Communicating Emotions in Online Chat Using Physiological Sensors and Animated Text

Hua Wang
The University of Tokyo
7-3-1 Hongo, Bunkyo-ku
Tokyo 133-0033, Japan
wanghua@is.s.u-tokyo.ac.jp

Helmut Prendinger
The University of Tokyo
7-3-1, Hongo, Bunkyo-ku,
Tokyo 113-8656, Japan
helmut@acm.org

Takeo Igarashi
The University of Tokyo/JST PRESTO
7-3-1 Hongo, Bunkyo-ku
Tokyo 133-0033, Japan
takeo@is.s.u-tokyo.ac.jp

ABSTRACT
We present a chat system that uses animated text associated with emotional information to show the affective state of the user. The system obtains the affective state of a chat user from a physiological sensor attached to the user's body. This paper describes preliminary experiments and provides examples of possible applications of our chat system. Observations from informal experiments comparing our animated chat system with a conventional system suggest that an online interface that conveys emotional information helps online users to interact with each other more efficiently.

CATEGORIES AND SUBJECT DESCRIPTORS
H.5.2 [Information Interfaces and Presentation]: User Interfaces — Interaction styles, Screen, design, Theory and methods ; H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems — Animations; H.4.3 [Information Systems Applications]: Communications Applications — Computer conferencing;

GENERAL TERMS
Human Factors

KEYWORDS
Physiological sensor, animated text, kinetic typography, affective computing, online chat

INTRODUCTION
Online communication has become a major tool worldwide, but it still has many shortcomings. The online interfaces in use today, such as chat systems, are far more primitive than face-to-face conversation, making it difficult to convey many basic cues, such as emotions.

The successful deployment of computer-supported communication systems requires the consideration of social factors, such as getting to know other people's feelings, emotions, *etc*. Birdwhistell's [1] linguistic analogy suggests that the information conveyed by words amounts to only 20-30% of the information conveyed in a conversation. These observations demonstrate the importance of non-verbal information, such as emotions, which are essential for human cognition and influence different aspects of peoples' lives. For example, when we are excited, our perception is biased toward the selection of happy events, while the opposite is true for negative emotions. This indicates that online communication would benefit from knowledge of other's affective state.

There is currently considerable interest in how to assess the emotions of a user within a real-time environment [2]. With information on users' affective states, a system can interact more effectively with the users. Conati *et al*. [3] reported that the amount of detail required depends on the task and application, ranging from the overall mood and level of arousal of the user's affective state to one or more specific emotions.

There have been several attempts to use physiological signals to detect specific emotions, such as anxiety, joy, etc. However, none has attempted to use physiological sensors with real-time online chat systems. Therefore, it is not clear how effective the use of physiological signals would be in helping chat users to interact. Here, we describe a new chat system that uses physiological data and animated text to convey feelings and nuanced information during an online chat. We also report the results of a preliminary study that was designed to reveal how the use of affective biometric signals affects online communication in real-time.

RELATED WORK
Animated text
Animated text, also known as kinetic typography, refers to text messages that dynamically change their appearance such as location and shape. The earliest known use of animated text occurred in films and TV advertising. They used animated text because they are capable of conveying emotive content, and thus meets the goals of advertising. Ford *et al*. [4] showed that animated text was effective at conveying a speaker's tone of voice, affect and emotion. Ishizaki [5] demonstrated that, in some cases, animated text could explicitly direct or manipulate the attention of the viewer. Lee *et al*. [6] improved kinetic typography libraries. The Kinedit System (Forlizzi *et al.* [7]) provides a way to create and control kinetic typography and gives users easy methods to control the size, font, and placement of text characters. Bodine *et al.* [8] applied kinetic typography to Instant Messaging.

CHI'04, April 24–29, 2004, Vienna, Austria.
ACM 1-58113-703-6/04/0004.

Affective Computing Interface

Recently, there have been several attempts to build user models that consider emotion and affective state. Conati's probabilistic user model [3], which is based on a dynamic decision network, represents the affective state of the user interacting with an educational game, as well as the user's personality and goals. Klein *et al*. [9] reported an interactive system that responds to the user's self-reported frustration while the user is playing a computer game. Prendinger *et al.* [10] derived the user's emotional state by monitoring physiological signals when interacting with a character-based interface. The study suggests that use of an interface agent with empathic behavior may decrease user stress. Vyzas *et al*. [11] studied the physiological changes that occurred in a human actor when she intentionally induced eight different emotional states. Their classifiers reached an 80% success rate when discriminating among eight emotions.

