

Emotion Recognition from Electromyography and Skin Conductance

Arturo Nakasone¹, Helmut Prendinger², Mitsuru Ishizuka¹

¹Dept. of Information and Communication Eng., Graduate School of Information Science and Technology, University of Tokyo, Tokyo, Japan

²National Institute of Informatics, Tokyo, Japan

Abstract— The evocation and detection of the user’s emotional state is becoming a crucial element in the aim for developing more effective interfaces between humans and computers, especially in applications such as games and e-learning tools. In this paper, we describe a model that allows us to determine emotion in real time, based on Electromyography (EMG) and skin conductance. The developed emotion recognition component has been used in a joint research project with the University of Bielefeld [1] that intends to address affective gaming by adding real time emotion detection to a game scenario between a human user and a 3D humanoid agent called Max.

Keywords—Emotion Recognition, Biosignal Interpretation, Affective Computing.

I. INTRODUCTION

For many years, Embodied Conversational Agents (ECA) research has concentrated its efforts into developing computer-generated, humanoid characters able to communicate in a natural way with humans in the context of computer applications [2,3]. In order to increase the believability of ECAs, emotional components based on well known psychological theories [4] have been added to these characters and, due to recent advances in machine emotion recognition, user feedback perception and interpretation [11] is also being incorporated. By analyzing emotional state inputs, agents may be able to adapt their behavior according to the state of the human user and, therefore, be experienced as a more sensible communication partner.

Affective gaming research aims at creating a new type of game experience by adapting the game development to the human player’s affective state [5]. From the perspective of emotion recognition technology, affective gaming is closely related to biofeedback systems, which use physiological signals like blood flow, muscular activity, and brain waves to control several aspects of the gaming experience.

In this paper, two of those physiological signals, namely electromyography (EMG) and skin conductance, were considered to develop a real time emotion recognition component, and integrated to a game where the user plays a simple card game called “Skip-Bo” against the ECA Max. Based on the current status of the game, the component is able to detect what emotion the user might be experiencing at any moment and then provide Max with this input, so that he can adapt his emotion expression and game play based on his own internal emotion model.

The rest of the paper is organized as follows. The next section provides the details of our implementation of a real time emotion recognition component. Section III discusses some methodological issues in real time emotion recognition from bio-signals. Section IV concludes the paper.

II. IMPLEMENTATION

In this section, we describe the module that recognizes emotions inferred from biometric measures in real-time. The module is based on a module that was used in our Empathic Companion system [6]. We start with explaining how a user’s physiological activity can be interpreted as emotional states.

A. Relating Physiological Signals from Emotions

Lang [7] claims that all emotions can be characterized in terms of judged valence (pleasant or unpleasant) and arousal (calm or aroused). Figure 1 shows some named emotions as coordinates in the arousal–valence space. The relation between physiological signals and arousal/valence is established in psychophysiology that argues that the activation of the autonomic nervous system (ANS) changes while emotions are elicited [8]. The following two signals have been chosen for their high reliability

- Galvanic skin response (GSR) is an indicator of skin conductance (SC), and increases linearly with a person’s level of overall arousal;
- Electromyography (EMG) measures muscle activity and has been shown to correlate with negatively valenced emotions.

