An Environment to Acknowledge the Interface between Affect and Cognition

Christine L. Lisetti
Department of Computer Science
Stanford University
Stanford, CA 94305 USA
lisetti@psych.stanford.edu

David E. Rumelhart
Department of Psychology
Stanford University
Stanford, CA 94305 USA
der@psych.stanford.edu

Mark Holler
Microcomputer Research Labs
Intel Corporation
Santa Clara, California 95052 USA
mark_holler@ccm.sc.intel.com

Abstract

Human intelligence is being increasingly redefined to include the all-encompassing effect of emotions upon what used to be considered 'pure reason'. With the recent progress of research in computer vision, speech/prosody recognition, and bio-feedback, real-time recognition of affect could very well prove to enhance human-computer interaction considerably, as well as to assist further progress in the development of new emotion theories. We propose an adaptive system architecture designed to integrate the output of various multimodal subsystems. Based upon the perceived user's state, the agent can adapt its interface by responding most appropriately to the current needs of its user, and provide intelligent multi-modal feedback to the user. We concentrate on one aspect of the implementation of such an environment: *facial expression recognition*. We give preliminary results about our approach which uses a neural network.

Introduction

The field of human-computer interaction has recently witnessed the explosion of adaptive and customizable human-computer interfaces which use cognitive user modeling, for example, to extract and represent a student's knowledge, skills, and goals, to help users find information in hypermedia applications, or to tailor information presentation to the user. They can also adapt their interface to a specific user, choose suitable teaching exercises or interventions, give the user feedback about the user's knowledge, and predict the user's future behavior such as answers, goals, preferences, and actions (Jameson, Paris, and Tasso 1997).

New theories of cognition, however, emphasize the tight interface between affect and cognition. Given the increasing use of computers which support the human user in many kinds of task, issues in affective computing (Picard 1997) – " computing that relates to, arises from, or deliberately influences emotions" – necessarily begin to emerge. Until recently, information has been conveyed from the computer to the user mainly via the visual channel, whereas inputs from the user to the computer are made from the keyboard and pointing devices via the user's motor channel.

The recent emergence of bilateral multi-modal interfaces as our everyday tools, might restore a better balance between our physiology and sensory/motor skills, and impact (for the better we hope), the richness of activities we will find ourselves involved in. Given some of the new progress in user-interface primitives composed of gesture, speech, context and affect, it seems possible to design environments which do not impose themselves as *computer environments*, but have a much more natural feeling associated with them.

In particular, as we reintroduce the use of all of our senses within our modern computer tools – or at least the visual, kinesthetic, and auditory (**V**, **K**, **A**) – the possibility to take into account the recent neurological findings about the primordial role of emotions in human cognition opens up. We might also reach an opportunity to better understand ourselves by building instructive multi-modal tools and models focussed around emotional systems (Lisetti 1998), (Rumelhart and Lisetti 1998).

As pointed out in the following section, emotions are strongly related with our physiology. They interface with many important cognitive processes including motivation, attention, memory, perception, as well as rational decision-making. In addition, emotions also play a crucial role in communication patterns, where their expression can convey important information about implicit contextual information in a communicative exchange.

In the remainder of this paper, we give a brief overview of the interface between affect and cognition. We propose a system architecture for a multi-modal intelligent interface which acknowledges the interface between affect and cognition. We describe one subcomponent of the interface which deals with recognizing the facial expressions of the user in the environm-

nent. Lastly we explain future extensions and applications of our system such as instantiation of motivational states for the agent interacting with the user, in order to self-adjust the modalities of its interface. In that manner, the environment can respond most appropriately, as the user's states change over time.

The Affect-Cognition Interface

Affect Representation

Emotions are elicited by a combination of events: sensory, cognitive, and biological. In Zajonc and Markus' theory (Zajonc and Markus 1984b), it is assumed that emotional states can be elicited either by sensory inputs: the sight of flowers might elicit joy, and a strident noise could arouse alertness. Cognitive inputs can also elicit emotions. For example, the memory of a loss might elicit sorrow. Biological events as well can elicit affective reactions: various drugs can affect our emotional states differently.