There have been some studies on the detection of affective state, but none have applied this information to online communication systems, such as online chat. Our approach detects emotion in real-time and applies it to an online chat interface. By showing the animated chat associated with the users' affective information, users can determine their chat partners' affective states. One might argue that one can easily communicate affective information in video chat. We focus on text messages because they still have a number of advantages over videos (simplicity, small data size, less privacy concern, etc.).

DETECTION AND PRESENTATION OF AFFECTIVE STATE

Figure 1 depicts an overview of our online chat system. The system estimates the affective state of the user by using data from physiological sensors and manually specified animation tags. Then, the estimated affective state is presented to other users as animated text[1].

A detailed description of our system is presented below.

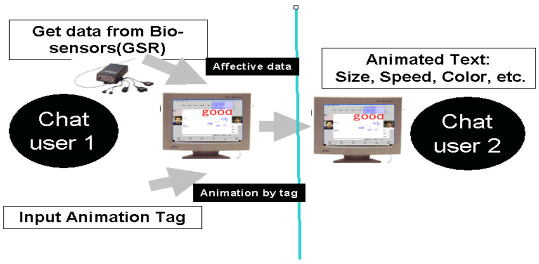

Figure 1 System overview

Lang [12] proposed a two-dimensional emotion map where one axis represents affective arousal and the other represents affective valence (Figure 2). The arousal shows the extent of nervousness or stress. The arousal increases with anxiety and stress. The emotional valence shows whether the emotion is positive or negative. The negatively valenced emotions are fear or sadness, and the positively valenced ones are happiness, excitement, etc. (see the discussion in Affective Computing [2]). Ideally, we would like to determine the affective state of the user from physiological sensors only, without user intervention. However, as we discuss in the next paragraph, it is difficult to obtain valence information from physiological sensors. Therefore, we use a manually specified animation tag to obtain the valence information (emotion type: happy, sad, etc.) and use physiological data to detect the emotional arousal (intensity of emotion: size, speed change, etc.)

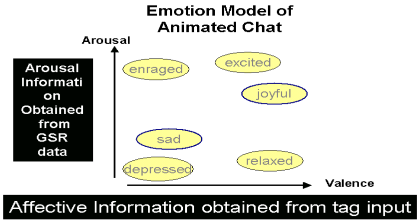

Figure 2 Lang's 2D model in our system

Arousal and valence can be determined from galvanic skin response (GSR) data and valence from blood volume pulse (BVP). BVP increases with negative valence emotions, such as fear and anxiety, and decreases with relaxation (Picard [2]). However, BVP is too sensitive to obtain meaningful data and thus it is difficult to detect affective valence. Since GSR data have a less noisy signal, we decided to use the GSR sensor to detect the arousal and to use manual control to specify the valence. GSR sensors that measure skin conductivity (SC) are attached to the middle and index fingers of the user's non-dominant hand, and the signals are recorded with a ProComp+ unit [13]. SC varies linearly with the overall level of arousal and increases with anxiety and stress. Specifically, we analyze the peaks and troughs of the GSR data to detect arousal.

In addition to GSR data, we use user-specified animation tag input to obtain the affective valence information. The affective valence decides the animation type to use, such as happy or sad and the arousal decides the speed, size, color, and interaction of animated text. When GSR data suddenly increase and show a peak, text animation simultaneously has a higher pace, increases in size and becomes lighter in color. When GSR data decrease, the animation becomes slower, the size is reduced, and the color darkens.

We implemented 20 types of animation (Figure 3), which the user specifies by using a button, a short-cut key, or a tag embedded in a text message. The tags before sentences direct the text animations. For example, "<happy> I am happy!" presents "I am happy!" with happy motion.

[1] The communication of the affective state should be two-way, but our current implementation is one-way because we presently have only one signal capture unit.

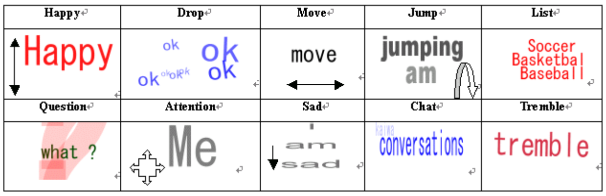

Figure 3 Examples of animations

IMPLEMENTATION

Our system uses a two-dimensional graphical window to display chat animation and information (Figure 4).

