

Emotion Recognition from Physiological Signals for User Modeling of Affect

Fatma Nasoz¹, Christine L. Lisetti¹, Kaye Alvarez², and Neal Finkelstein³

¹University of Central Florida, Department of Computer Science,
Orlando, FL 32816-2362 Tel: 1-407-823-{3931,3537}
{fatma, lisetti}@cs.ucf.edu

²University of Central Florida, Department of Psychology
Orlando, FL 32816 Tel: 1-407-823-3931
kayealvarez@earthlink.net

³ Operations Simulation Technology Center, 12423 Research Parkway,
Orlando, FL, 32826-3276 Tel: 1-407-384-5107
Neal_Finkelstein@peostri.army.mil

ABSTRACT

In this paper, we describe algorithms developed to analyze physiological signals associated with emotions, in order to recognize the affective states of users via non-invasive technologies. We propose a framework for modeling user's emotions from the sensory inputs and interpretations of our multi-modal system. We also describe examples of circumstances in which this kind of system can be applied to.

Keywords: Emotion recognition, affective intelligent user interfaces, user-models of emotions.

1. MOTIVATION

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 [1][8][25][30], news access [2], e-commerce [13], and health-care [31].

However, none of these conventional models includes a very important component of human intelligence: *affect* and *emotion*.

Emotions are essential for human cognition and they influence different aspects of people's lives including:

- **organization of memory and learning:** we recall an event better when we are in the same mood as when the learning occurred [4];
- **perception:** when we are happy, our perception is biased at selecting happy events, likewise for negative emotions [4];
- **categorization and preference:** familiar objects become preferred objects [32];
- **goal generation, evaluation, and decision-making:** patients who have damage in their frontal lobes (cortex communication with limbic system is altered) become unable to feel, which results in their complete dysfunctionality in real-life settings where they are unable to decide what is the next action they need to perform [9]. Normal emotional arousal, on the other hand, is

intertwined with goal generation and decision-making;

- **strategic planning:** when time constraints are such that quick action is needed (as in fear of a rattle snake), neurological shortcut pathways for deciding upon the next appropriate action are preferred over more optimal but slower ones [21];
- **focus and attention:** emotions restrict the range of cue utilization such that fewer cues are attended to [10];
- **motivation and performance:** an increase in emotional intensity causes an increase in performance, up to an optimal point (inverted U-curve Yerkes-Dodson Law);
- **intention:** not only are there positive consequences to positive emotions, but there are also positive consequences to negative emotions - they signal the need for an action to take place in order to maintain, or change a given kind of situation or interaction with the environment [14];
- **communication:** important information in a conversational exchange comes from body language [3], voice prosody and facial expression revealing emotional content [12], and facial displays connected with various aspects of discourse [5].
- **learning:** people are more or less receptive to the information to be learned depending their liking (of the instructor, or the visual presentation, or of how the feedback is given). Moreover, emotional intelligence is learnable [15].

Given the strong interface between affect and cognition on the one hand, and given the increasing versatility of computer agents on the other hand, the attempt to enable our computer tools to acknowledge affective phenomena rather than to remain blind to them appears desirable.

In the last five years, there has been a significant increase in number of attempts to build user models that include emotion and affect at some

level in the user model. Conati's [7] probabilistic user model, which is based on Dynamic Decision Network, represents the emotional state of the user interacting with an educational game, as well as her personality and goals. ABAIS, created as a rule-based system by Hudlicka and McNeese [19], assesses pilots' affective states and active beliefs and takes adaptive precautions to compensate for their negative affects. Klein et al.'s [20] interactive system responds to the user's self-reported frustration during an interaction with a computer game. All these systems are created to adapt to the user's affective state based on the current context, however none of them performs any emotion recognition stage. Our multimodal system that recognizes the emotional state of the user will be a good complement to these systems and other systems that model user's emotions.

Maybury [24] defines the purpose of human computer interaction as enabling users to perform complex task more quickly, with more accuracy and improving user satisfaction. New interfaces can dynamically adapt to the user, the current context and situation [24].