The actual generation of the basic emotional state – including its autonomic arousal, visceral and muscular activity – depends upon a number of gating processes, such as attention, existing conflicting emotional states which might interfere with the new one, competing muscular engagement, or cognitive conscious or unconscious suppression.

Emotion generation is associated with three phenomena: autonomic nervous system (ANS) arousal, expression, and subjective experience (Zajonc and Markus 1984b). More recent views emphasize the plasticity of the brain and include mechanisms previously considered as results of emotional arousal (e.g. facial actions, ANS activity and breathing patterns) as sources of arousal as well (Zajonc 1989).

It has been traditionally thought that the interface between cognition and affect happened principally at the level of internal mental representation. If affect were to influence information processing, there had to be some affective represention (i.e. the subjective experience of emotion) which intertwined with the cognitive representation being processed. Thus the interaction of affect and cognition was studied by focussing on the associative structures that represent both types of elements.

Alternatively, it has been considered (Zajonc and Markus, 1984b) that affect and cognition can *both* be represented in multiple ways, including in the motor system. This is very observable in the case of affect. While moods and emotions eventually result in cognitive states, they can easily be identified by responses of the motor and visceral system: smiling or frowning faces, embarassed grin, tapping fingers, queasy stomachs, pounding hearts, or tense shoulders.

Cognition, like affect, may also be represented within the organism's activity. While involved in thinking, problem solving, or recalling, people are often found looking for the answers in the ceiling, scratching their heads for inspiration, stroking their chin, or biting their lips. The suggestion here is that both cognition and affect can be represented in the motor system. By building tools to observe, and measure the motor systems, we expect to build a rich database of revealing affective and cognitive phenomena, as well as to enrich the interaction with these tools.

Affect and Cognition

As a result of recent findings, emotions are now considered to be associated with adaptive, organizing, and energizing processes. We mention a few already identified phenomena of interaction between affect and cognition, which we expect will be further studied and manipulated by building intelligent interfaces which acknowledge such an interaction:

- *organization of memory and learning*: we recall an event better when we are in the same mood as when the learning occured (Bower 1981);
- *perception*: when we are happy, our perception is biased at selecting happy events, likewise for negative emotions (Bower 1981);
- *categorization and preference*: familiar objects become preferred objects (Zajonc 1984);
- *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 (Damasio 1994). 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 (Ledoux 1992);
- *focus and attention*: emotions restrict the range of cue utilization such that fewer cues are attended to (Derryberry and Tucker 1992);
- *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 (Frijda 1986);
- *communication*: important information in a conversational exhange comes from body language (Birdwhistle 1970), voice prosody and facial expression revealing emotional content (Ekman 1975), and facial displays connected with various aspects of discourse (Chovil 1991).
- *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 (Goleman 1995).

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.

An Intelligent System for Affect and Human Computer Interaction

Overall System Architecture

As shown in figure 1, we propose an architecture for a system which can take as input both mental and physiological components associated with a particular emotion. Physiological components are to be identified and collected from observing the user using receiving sensors with different modalities: visual, kinesthetic, and auditory (**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.

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 feed-back to the user about his and her current state, including a suggestions as to what next action might be possible to change state. As discussed in the last section, the system is designed to be extended by providing appropriate multimodal feedback to the user depending upon his/her current state. Examples of these adjustments are: changing the agent's voice intonation, changing the pace of a tutoring session, changing the facial expression of an animated agent.