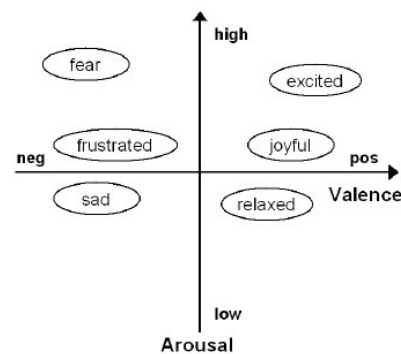


Fig. 1. Some named emotions in the arousal-valence space

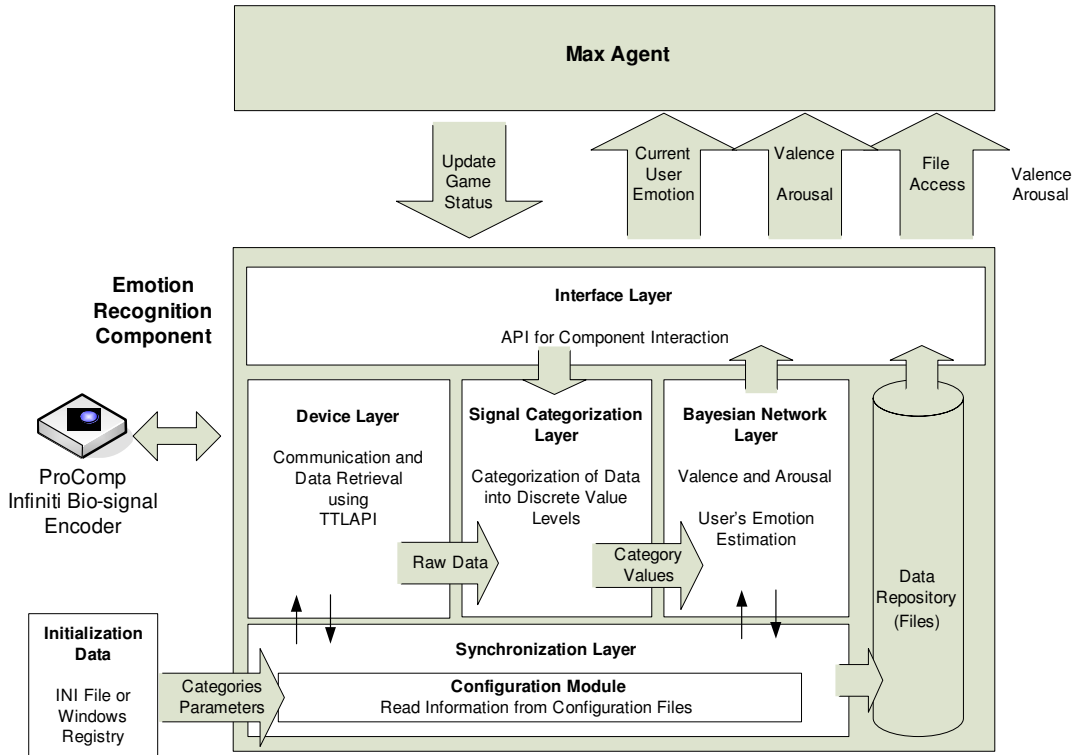


Fig. 2. System Architecture for real time emotion recognition

B. Real-time Emotion Recognition

The module architecture depicted in Fig. 2 shows how the system handles real-time emotion recognition. It has been implemented as an ActiveX component on a Windows XP platform and a proxy program, in order to allow data exchange between the emotion recognition component and the Max agent system. Each of the main components will be explained below.

THE INITIALIZATION DATA FILE

The Initialization Data file contains parameter definitions that are used by the module. This configuration scheme allows the emotion recognition component to have the required flexibility. The parameters defined in this file are: the sampling rate, the size of the value queues, the name of the physical file in which data will be stored and the category definitions.

THE SYNCHRONIZATION LAYER

This layer is used to initialize the emotion recognition component's parameters and to provide a framework for coordination between the other layers on data acquisition, use, and storage.

THE DEVICE LAYER

The user is attached to sensors of the ProComp Infiniti unit from Thought Technologies [9]. The ProComp Infiniti encoder is able to use input from up to eight sensors simultaneously. Currently, we only use galvanic skin response (GSR) and electromyography (EMG) sensors. Input from the sensors is digitally sampled at the rate of 20 samples/sec, but this value can be changed (up to 256 or 2048s/s) by modifying the corresponding parameter in the Initialization file. In order to perform the data acquisition, this layer makes use of the ProComp Infiniti data capture library, known as TTLAPI. The data acquisition process is performed periodically once it has been activated, and it maintains queues for the retrieved data of each sensor, storing new values and discarding old ones. The size of these queues can also be configured using the Initialization Data file.

THE SIGNAL CATEGORIZATION LAYER

When prompted by Max through the emotion recognition component's interface, this layer evaluates the data currently stored in the Device Layer queues. Given the baseline information for skin conductance (GSR signal) and muscle activity (EMG signal), changes in ANS activity are