We aim at creating intelligent systems that achieve the goals of human-computer interaction defined by Maybury [24] by developing algorithms for real-time emotion recognition. For example, when our system recognizes a learner's frustration, confusion or boredom and adjusts the pace of the training accordingly, the learning task will be improved. Similarly, when our system recognizes the anger or rage of a driver and takes action to neutralize her negative emotions, the driver's safety will be enhanced. Finally when our system recognizes the depressive patterns of sadness telemedicine patients might experience in their homes and communicates these to the health-care providers monitoring them, the patients' satisfaction will be improved from the better treatment received.

2. OUR APPROACH

The first step in our approach to model the emotional state of the user focuses on recognizing the user's emotional state accurately. Our system is shown in Figure 1 and is designed to [22]:

- gather multimodal physiological signals and expressions of a specific emotion to recognize the user's emotional state;
- give feedback to the user about her/his emotional state;
- adapt its interface to the user's affective state based on the user's preferences of interaction in the current context or application.

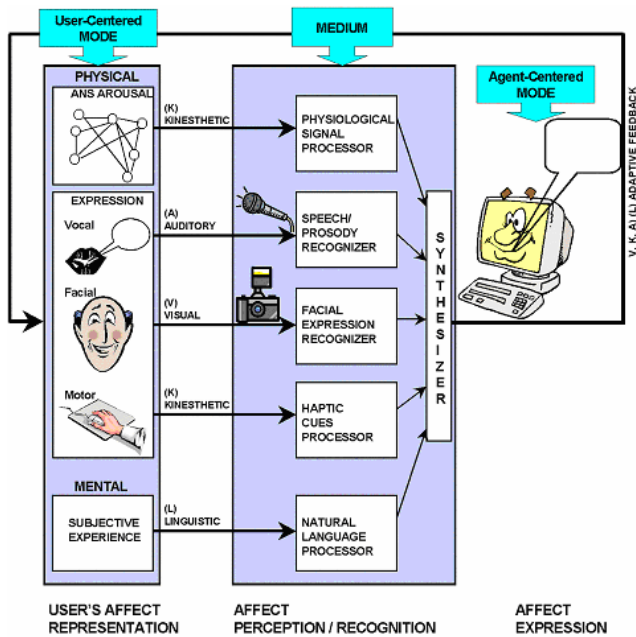


Figure 1: Human Multimodal Affect Expression matched with Multimedia Computer Sensing

Currently, we are working on recognizing the users' emotions by interpreting the physiological signals (temperature, heart rate, galvanic skin response) and mapping these physiological signals to their corresponding emotions.

3. RELATED WORK ON EMOTION ELICITATION AND RECOGNITION

3.1 Emotion Recognition from Physiological Signals Previous Research

Both manual analysis and statistical analysis of physiological signals collected during emotional experience have been implemented to understand the connection between emotions and physiological arousal.

3.1.1 Research with Manual Analysis

Ekman's study [11] used manual analysis to interpret the physiological signals (finger temperature, heart rate and skin conductance) that occurred when anger, fear, disgust, and happiness were elicited. The results demonstrated that anger, fear and sadness produced a larger increase in heart rate than disgust did, whereas anger produced a larger increase in finger temperature than fear did. It was also found that anger and fear produced larger heart rate than happiness, while fear and disgust produced larger skin conductance than happiness.

In Gross and Levenson [16] amusement, neutrality, and sadness were elicited by showing films. Skin conductance, inter-beat interval, pulse transit times and respiratory activation were measured. Their results showed that inter-beat interval increased for all three states, but for neutrality it was less than for amusement and sadness. Skin conductance increased after the amusement film, decreased after the neutral film and stayed the same after the sadness film.