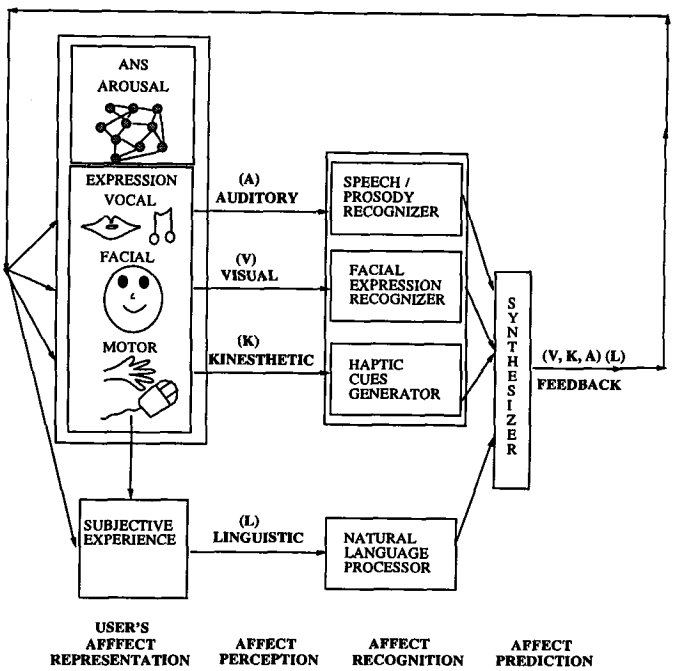

Figure 1: Affect and HCI Interface

Some of the latest progress in AI such as machine vision, speech recognition, and haptic output, make possible the integration of these diverse AI techniques for building intelligent interfaces that reflect the importance of emotion in human intelligence (Picard, 1997). Using the three main sensory systems of machine perception *visual* (**V**), *kinesthetic* (**K**), *auditory* (**A**), and via natural language processing (**L**), it is possible to process computationally:

- **affect perception and recognition** via:
  - (**V**) facial expression (Essa, Darrell, and Pentland 1994), (Black and Yacoob 1995), (Terzopoulos and Waters 1990), and (Kearney and McKenzie 1993);
  - (**A**) vocal emotion (Murray and Arnott 1993);
  - (**K**) advanced bio-feedback via wearable computers (Picard 1997), or via adapted versions of multimodal user interfaces using haptic output (Munch and Dillman 1997);
  - (**L**) spoken or written natural language (O'Rorke and Ortony 1994).

Recognition here needs to be understood in terms of measuring observable behaviors of the motor system which correspond with high *probabilities* to one emotion experienced subjectively. Having perceived some observations of the user's motor system, along

with linguistic patterns, the system recognizes the most likely emotion associated with those patterns, and can make an intelligent description of the user's state.

- **affect prediction/generation** via:
  - low-level "bodily" processes best emulated with neural networks (Rumelhart 1994), (Lisetti 1996) (Lisetti and Rumelhart 1997b);
  - cognitive generation (Elliott 1994) (Frijda and Swagerman 1987), (Dyer, 1987);
  - a combination of both with schemata-like representations (Lisetti 1997).

  Generation is needed in order to enable a computer system to have some level of understanding of what is it like for the user to have emotional states. Generating states for a computer system with similar properties and/or functions to human emotions, might be one of the first steps in building such an understanding.

- **affect expression** via:
  - vocal prosody (Cahn 1990), (Fleming 1997);
  - expressive believable agents (Bates 1994), (Essa, Darrell, and Pentland 1994);
  - semantic descriptions as in computational scripts (Lisetti 1997);
  - speech dialog with facial displays (Nagao and Takeuchi 1994).

  The agent can adapt its interface appropriately by adjusting its output (expressions, gestures, voice inflection) depending upon the recognized user's state, and does so from the correspondingly generated computer state.

It is worth mentioning here, that while different modalities are usually processed separately, sensory integration – for example integrating acoustic and visual information – sometimes offers promising improvements (Stork and Henneke 1996). It is also important to note, that while healthy humans have all these different components, computers can be implemented with one or more of these components (Picard 1997) – for example, with recognition capabilities but without generation capabilities–, and as a result exhibit a different performance.

Because affect recognition is necessary for any further progress to happen in the computational processing of affective states, we present some preliminary results on facial expression recognition.