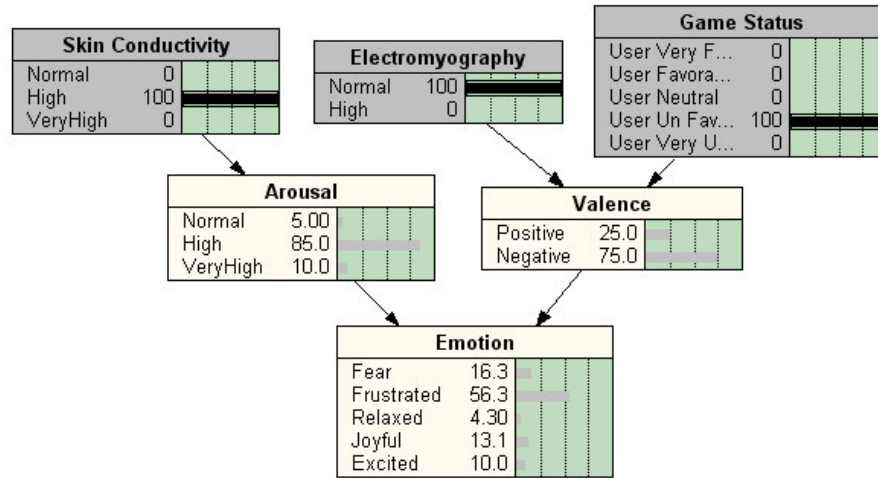


Fig. 3. Bayesian network for user's emotional state detection

computed by comparing the current mean signal values to the baseline value. The baseline is obtained during a relaxation period preceding the interaction (three minutes). The current mean value is derived from a segment of five seconds, the average duration of an emotion [8], though this value can also be configured. If the skin conductance is 15–30% above the baseline, is assumed as “high”, and for more than 30% as “very high”. If muscle activity is more than three times higher than the baseline average, it is assumed as “high”, else “normal”. Here, the current categorization is fixed due to the Bayesian Network model used in the following layer. However, the design of this layer allows a more flexible way to define categories based on the information stored in the Initialization Data File.

THE BAYESIAN NETWORK LAYER

Once the raw data from the sensors has been categorized, a Bayesian network (implemented with Netica [10]) is used to combine the categorized information from the bio-signals and other facts about the interaction and determine the user's emotion based on these values. This network is shown in Fig. 3. Specifically, the Bayesian network is used to derive the user's emotional state by first relating skin conductance to arousal, and EMG and the current state of the game from the user's perspective to valence, and then inferring the user's emotional state by applying the model of [7]. The probabilities have been set in accord with the literature (whereby the concrete numbers are made up). Some examples are: “Relaxed (happiness)” is defined by the absence of autonomic signals, i.e. no arousal (relative to the baseline), and positive valence; “Joyful” is defined by increased arousal and positive valence; “Frustrated” is defined by increased arousal and negative valence. The node “Game Status” represents situations where the game is one of the following states: very favorable for the user, favorable (for the user), neutral, unfavorable, or very

unfavorable. This (‘nonphysiological’) node was included to the network in order to more easily hypothesize the user's positive or negative appraisal of the current situation of the game, as the user's EMG value changes are often too small to evaluate valence. EMG activity is typically seen for strong emotions only.

THE INTERFACE LAYER

This layer provides the interface functions for communication between Max and the emotion recognition component. The functions defined in this layer allow Max to retrieve the user's current state of emotion, as well as the valence and arousal values that were used to determine that emotion. It also allows Max to update the current state of the game in the component, and control the sensor data acquisition process

III. ISSUES IN REAL-TIME ASSESSMENT OF PHYSIOLOGICAL DATA

The psycho-physiological literature discusses a range of problems related to the (real-time) assessment of a person's physiological information [8]. Most important among them are the ‘Baseline Problem’, the ‘Timing of Data Assessment Problem’, and the ‘Intensity of Emotion Problem’.

A. Baseline Problem

The Baseline Problem refers to the problem of finding a condition against which physiological change can be compared – the baseline. An obvious choice is a ‘rest’ period where the subject can be assumed to have no particular emotion. However, as [8] (p. 24) notes, emotion “is rarely superimposed upon a prior state of ‘rest’”. Instead, emotion occurs most typically when the organism is in some prior activation.” Consequently, he suggests adopting a baseline procedure that generates a moderate level of

ANS activity. Hence, in our experiment, we used an initial relaxation period of three minutes with a moderate level of ANS activity, where the subject listens to calm music. This procedure guarantees some independence of subjects' individual ANS activity levels as well as independence of situational factors, such as room temperature.