3.1.2 Research with Statistical Analysis

Collet et al. [6] showed naturally and emotionally loaded pictures to participants to elicit happiness, surprise, anger, fear, sadness, and disgust and measured skin conductance, skin potential, skin resistance, skin blood flow, skin temperature, and instantaneous respiratory frequency. Statistical comparison of data signals was performed pairwise, where 6 emotions formed 15 pairs. Out of these 15 emotion-pairs, electrodermal responses distinguished 13 pairs, skin conductance ohmic

perturbation duration indices separated 10 pairs, and conductance amplitude could distinguish 7 pairs successfully.

Healey and Picard conducted a study [18] in which physiological signals (electrocardiogram, electromyogram, respiration, and skin conductance) of the drivers were measured in order to detect their stress level. They used SFFS (sequential forward floating selection) algorithm to recognize pattern of drivers' stress. They recognized the intensity of driver's stress by 88.6%.

Picard et al. conducted another study [29] by showing emotion specific pictures to elicit happiness, sadness, anger, fear, disgust, surprise, neutrality, platonic love, and romantic love. The signals measured were galvanic skin response, heartbeat, respiration, and electrocardiogram. The algorithms used to analyze the data were SFFS, Fisher Projection, and hybrid of these two. The best classification achievement was gained by the hybrid method, which resulted in 81.25% accuracy overall.

3.2 Emotion Elicitation with Movie Clips

In 1995, Gross and Levenson [17] conducted a study to find an answer to the question of whether movie clips could be used to elicit emotions. They showed 78 movie clips to 494 subjects. They chose 16 of these 78 film clips as being the best films based on discreteness and intensity for eight target emotions (amusement, anger, contentment, disgust, fear, neutrality, sadness, and surprise), 2 best films for each emotion. The study showed that these 16 film clips could successfully elicit the above 8 emotions [17].

Since in Gross and Levenson's study [17] these 16 movie clips were validated to elicit these emotions successfully, we used their results to guide the design of our experiment described next.

4. RECOGNIZING EMOTIONS FROM PHYSIOLOGICAL SIGNALS

We designed experiments where we elicited emotions from participants via multi-modal input and measured their physiological signals. We analyzed these physiological signals by implementing and testing pattern recognition algorithms.

4.1 Designing Our Experiment

In order to map physiological signals to certain emotions we designed an experiment in which we elicited six emotions, and measured three physiological signals.

Emotions Elicited: Sadness, Anger, Surprise, Fear, Frustration, and Amusement.

Movie Clips Used: Before choosing the movie clips we used in this experiment we conducted a panel study to measure how effectively they elicited emotions according to the reports of the participants. The chosen movie clips are from the following movies:

- **The Champ** for sadness (same as one of the sadness movie clips from [17])
- **Schindler's List** for anger (different from [17] since the movie clips for anger in [17] gained lower agreement level than it)
- **Capricorn One** for surprise (same as one of the surprise movie clips from [17])
- **Shining** for fear (same as one of the fear movie clips from [17])
- **Drop Dead Fred** for amusement (although the movie clip from When Harry Met Sally from [17] gained the same agreement, chose not to use it since most of the participants experienced embarrassment as well as amusement)

Body signals Measured and Equipment Used: While the above emotions were elicited, participants' galvanic skin response (GSR), heart rate, and body temperature were measured with the non-invasive wearable computer BodyMedia

SenseWear Armband shown in Figure 2 (for GSR and temperature) and a chest band (for heart rate) that works in compliance with the armband. Since the armband is wireless and non-invasive, it can easily and efficiently be used in real life scenarios without distracting the user.



Figure 2 BodyMedia SenseWear Armband

Participants: The sample included 31 undergraduate students from different genders, age groups and ethnicities who enrolled in a computer science course.

Procedure: One to three subjects participated in the study during each session. After signing consent forms, the armband was placed on the subjects' right arm. 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 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: (1) to find a comfortable sitting position and try not to move around until answering an item on the questionnaire, (2) to wait for the slide show to instruct them to answer questions on the in-study questionnaire, (3) to please not look ahead at the questions, and (4) that someone would sit behind them at the beginning of the study to activate the armband.