Facial Expression Recognition

It is expected that three to ten years from now, the price of cameras will have dropped considerably. This will make visual awareness research for artificially intelligent systems a very interesting alley for developing environments.

Quite a lot of research has been done already in the field of face and gesture recognition with some outstanding results. Recently there has been increasing interest in the field of computer vision to recognize facial expressions. A number of systems have dealt with facial expressions using different technical approaches such as the memory-based rule system, JANUS (Kearney and McKenzie 1993), spatio-temporal templates (Essa, Darrell and Pentland 1994), image motion (Black and Yacoob 1995), among others.

One of our motivations is to explore the potential that neural networks offer for this type of pattern recognition problem. An early neural networks which dealt with facial expressions was the single perceptron which could classify smiles from frowns, and which was tested on one person only (Stonham 1986). Since then, other connectionist systems have been implemented to classify and recognize facial expressions with some success (Cottrell and Metcalfe 1991), (Rosenblum, Yacoob, and Davis 1994). Our research project continues to explore the potential of neural networks to recognize facial expressions.

Exactly What Expressions are Relevant?

Facial expressions can be viewed as being communicative signals (Chovil 1991), or they can be considered as being expressions of emotion (Ekman and Friesen 1975). When related with emotions, they raise ongoing debates about their universality. Ekman and Friesen (1975) identified six basic universal emotions: *fear, anger, surprise, disgust, happiness, and sadness.* Another approach emphasizes that what we *refer to* as basic emotions with *labels* such as "fear", have concepts which may very well be culturally determined (Wierzbicka 1992). Studying these concerns is beyond the scope of this paper, and need to be addressed in further details when dealing with particular applications.

Relevant expressions and their interpretations may indeed vary depending upon the chosen type of application. From a computer-interaction perspective, for example, it would be interesting to work with expressions corresponding to *surprise, confusion, frustration* and *satisfaction.* It is also very feasible to work with user-specific expressions corresponding to some of the user's most frequent affective-cognitive states experienced while interacting with the environment. In the

present paper, we work with two expressions, namely *neutral* and *smiling*.

The Data Base of Images

We used the FERET data base of faces which included smiling faces and neutral faces. [1] The FERET database includes pictures of faces with various poses (such as full face, profile, and half profiles) for each person, pictures which are useful to build reliable face recognition algorithms in terms of person identification.

Since we are presently principally interested in facial expressions, however, we built a sub-set of the FERET data base to include only two different poses per person: namely one full face with a neutral expression, and the other full face with a smile. Not every one of the pictures had the same degree of neutrality, and not the same degree of "smilingness". We have designed various different approaches to test this scalability among images.

By contrast with non-connectionist approaches which usually use geometrical face codings, connectionist approaches have typically used image-based representation of faces in the form of 2D pixel intensity array. While this model has the advantage of preserving the relationship between features and texture, it has a very high sensitivity to variations in lighting conditions, head orientation, and size of the picture (Valentin, Abdi, O'Toole, and Cottrell 1994). These problems justify a large amount of work in preprocessing the images. Normalization for size and position is necessary and can be performed automatically with algorithms for locating the face in the image and rescaling it (Turk and Pentland 1991).

In order to isolate relevant portions of the face in our initial approach, we manually reduced the amount of information within a picture to different areas of the face as described later. We also used various techniques to crop, reduce and histogram the images.

Interpreting various facial expressions of an individual in terms of signaled emotions requires us to work with minute changes of some features with highly expressional value. Some examples of those are found in the mouth such as the orientation of the lips (up or down), in the eyes such as the openess of the eyes, etc.

We have been testing various approaches to isolate relevant information in such a way as to work with the smallest amount of data possible (to keep the size of the images workable). There are three areas of the face capable of independent muscular movement: the brow/forehead; the eyes/lids and root of the nose; and the lower face including the cheeks, mouth, most of the nose and the chin (Ekman 1975). Not every expression is shown in the same area of the face. For example *surprise* is mostly shown in the upper part of the face with a lot of wrinkles in the forehead, while *smile* is mostly shown in the lower face. We designed different strategies to test which approach would be better fit to recognize facial expressions in terms of signaled emotions.

