Affective Computing: Are we there yet?

Poika Isokoski, Veikko Surakka, Martti Juhola, Jouni Laakso, Timo Partala, Anne Aula, and Marko Illi
Department of Computer and Information Sciences
FIN-33014 University of Tampere, Finland
+358 3 215 8562
poika@cs.uta.fi

ABSTRACT
The ultimate goal of affective computing research certainly seems worthwhile. However, the gap between existing technology and computers capable of appropriate affective response seems wide. We give a brief overview of our work on affective and other physiological computing with emphasis on the real-world applicability of our results.

Keywords
Electromyography, pupil size, emotional feedback, eye tracking, facial muscle activity

INTRODUCTION
For some years now affective computing has been one of the active areas of research at the Tampere Unit for Computer-Human Interaction. Based on our experience it is relatively easy to use conditioned and digitized physiological signals in experimental set-ups. It is also possible for humans to infer the state of the user in off-line analysis based on well-known correlates of certain emotional reactions. However, we have not been able to build systems that could express truly useful affective behavior as mentioned in visionary goals of affective computing.

The piece that seems to be missing is a reliable computable algorithm that could simulate human affective behavior including the differences between individuals. A system running this algorithm would be able to recognize and synthesize believable affective behavior. We are aware of attempts to build systems like this (like those described in [8]), but are not convinced enough on the merits of known systems to commit to building one. Thus, our interest lies in learning about experiences with some of the frameworks and algorithms that have been developed for the affective computing.

Our work has mainly revolved around experiments on human responses to emotional stimuli and feedback. On the technical side we have used eye trackers for pupillometry and facial electromyography (EMG) measurements. Experiments utilizing EEG data interactively are being planned, but so far we have stayed away from electroencephalography (EEG) somewhat because of the constraints that reliable EEG measurements place on the laboratory setting. Immobilizing users in order to get good quality EEG input does not seem like a realistic option in a general-purpose human-computer interface.

In the remainder of this paper we give a short overview of the work (published and unpublished) that has been done in our unit. We focus on the implications of our experience on the robustness of computing systems if built using the technology that we have used.

AFFECTIVE COMPUTING

Pupil Size
When eye trackers are used in user interfaces, the main purpose is usually to follow the direction of the user's gaze. In addition to the gaze direction data video based trackers usually produce information on the relative size of the pupil. To gain experience on the reliability of pupil behavior as an indicator of the user's affective state, we have analyzed the behavior of the pupil in experiments that have involved eye-trackers.

Partala et al. [10] conducted an experiment to study pupillary responses to auditory stimuli. The results showed that an emotionally neutral auditory stimulus evokes an approximately 0.25 mm increase in pupil diameter. Additionally, on average, stimuli with emotional valence (positive or negative) tended to cause larger pupil dilation. However, the absolute magnitude of this additional difference (when present) was in the order of 0.05 mm, which means that it will be difficult to detect reliably in real usage situations with varying lighting conditions and distractions external to human-computer interaction.

In an experiment with spoken emotional feedback in the context of simple mathematical tasks Aula and Surakka [2,3] found essentially the same pupil behavior. The feedbacks caused remarkable pupil dilations, but the differences depending on the content of the feedback were not equally clear. The observed relatively poor discrimination between different feedback conditions can be partly explained by the experimental conditions. The feedback phrases were spoken by a speech synthesizer with a rather monotonous tone that did not change according to the content of the phrase. Also, the content of the feedback did not depend on the user's actual performance but was given randomly.

Electromyography
Facial muscle activity related to user's emotions can be studied non-invasively by using cameras and human viewers or sophisticated image processing algorithms to recognize the features of interest. The other way to gain

roughly the same information is to attach electrodes on the user's face and measure the electrical activity associated with muscle contractions. We have used the latter method mostly because we do not have resources for building the image-analysis software needed for on-line analysis of the video stream, but also because we hope to be able to measure activity that is too small to cause perceptible motion of the muscle and tissues attached to it. Further, because nonverbal visually observable behaviour is in many cases controlled and masked by various factors like social norms, situational and personality factors it may be an unreliable source of information. Thus, the monitoring of physiological signals can be more reliable in respect to the user's spontaneous emotional reactions [11].

A simple approach for studying the EMG data is to make sure that the background noise is low enough and study the amplitude of the signal. This is what we have done in most of our work involving EMG. However, it is unlikely that real-world electrode placement and contact is as good as in our laboratory experiments. Therefore, Laakso et al. [7] have begun work on recognizing bursts of muscle activity based on frequency characteristics rather than overall amplitude.

