic Emotions. *Journal of theAutonomic Nervous System*, 62(1-2):45-57, 1997.

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.

A. Damasio. *Descartes' Error*. Avon Books, New-York, NY, 1994.

D. Derryberry and D. Tucker. Neural Mechanisms of Emotion. *Journal of Consulting and Clinical Psychology*, 60(3):329-337, 1992.

P. Ekman and W. V. Friesen. *Unmasking the Face: A Guide to Recognizing Emotions from Facial Expressions*. Prentice Hall, Inc, New Jersey, 1975.

P. Ekman, R. W. Levenson, and W. V. Friesen. Autonomic Nervous System Activity distinguishes Between Emotions. *Science*, 221(4616):1208-1210, 1983.

N. H. Frijda. *The Emotions*. Cambridge University Press, New York, 1986.

R. Fuller and J. A. Santos, editors. *Human Factors for Highway Engineers*. Elsevier Science, Oxford, UK, 2002.

J.J. Gross and R. W. Levenson. Emotion elicitation using films. *Cognition andEmotion*, 9:87-108, 1995.

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.

D. Goleman. *Emotional Intelligence*. Bantam Books, New York, 1995.

M. T. Hagan and M. B. Menhaj. Training Feedforward Networks with the Marquardt Algorithm. *IEEE Transactions on Neural Networks*, 5(6):989-993, 1994.

J. Healey. *Wearable and Automotive Systems for Affect Recognition from Physiology*. PhD thesis, Massachusetts Institute of Technology, May 2000.

L. James. *Road Rage and Aggressive Driving*. Prometheus Books, Amherst, NY, 2000.

J. Larson and C. Rodriguez. *Road Rage to Road-Wise*. Tom Doherty Associates, Inc., New york, NY, 1999.

J. Ledoux. Brain Mechanisms of Emotion and Emotional Learning. *Current Opinion in Neurobiology*, 2:191-197, 1992.

C. L. Lisetti. A User Model of Emotion-Cognition. In *Proceedings of the UM'99Workshop on Attitude, Personality, and Emotions in User-Adapted Interaction*, Banf, Canada, June 1999.

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.

T. M. Mitchell. *Machine Learning*. McGraw-Hill Companies Inc., 1997.

A. A. Nicol and P. M. Pexman. *Presenting Your Findings: A Practical Guide for Creating Tables*. American Physiological Association, Washington, DC, 1999.

R. W. Picard, J. Healey, and E. Vyzas. Toward Machine Emotional Intelligence Analysis of Affective Physiological State. *IEEE Transactions on Pattern Analysisand Machine Intelligence*, 23(10):1175-1191, 2001.

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.

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.

R. Zajonc. On the Primacy of Affect. *American Psychologist*, 39:117-124, 1984.