[1] Portions of this research in this paper use the FERET database of facial images collected under the ARPA/ARL FERET program.

Some Experimental Results

Full Face Processing: In our initial stage, we preprocessed the images manually, cropped the relevant portions of the face, and performed histogram equalization for normalizing intensity across the different images. In our particular data set, there was no need for rescaling, as all the images were consistent with each other along that dimension.

Table 1: Results of full face training

PERSON ID	MEASURE NEUTRAL	NETWORK OUTPUT	MEASURE SMILE	NETWORK OUTPUT
s197	.2	.199	.5	.493
s198	.2	.181	.2	.220
s199	.1	.172	.6	.588
s200	.2	.251	.4	.378
s201	.2	.258	.8	.662
s202	.4	.383	.5	.533
s203	.1	.163	.3	.305
s204	.1	.190	.3	.261
s205	.2	.249	.4	.369
s206	.2	.197	.6	.650
s207	.2	.174	.3	.282
s208	.1	.201	.8	.659
s209	.3	.313	.3	.232
s210	.2	.240	.8	.663
s211	.2	.230	.8	.665

Table 2: Results of full face generalizing

PERSON ID	MEASURE NEUTRAL	NETWORK OUTPUT	MEASURE SMILE	NETWORK OUTPUT
s192	.2	.221	.6	.500
s193	.1	.166	.4	.168*
s194	.1	.259	.6	.514
s195	.2	.195	.3	.210
s196	.3	.180	.3	.493*

In an early experiment, we used input images of the entire face without head hair (as opposed to hair from beard and mustache which remained). We further reduced the image size by reducing the resolution of the image resulting in input images of size (68X68). We used a connectionist network with one hidden layer, and the *backpropagation* algorithm. The network had 40 hidden units, one input unit per pixel, and one single output unit. We trained the network on 40 input images. We had images of 20 different persons, two im-

ages per person. We included 30 images in our training set and 10 images in our testing set for generalization. Target values ranged from .1 to .6 (similarly to psychological tests used in emotion research), to indicate the level of expressiveness of the image.

The results of the training shown in table 1 indicate that the network learned accurately. The first column identifies the person. The second and fourth columns list the target values evaluating the degree of expression of each image. The third and fifth column give the output of the network.

Generalization on new faces of persons that the network had never "seen" before was then tested. The generalization results are shown in table 2, indicating that the network did not generalize completly on two cases (marked by asterisks '*').

Lower Face Processing: Happiness and satisfaction are typically shown in the third area of the face capable of independent movement. It includes the mouth, most of the nose and the chin (Ekman 1975). We therefore isolated this area of the face and generated cropped images including the lower face only.

We designed two procedures to test how well the network could generalize on new images that it had not been trained with: (1) testing whether the network could generalize if it had been trained on the person with one expression but not on the other, (2) testing whether the network could generalize on people to which it had never been exposed to (i.e. with neither expression).

Test 1: In this case, we selected intermittently neutral faces, and smiling faces as illustrated in table 3. That is, instead of training the network on both expressions for each individual, we trained the network on each individual, sometimes including both expressions but sometimes witholding either one of the two expressions so that we could test how well it generalized at a later state, without having been exposed to both expressions of the same person during the training.

Once again the training was very successful as the network approximated its output to be very close to the values we had given it as targets. The network generalized very accurately for each of the test cases as can be observed from table 4. We then wanted to know if the network could generalize on smiles of people that it had never been exposed to.