The 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. The following slides were pictures of the ocean, mountains, trees, sunsets, and butterflies. 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. Starting again with the slide asking the subjects to relax while listening to soothing music, this process continued for the anger, surprise, fear, 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.

The in-study questionnaire included 3 questions for each emotion. The first asked, "Did you experience 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 on an open basis.

4.2 Analyzing the Data

We segmented the collected data, according to the emotions elicited at corresponding time slots, and then we stored the data 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).

We used two different algorithms to analyze the data we collected.

One of them is the *k-Nearest Neighbor Algorithm* [26]. It calculates the distance vectors between test data (test data includes one instance of GSR, temperature, and heart rate values measured

when a particular subject was experiencing a particular emotion) and every instance of the training data sets (with known emotions).

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 $d(x_i, x_j)$ where,

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^n (a_r(x_i) - a_r(x_j))^2}$$

Then the algorithm finds the k smallest distance vectors. The corresponding k emotions of these k instances of data sets are used to determine the corresponding emotion of our test data. The emotion we are looking for is the emotion with the highest frequency in these k emotions.

The other algorithm developed is based on *Discriminant Function Analysis* [28], which is a statistical method to classify data signals by using linear discriminant functions.

Let x_{data} be the average value of a specific data signal, the functions that we are going to solve the coefficients of will be,

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

The coefficients of these functions are calculated from the covariance matrices of the data matrix. Data labeled with emotions are entered into the discriminant function and as a result a new data set clustered by emotions is obtained. By using the k-nearest neighbor algorithm with these emotion clusters, the input signals are mapped to the corresponding emotions.

4.3 Results

As shown in Figure 3, with k-Nearest Neighbor algorithm (KNN, henceforth) the recognition accuracy that we obtained was: 78.58% for anger, 75.00% for sadness, 70.00% for fear, 66.67% for

surprise, 58.33% for frustration, and finally 43.75% for amusement.

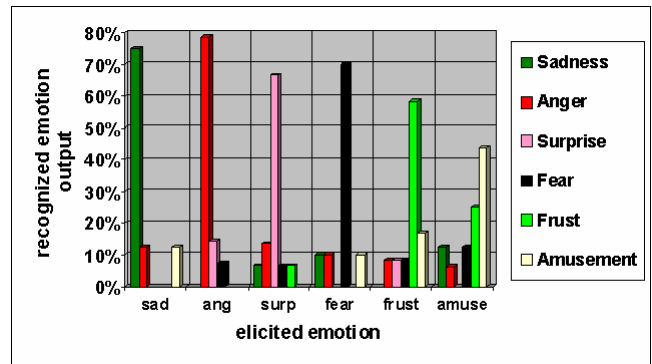


Figure 3 Emotion Recognition Results with KNN Algorithm

Similarly, Figure 4 shows the results of accuracy we gained with Discriminant Function Analysis (DFA) algorithm. The DFA algorithm recognized fear with 90.00%, sadness with 87.50%, anger with 78.58%, amusement with 56.25%, surprise with 53.33%, and frustration with 50.00% success rates.

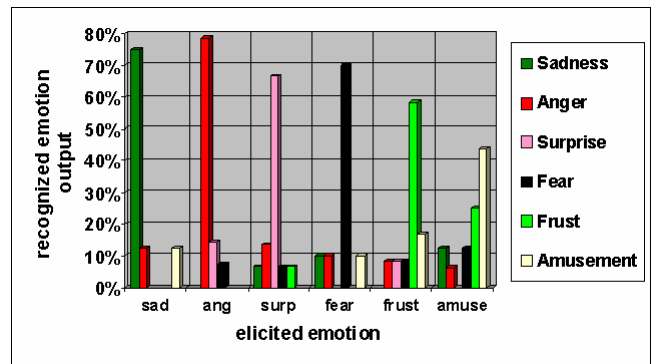


Figure 4 Emotion Recognition Results with DFA Algorithm

The accuracy we gained with our DFA algorithm is better than the KNN Algorithm for fear, sadness, anger, and amusement. KNN performed better for frustration and surprise.