B. Timing of Data Assessment Problem

The Timing of Data Assessment Problem refers to the temporal dimension of emotion elicitation, including onset (the onset of an emotion indicates how fast an emotion is elicited) and duration of emotions. [8] (p. 30) suggests 0.5-4 seconds as an approximation for the duration of emotions, which locates them duration-wise between (orienting) reflexes (i.e. an organism's response to novelty) and moods. In our study, we take values every 50 milliseconds, for a period of five seconds.

As pointed out in [8], when measuring at the wrong time the emotion might be missed or, different emotions might be covered when too long periods are measured. While the ANS is sometimes considered a slow reacting system, latency of onset for autonomic activity related to emotions can be very short, e.g. with surprise. On the other hand, an emotion like anger may build up over time and blur the actual 'start' of the anger emotion.

C. Intensity of Emotion Problem

The Intensity of Emotion Problem concerns the question how the intensity of an emotion is reflected in the physiological data. While at a low level of emotion intensity no informative ANS activity occurs, a very high intensity level may destroy the pattern of ANS activity associated with an emotion [8]. In our game application, negatively valenced emotions (indicated by EMG) would hardly be indicated, whereas different arousal levels (indicated by GSR) could be easily shown. In practice, emotions with little autonomic activity ('relaxed happiness') or moderate intensity levels seem to occur most frequently. To date, issues in emotion intensity remain largely unsolved [8].

IV. CONCLUSION

In this paper, we have presented a simple model using a Bayesian network to determine user's emotions based on EMG and skin conductance signals, and discussed some methodological problems of real time emotion recognition. We expect that the capability of online emotion recognition will be a key feature of future human-centered interactive applications such as e-learning systems and games. Currently, we are conducting a study using our emotion

recognition component and the Max agent to evaluate users' experience in a game scenario based on the "Skip-Bo" card game. We hope to present the results of this study in the near future.

ACKNOWLEDGMENT

We are indebted to Mr. Christian Becker from the University of Bielefeld, for providing the agent and game scenario for the testing of the emotion recognition component.

REFERENCES

- [1] C. Becker, S. Kopp, and I. Wachsmuth. Simulating the emotion dynamics of a multimodal conversational agent. In Proceedings Tutorial and Research Workshop on Affective Dialogue Systems (ADS-04), pp. 154-165, Springer Verlag, 2004.
- [2] J. Cassell, J. Sullivan, S. Prevost, and E. Churchill. Embodied conversational agents. The MIT Press, Cambridge, MA, 2000.
- [3] H. Prendinger and M. Ishizuka. Life-like characters. Tools, Affective functions, and Applications. Springer Verlag, 2004.
- [4] A. Ortony, G. Clore, and A. Collins. The cognitive structure of emotions. Cambridge University Press, 1988.
- [5] K. M. Gilleade, A. Dix, and J. Allanson. Affective videogames and modes of affective gaming: Assist me, challenge me, emote me ACE. In The 2005 International Conference on Changing Views: Worlds in Play, 2005.
- [6] H. Prendinger, H. Dohi, H. Wang, S. Mayer, and M. Ishizuka. Empathic embodied interfaces: Addressing users' affective state. In Proceedings Tutorial and Research Workshop on Affective Dialogue Systems, pp. 53-64, Springer Verlag, 2004.
- [7] P. J. Lang. The emotion probe: Studies of motivation and attention. *American Psychologist*, 50(5) pp. 372-385, 1995.
- [8] R. W. Levenson. Emotion and the autonomic nervous system: A prospectus for research on autonomic specificity. In H. L. Wagner, editor, *Social Psychophysiology and Emotion: Theory and Clinical Applications*, pp. 17-42. John Wiley & Sons, 1988.
- [9] Thought Technology Ltd., 2002. URL: <http://www.thoughttechnology.com>.
- [10] Netica. Norsys Software Corp., 2003. URL: <http://www.norsys.com>
- [11] R. Picard, E. Vyzas, and J. Healey. Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, number 10, pp. 1175-1191, 2001

Address of the corresponding authors:

Arturo Nakasone, Mitsuru Ishizuka
Department of Information and Communication Engineering
Graduate School of Information Science and Technology
University of Tokyo
7-3-1 Hongo, Bunkyo Ku, Tokyo 113-8656
Tel: +81-(0)3-5841-6347
Email: arturo@miv.t.u-tokyo.ac.jp

Helmut Prendinger,
National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda Ku, Tokyo 101-8430
Tel: +81-(0)3-5841-6347
Email: helmut@nii.ac.jp