Test 2 – Without prior exposure to the person: We trained the network on 116 input images of size (74X73). We included images of 58 different persons, two images per person. We selected 94 images for our training set and 22 images for our testing set for generalization. Plots in figures 2 and 3 compare:

Table 3: Results of lower face training

PERSON ID	MEASURE NEUTRAL	NETWORK OUTPUT	MEASURE SMILE	NETWORK OUTPUT
s192			.6	.599
s193	.1	.178	.4	.362
s194	.1	.199		
s195	.2	.185	.3	.280
s196			.3	.198
s197	.2	.184	.5	.524
s198	.2	.191		
s199	.1	.226	.6	.572
s200			.4	.398
s201	.2	.213	.8	.776
s202	.4	.407		
s203	.1	.168	.3	.302
s204			.3	.266
s205	.2	.196	.4	.407
s206	.2	.206		
s207	.2	.170	.3	.279
s208			.8	.775
s210	.2	.192		
s211	.2	.225	.8	.751

Table 4: Results of lower face generalizing

PERSON ID	MEASURE NEUTRAL	NETWORK OUTPUT	MEASURE SMILE	NETWORK OUTPUT
s192	.2	.253		
s194			.4	.561
s196	.3	.339		
s198			.3	.227
s200	.2	.210		
s202			.5	.438
s204	.1	.178		
s206			.6	.672
s208	.1	.255		
s210			.8	.781

(1) the given target expressed in terms of the degree of expressiveness of a particular image (dotted line) with (2) the actual output of the network generalization and training (plain line). In both of these graphs, the training process starts at person ID 11, while the generalization is plotted over person 1-10. As can be seen from the graphs, the network matched the given target values quite acurately.

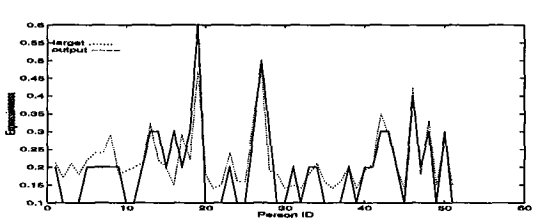

Figure 2: Neutral Expression: Generalization and Training

Discussion

These results indicate that zooming in particular areas of the face for expression detection offers better results

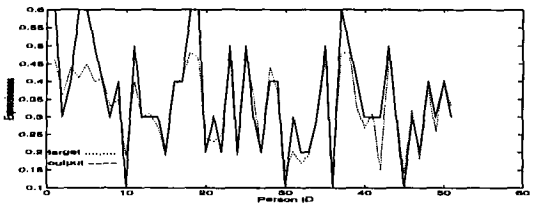

Figure 3: Smile Expression: Generalization and Training

than processing the full face as a whole. Our next approach will be to isolate various areas of the face, and combine their inputs into a single network to increase precision of our recognition algorithm.

One extension of the facial expression system will be the integration of the recognition scheme with a real-time tracker. This coupling is planned to enable the system to perform real-time recognition of facial expressions.

Future Research

An interface to *recognize* and *express* affective states is attractive in that it could improve the communication from the computer to the user by:

- rendering the computer more human-like in its interaction to help the user develop trust/liking for the computer;
- adapting its interface to induce various emotions;
- recording and remembering the user's states during an interaction;
- changing the pace of tutoring session based upon the monitored user's cognitive and emotional states (i.e. bored, overwhelmed, frustrated, etc.);
- guiding the user to avoid cognitive/emotional paths where [s/]he] gets blocked;
- implicitly adapting its interface using multi-modal devices (expression, posture, vocal inflection) to provide adaptive feedback;

From the user to the computer, it might be desirable to:

- give computer agents awareness of what emotional state the user might be feeling so that inferences can be drawn about the motivations implied in them;
- motivate agents to initiate actions (i.e. search and retreive items) using explicitly-set agent's competence level. The level is evaluated in terms of the accuracy of its predictions made from observations of the user (Maes 1994);
- explicitly change some aspects of agent's interface depending on user's state;

From an AI perspective, *affect simulation and generation* might lead to the development of computational models of emotion in order to:

- test emotion theories by providing an artificial environment for exploring the nature and development of emotional intelligence;
- learn from naive psychology: explain, understand, predict behaviors of others, and build user models;
- improve AI process-control: control of cognition, attention, and action;
- choose various planning algorithms under different time pressures signaled by intensity of artificial motivational state;
- develop pro-active intelligent agents: self-motivated software agents or robots with motivational states (Sloman 1990);
- self-adjust the commitment to an ongoing activity based upon valence of current state (negative: slow down waste of energy and reevaluate context, positive: continue in the same direction);

Conclusion

In this paper, we have emphasized the primordial role of emotions on 'high-level' cognitive processes. We discussed a possible architecture for a system able to acknowledge the interface between affect and cognition and to provide multi-modal intelligent feedback accordingly. We have shown results about the implementation of a portion of the system responsible for *facial expression recognition.* Our results indicate that neural networks are very promising for facial expression recognition, a growing area of interest in computer vision, and human-computer interaction. We have also sketched some possibly useful applications of computational processing of emotional phenonema.

Much work is needed to model affect in a computer systems, yet affective computing might very well be a determining factor for the future of the next generation of intelligent computers.

Acknowledgments.

We would like to acknowledge *Intel Corporation* for partial funding for this research.

References

Bates, J. 1994. The Role of Emotions in Believable Agents. *Communications of the ACM* 37(7): 122-125.

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

Black, and Yacoob, Y. 1995. Recognizing Faces Showing Expressions. In Proceedings of the International Workshop on Automatic Face and Gesture Recognition, IEEE Press.

Black, M., and Yacoob, Y. 1995. Tracking and Recognizing Rigid and Non-Rigid Facial Motions using Local Parametric Models of Image Motion. In Proceedings the International Conference on Computer Vision, 374-381.

Bower, G. 1981. Mood and Memory. *American Psychologist*, 36(2).

Cahn, J. 1990. The Generation of Affect in Synthesized Speech. *Journal of the American Voice I/O Society*, 8: 1-19.

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

Cottrell, G. and Metcalfe 1991. EMPATH: Face, Emotion, and Gender Recognition using Holons. In *Advances in Neural Information Processing*, Morgan Kaufmann Publishers.

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

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

Dyer, M. 1987. Emotions and their Computations: Three Computer Models. *Cognition and Emotion*, 1(3): 323-347.

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.

Elliott, C. 1994. Components of Two-way Emotion Communication between Humans and Computers Using Broad, Rudimentary Model of Affect and Personality. *Cognitive Studies: Bulletin of the Japanese Cognitive Society*, 1(2): 16-30.

Essa, I., Darrell, T. and Pentland, A. 1994. Tracking Facial Motion. In Proceedings of the IEEE Workshop on Nonrigid and Articulate Motion 1994.

Fleming, M. A. 1997. Neural Network Model of Micro and Macroprosody. Ph.D. diss., Dept. of Psychology, Stanford University.

Frijda, N. and Swagerman J. 1987. Can Computers Feel? Theory and Design of an Emotional System. *Cognition and Emotion*, 1(3): 235-257.

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

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

Jameson, A., Cecile Paris and Tasso, C. 1997. *User Modeling: Proceedings of UM'97*, New-York: Springer Wien New York.

Kearney, G. and McKenzie, S. 1993. Machine Interpretation of Emotion: Design of a Memory-Based Expert System for Interpreting Facial Expressions in Terms of Signaled Emotions, *Cognitive Science* 17.

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

Lisetti, C., 1998. Acknowledging the Interface between Affect and Cognition: Panel on Affect and Emotion in the User Interface. In Proceedings of the 1998 International Conference on Intelligent User Interfaces, 91-94. New York, NY: ACM.

Lisetti, C. 1997. Motives for Intelligent Agents: Computational Scripts for Emotion Concepts. In Proceedings of Sixth Scandinavian Conference on Artificial Intelligence, 59-70. Amsterdam, Netherlands: IOS Press Frontiers in Artificial Intelligence and Applications.

Lisetti, C. and Rumelhart, D. 1997b. A Connectionist Model of Emotional Arousal and Regulation. Technical Report, PDP-97-01. Dept. of Psychology. Stanford Parallel Distributed Processing Research Group, Stanford University.