5. VIZUALIZING USER EMOTIONAL STATES

Finally, we designed MAUI [22], a prototype multimedia affective user interface shown in Figure 5, to visualize the output of our various multimodal recognition algorithms.

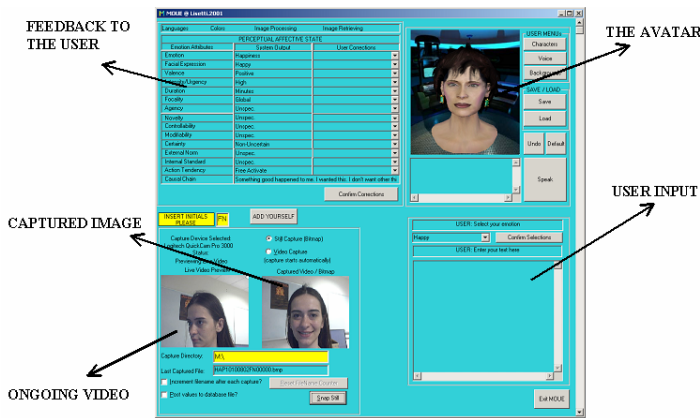


Figure 5: MAUI – Multimedia Affective User Interface

The mode of feedback given to the user is context and application dependent. For example, whereas an interface agent for a car could automatically adjust the radio station or roll down the windows if the driver is falling asleep, an interface agent for a tutoring application can display empathy via an anthropomorphic avatar who adapts its facial expressions and vocal intonation according to the user’s affective state (see Figure 5 upper right).

Our avatar can be used to

- assist the user in understanding his/her emotional state by prompting the user with simple questions, comparing the various components of the states he/she believes to be in with the system’s output (since self-report is often self-misleading)
- mirror the user’s emotions with facial expressions to confirm the user’s emotions (Figure 6)

- animate a previously text-only internet-based chat session showing empathic expressions



Figure 6 Avatar Mirroring User's Affective State

6. CONCLUSION AND FUTURE WORK

We made substantial progress toward recognizing emotions accurately by mapping them to physiological signals with two different algorithms (KNN and DFA). We proposed a way to use this affective knowledge to model the user’s affective state.

Future work will include 1) designing and conducting more experiments and measuring the physiological signals with different equipment; 2) analyzing the collected data with different pattern recognition algorithms; and 3) integrating the different modalities of emotion recognition with recognition from physiological signals including facial expression recognition, vocal intonation recognition, and natural language understanding; 4) adapting the interface to the user appropriately with respect to the current application, e.g. soldier training, driver’s safety [27], and telemedicine [23].

7. REFERENCES

[1] Barker, T., Jones, S., Britton, J., and Messer, D. (2002). 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.

[2] Billsus, D. and Pazzani, M. J. (2000). User Modeling for Adaptive News Access. *User Modeling and User-Adapted Interaction* 10: 147-180.