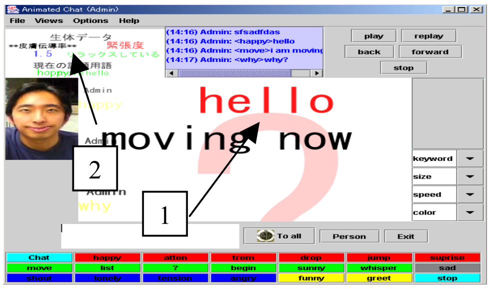

Figure 4: Animated Chat system

Animation Chat Screen

This shows animations of specific words or sentences. The movement of kinetic text changes according to the emotional state detected from the tags and GSR data. (Figure 4)

Affective State Monitor

The screen shows the physiological data obtained from the GSR and provides a means of evaluating whether the chat partner is aroused or calm. The state of arousal is measured from the GSR. When the GSR indicates that the partner is in a state of high mental tension, "Very tense" is displayed on the screen. When the user is in a state of low mental tension, "Relaxed" is shown on the screen.

INFORMAL TESTS AND PRELIMINARY FINDINGS

We performed two preliminary experiments using our system. The first was performed during a short conversation, and the second was conducted in an online educational setting. Detailed descriptions of each test are given below.

Short Conversations

We conducted a preliminary usability study of this system. Six subjects participated in the preliminary experiment. The subjects were college students who used both our system and the conventional Microsoft Messenger. The subjects were assigned to use both the chat systems at a random turn. They were in the same building but were not supposed to see each other. They were separated into three pairs and talked with the partner. Each conversation lasted for more than one hour. They had conversations about school courses, sports, *etc*. After a conversation, the subjects answered questionnaires and commented on the system.

First, we examined the GSR data and the reactions of the subjects. They were asked when they felt most involved in the chat, and we compared their answers with the GSR results. There was a good correlation between the GSR data and the user-reported tension. The subjects reported that they concentrated on the conversations gradually, and the GSR showed similar changes (Figure 5).

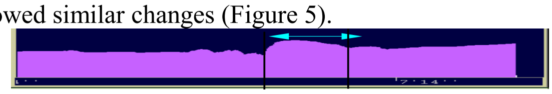

Figure 5 The GSR increased as the subjects concentrated on the conversations more. (The graph records the user's GSR every minute)

This indicated that the GSR can be used to determine changes in mental tension in real time during an online conversation. The results also suggest that emotional information might be able to increase the subject's involvement in the conversation. The emotional information from the chat partner gave users the feeling that they were not only exchanging textual information, but were also communicating their feelings to each other, which allowed them to become more involved in the conversation. One drawback that some subjects mentioned was that they did not always want others to see their GSR information. Table 1 shows some of their answers.

Table 1: Comparison of the systems with and without emotional feedback. Mean scores for questions about the interactive experience in chat systems with and without emotional feedback (EF and Non-EF). The possible responses ranged from 1 (disagree) to 5 (agree) (6 subjects).

Question	EF	non-EF	SD (EF)	SD no-EF
I became very involved in the conversation	4.7	4.0	0.47	0.58
I enjoyed using the system	4.5	4.2	0.50	0.55
I got to know the emotional states of my chat partner	4.2	3.5	0.61	0.50
I was more interested in talking	4.2	4.0	0.49	0.58
It was easy to detect my partner's emotions	3.8	3.5	0.37	0.50

Table 1 shows that the chat system with emotional feedback influenced the subjects' impressions of the other users' emotional states. This may make them feel more eager to talk with their partner and allow them to become more involved in the conversation.

Teacher-Student Interaction in Online Education

We also tested our system in an online education setting. There is already some research on the use of facial expressions in online education. Natatsu *et al*. [14] indicated that by sharing non-verbal information during online education, teachers can interact with students more easily. We tested teacher-student interactions to determine the usefulness of emotional information in an educational setting. We used online education because it is an application in which the more the system knows about the emotions of the user, the better the partner can react, by tailoring the interaction so as to promote learning, while maintaining a high level of engagement.

One teacher and one student took part in the experiment. The physiological sensor was attached to the student, so that the teacher could observe the student's affective state during the session. The teacher presented information about the theme "cartoon history" and asked the student questions about the topic. They were in the same building but were not supposed to see each other.. The online class lasted about 40 minutes. The subjects answered questionnaires and commented on the system. We also analyzed the chat log that accompanied the GSR records. The teacher tried to use the physiological information to gauge the student's reactions. We found that if the teacher suddenly became very strict, the student's GSR data changed rapidly. (Figure 6)

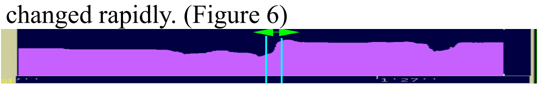

Figure 6 Sudden arousal revealed in a student's GSR (the graph plots the student's GSR every minute)

The subjects also suggested possible uses for this system. Their suggestions can be classified into two categories: 1) Real-time feedback regarding others' tension in a conversation can be useful. From the degree of tension in the conversation, the teacher can identify the student's mood, *e.g*., whether they are concentrating or bored. 2) Feedback can be used later to compare the student's motivation in different classes. By monitoring the changes in their students' concentration, teachers can learn which topics are of interest.