Lisetti, C. 1995. *Emotions along the Autonomic Nervous System: A Connectionist Simulation.* Ann Arbor, MI: UMI.

Maes, P. 1994. Agents that Reduce Work and Information Overload. *Communications of the ACM* 37(7): 31-40.

Munch, S. and Dillman, R. 1997. Haptic Output in Multimodal User Interfaces. In Proceedings of the 1997 International Conference on Intelligent User Interfaces, 105-112. New York, NY: ACM Press.

Murray, I. and Arnott, J. 1993. Toward the Simulation of Emotion in Synthetic Speech: A Review of the Litterature on Human Vocal Emotion. *Journal Acoustical Society of America*, 93(2): 1097-1108.

Nagao, K., and Takeuchi, A. 1994. Speech Dialogue with Facial Displays: Multimodal Human-Computer Conversation. In Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics.

O'Rorke, P. and Ortony, A. 1994. Explaining Emotions *Cognitive Science* 18(2): 283-323.

Picard, R. 1997. *Affective Computing.* Cambridge, MA: MIT Press book.

Rosenblum, M., and Yacoob, Y., and Davis, L. 1994. Human Emotion Recognition from Motion Using a Radial Basis Function Network Architecture. In *IEEE Workshop on Motion of Non-Rigid and Articulated Objects.*

Rowley, H. A., and Baluja, S., and Kanade, T. 1996. Neural Network-Based Face Detection. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 285-319.

Rumelhart, D. 1995. Neuro-Modulation: A Connectionist Approach. In H. Morowitz and J. Singer (Eds.), *The Mind, the Brain, and Complex Associative Systems.* 145-153. Readings, MA: Addison-Wesley.

Rumelhart, D. and Lisetti, C. 1998. The Computational Approach to Cognition. In Martin, B. and Rumelhart, D. Ed., *Handbook of Cognition and Perception Volume 2: Cognitive Science.* San Diego, CA: Academic Press. Forthcoming.

Sloman, A. Motives, Mechanisms, and Emotions. 1990. In Boden, M. Ed., *The Philosophy of Artificial Intelligence.* New York, NY: Oxford University Press.

Stonham, T. J. 1986. Practical face recognition and verification with Wisard. In H. Ellis & M.A. Jeeves (Eds.), *Aspects of face processing.* Lancaster, England: Martinus Nijhoff.

Stork, D., and Henneke, M. 1996. Speechreading: An Overview of Image Processing, Feature Extraction, Sensory Integration and Pattern Recognition Techniques. In Proceedings of the Second International Conference on Automatic Face and Gesture Recognition, xvi-xxvi. Los Alamitos, CA: IEEE Computer Society Press.

Terzopoulos, D. and Waters, K. 1990. Analysis and Synthesis of Facial Images Using Physical and Anatomical Models. In Proceedings of the Internationa Conference on Computer Vision, 727-732.

Turk, M., and Pentland, A. 1991. Eigenfaces for Recognition. *Journal of Cognitive Neuroscience* 3(1): 71-86.

Valentin, T., and Abdi, H., and O'Toole, A., and Cottrell, G. 1994. Connectionist Models of Face Processing: A Survey. *Pattern Recognition* 27: 1290-1230.

Wierzbicka, A. 1992. Defining Emotion Concepts. *Cognitive Science* 16: 539-581.

Yacoob, Y., and Lam, H., and Davis, L. 1995. Recognizing Faces Showing Expressions. In Proceedings of the International Workshop on Automatic Face and Gesture Recognition, IEEE Press.

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

Zajonc, R. and Markus, H. 1984b. Affect and Cognition: the Hard Interface. In Izard, C. and Kagan, J. and Zajonc, R. Ed., *Emotion, Cognition and Behavior.* Cambridge, MA: Cambridge University Press.

Zajonc, R. 1989. Feeling and Facial Efference: Implications of the Vascular Theory of Emotion. *Psychological Review* 39: 117-124.