- [3] Birdwhistle, R. (1970). *Kinesics and Context: Essays on Body Motion and Communication*. University of Pennsylvania Press.
- [4] Bower, G. (1981). Mood and Memory. *American Psychologist*, 36(2).
- [5] Chovil, N. (1991). Discourse-Oriented Facial Displays in Conversation. *Research on Language and Social Interaction* 25: 163-194.
- [6] Collet, C., Vernet-Maury, E., Delhomme, G., and Dittmar, A. (1997). Autonomic Nervous System Response Patterns Specificity to Basic Emotions. *Journal of the Autonomic Nervous System*, 62 (1-2), 45-57.
- [7] Conati, C. (To Appear). Probabilistic Assessment of User's Emotions in Educational Games. *Journal of Applied Artificial Intelligence*.
- [8] Corbett, A., McLaughlin, M., and Scarpinato, S. C. (2000). Modeling Student Knowledge: Cognitive Tutors in High School and College. *User Modeling and User-Adapted Interaction* 10: 81-108.
- [9] Damasio, A. (1994). *Descartes' Error*, New-York: Avon Books.
- [10] Derryberry, D. and Tucker, D. (1992). Neural Mechanisms of Emotion. *Journal of Consulting and Clinical Psychology* 60(3): 329-337.
- [11] Ekman, P., Levenson, R.W., and Friesen, W.V. (1983). Autonomic Nervous System Activity Distinguishes Between Emotions. *Science*, 221 (4616), 1208-1210.
- [12] Ekman, P. and Friesen, W. V. (1975). *Unmasking the Face: A Guide to Recognizing Emotions from Facial Expressions*. Englewood Cliffs, New Jersey: Prentice Hall, Inc.
- [13] Fink, J. and Cobsa, A. (2000). 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.
- [14] Frijda, N. (1986). *The Emotions*. New York: Cambridge University Press. MIT Press book.
- [15] Goleman, D. (1995). *Emotional Intelligence*. New-York: Bantam Books.
- [16] Gross, J. J. Levenson, R. W. (1997). Hiding Feelings: The Acute Effects of Inhibiting Negative and Positive Emotions. *Journal of Abnormal Psychology*, 106 (1), 95-103.
- [17] Gross, J.J., & Levenson, R.W. (1995). Emotion elicitation using films. *Cognition and Emotion*, 9, 87-108.
- [18] Healey, J. and Picard, R. W. (2000). SmartCar: Detecting Driver Stress. In *Proceedings of ICPR'00*, Barcelona, Spain, 2000.
- [19] Hudlicka, E. and McNeese, M. D. (2002). 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.
- [20] Klein, J., Moon, Y., and Picard, R. W. (2002). This computer responds to user frustration: *Theory, design, and results*. *Interacting with Computers*, 14: 119-140.
- [21] Ledoux, J. 1992. Brain Mechanisms of Emotion and Emotional Learning. *Current Opinion in Neurobiology* 2: 191-197.
- [22] Lisetti, C. L. and Nasoz, F. (2002). MAUI: A Multimodal Affective User Interface. Submitted to the *ACM Multimedia International Conference 2002*, (Juan les Pins, France, December 2002).
- [23] Lisetti, C. L., Nasoz, F., Lerouge, C., Ozyer, O., and Alvarez K. (to appear). 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*.
- [24] Maybury, M.T. (2001). Human Computer Interaction: State of the Art and Further Development in the International Context – North America. Invited Talk. *International*

Status Conference on MTI Program.
Saarbruecken, Germany 26-27 October 2001.

- [25] Millan, E. and Perez-de-la-Cruz, J. L. (2002). A Bayesian Diagnostic Algorithm for Student Modeling and its Evaluation. *User Modeling and User-Adapted Interaction* 12: 281-330.
- [26] Mitchell, T. M. (1997). *Machine Learning*, McGraw-Hill Companies Inc.
- [27] Nasoz, F., Ozyer, O., Lisetti, C. L., and Finkelstein N. (2002). Multi-modal Affective Driver Interfaces for Future Cars. In *Proceedings of the ACM Multimedia International Conference 2002*, (Juan les Pins, France, December 2002).
- [28] Nicol, A. A. (1999). *Presenting Your Findings: A Practical Guide for Creating Tables*. Washington, DC: American Physiological Association.
- [29] Picard R. W., Healey, J., and Vyzas, E. (2001). Toward Machine Emotional Intelligence Analysis of Affective Physiological State. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23 (10), 1175-1191.
- [30] Selker, T. (1994). Coach: A Teaching Agent that Learns. *Communications of the ACM*, Vol (37): 7, July.
- [31] Warren, J. R., Frankel, H. K., and Noone, J. T. (2002). Supporting special-purpose health care models via adaptive interfaces to the web. *Interacting with Computers*, 14: 251-267.
- [32] Zajonc, R. (1984). On the Primacy of Affect. *American Psychologist* 39: 117-124.