DISCUSSION AND FUTURE WORK

We also explored other approaches for obtaining information on emotion, such as typing speed and auditory feedback. Typing speed can be a supplemental when it is inconvenient to use GSR data. When users are anxious or in a hurry, their typing speed tends to increase. In our system, the text animation can become faster and thereby give other users graphical feedback. In addition, sound is a good way to reflect affective state. In our system, the features of auditory feedback are associated with mental arousal, which is detected by the GSR signal. For instance, when a user is anxious or stressed, an unrestful sound is played to indicate his or her mood. Alternatively, a peaceful sound can be played to regain their calmness.

During the preliminary trials, we found that people tended to read their chat partner's physiological data and tried to influence their partner. They felt that our system provided more feedback on emotion and that this made it more engaging than plain text chat. Most people were interested in knowing whether their partner was nervous, and most said that they enjoyed the chat with accompanying physiological information more than that without it. This system has some limitations. It is not convenient to wear and carry the sensors, and there is still much noise in GSR data. However, we believe that miniaturization of the physiological sensors will make them more convenient to wear and will provide more reliable results. There are also some concerns regarding privacy; some people indicated that sometimes they did not want other people to see their GSR data, and they turned off the signal in order to feel comfortable.

The system could also be used in fields such as long-distance health care. The state of a patient can be determined in real-time from data on their online state together with physiological information. Patients could also become familiar with their own mental and physical states by comparing their own information logs for different sessions or comparing them with those of other people.

Currently, we are investigating methods for recognizing patterns in the physiological signals in online chat. In addition, we are investigating how to remove noise that is generated by external factors. We also plan to analyze video data recorded during the online chat sessions to search for correlations between the video and physiological data. In particular, by integrating animated text, physiological, and video data, we hope to create a new multimodal chat system that can demonstrate users' affective states more efficiently and conveniently.

REFERENCES

1. Birdwhistell, R. Kinetics and Context:Essays on Body Motion Communication, University of Pennsylvania Press, 1970.

2. Picard, R., Affective Computing. Cambridge: MIT Press, 1995.

3. Conati, C., Chabbal, R., Maclaren, H. A Study on using biometric sensors for monitoring user emotions in educational games. Workshop on Assessing and Adapting to User Attitudes and Affect: Why, When and How?, in conjunction with User Modeling (UM-03), 2003.

4. Ford, S., Forlizzi, J., and Ishizaki, S. Kinetic typography: Issues in time-based extended abstracts, In Proceedings of CHI97 Conference Extended Abstracts, 1997, pp 269-270.

5. Ishizaki, S. On Kinetic Typography Statements, American Center for Design Journal Vol. 12 No.1, 1998.

6. Lee, J., Forlizzi, J., and Hudson, S. The kinetic typography engine: an extensible system for animating expressive text, In Proceedings of UIST02, 2002, pp 81-90.

7. Forlizzi, J., Lee, J., and Hudson, S. Kinedit system: Affective messages using dynamic texts, In Proceedings of CHI, ACM Press, 2003, pp 377-384.

8. Bodine, K., and Pignol, M. Kinetic typography-based instant messaging, In Proceedings of CHI03, Demonstration, 2003.

9. Klein, J., Moon, Y., and Picard, R. This computer responds to user frustration: Theory, design, and results. Interacting with Computers, 2002, pp 119-140.

10. Prendinger, H., Mayer, S., Mori, J., and Ishizuka, M. Persona effect revisited. Using bio-signals to measure and reflect the impact of character-based interfaces. 4th Working Conference on Intelligent Virtual Agents,2003, pp 283-291.

11. Vyzas, E., Recognition of Emotional and Cognitive States Using Physiological Data, Mechanical Engineer's Degree Thesis, MIT, June 1999

12. Lang, P. J., The emotion probe: Studies of motivation and attention. American Psychologist, 50(5), 1995, pp 372.

13. ProComp UNIT, Thought Technology Ltd., http://www.thoughttechnology.com/

14 Natatsu, R., Tadenuma. M., Maekawa, T., Computer technologies that support Kansei expression using the body, In Proceedings of 9th ACM Multimedia, 2001, pp 358-364