In the experiment by Aula and Surakka [2,3] the activity over *zygomaticus major* and *corrugator supercilii* was also recorded and analysed for effects of the feedback. Despite significant effects of feedback on the performance of the users, the effects on facial activity were not clear-cut. Previous research has frequently shown that these muscle sites are activated especially to negative and positive emotional stimulation [4,11]. Thus, it is more likely that the results of [2,3] were compromised by the experimental procedure rather than by the failure of these muscle sites to communicate user's emotional state in a reliable way.

NON-AFFECTIVE PHYSIOLOGICAL COMPUTING

In addition to studies exploring the use of eye-trackers and EMG in tracking the user's emotional state, we have worked on user interface prototypes that utilize combined voluntary control of eyes and facial muscles.

This work was started by Partala et al. [9] in an experiment that compared mouse and the combination of gaze direction and facial muscle (*corrugator supercilii*) activity as pointing devices. Gaze direction was used to point and *corrugator supercilii* activity for clicking. The results on pointing speed were rather favorable to the new technique. The pointing times were the same or shorter for the new technique than for the mouse. However, the error rate with the mouse was below 1% while the new technique failed in 34% of the cases.

In an attempt to improve the performance of the new technique we ran another experiment [5,6]. Instead of off-line analysis of the results we implemented the system in a form that worked in real time. This enabled us to give the user feedback on the direction of his or her gaze and success or failure of clicking on the targets. However, we were able to reduce the error rate only to 26% with target size closest to that used by of Partala et al. in the earlier study (20mm). However, the relatively large error rate is mostly due to the inaccuracy of the eye tracker rather than failure of the EMG component or the combination of the input modalities. With 40 mm wide targets the error rate dropped below 10%.

In this and other still ongoing studies, we have found voluntary facial muscle activity to be a robust input method at least when measured with surface electrodes and the relatively high-end amplifier (Grass Model 15) that we have been using. Unfortunately, signs of user's emotions are often much weaker than voluntarily produced facial expressions.

CONCLUSIONS

Both pupil size and muscle activity are relatively easy to measure with proper equipment. In controlled laboratory experiments the data often reflects the user's spontaneous emotional reactions. However, in human-computer interaction situations the affective state of the user may be rather weakly reflected in the measured data. Furthermore, in real-world usage situations the environment is not controlled and thus the measured data is often confounded by many unknown variables. Consequently, automated intelligent on-line interpretation of user's affective state seems, in our experience, to be difficult. A popular approach is to measure more than one physiological reading simultaneously in order to gain a more robust basis for classifying the user's state [1,8]. Even with this approach the best guess on the user's emotions and thus the decision on an appropriate response is sometimes inaccurate or just plain wrong. Thus the most interesting question at the present is whether there are situations where clearly imperfect affective computing can be useful?

REFERENCES

1. Ark, W., Dryer, C. E., and Lu, D. J. The emotion mouse, *Proceedings of the Eighth International Conference on Human-Computer Interaction 1999*, 818-823.
2. Aula, A. *The Effects of Emotional Feedback in Human-Computer Interaction*, Master's thesis, Department of Psychology, Univeristy of Tampere.
3. Aula, A., and Surakka, V. Auditory Emotional Feedback Facilitates Human-Computer Interaction, *Submitted to HCI 2002 Conference*.
4. Hietanen, J. K., Surakka, V., and Linnankoski, I. (1998) Facial electromyographic responses to vocal affect expressions. *Psychophysiology*, 35, 530-536.
5. Illi, M., Isokoski, P., and Surakka V. Combined voluntary gaze direction and voluntary facial muscle activity as a new handsfree technique for human-computer interaction. *Poster at the 11th European Conference on Eye Movements (ECEM11, Turku, Finland)*, 2001.
6. Illi, M., Isokoski, P., and Surakka, V. (2002) Gaze based Human-Computer Interaction. Submitted for

publication in J. Hyönä, R. Radach & H. Deubel (Eds.) *The Mind's Eyes: Cognitive and Applied Aspects of Oculomotor Research*. Elsevier Science: Oxford.

7. Laakso, J., Juhola, M., Surakka, V., Aula, A., and Partala, T., Detection of electromyographic events with neural networks, *Proceendings of Medical Informatics 2001*, London, 489-492.
8. Picard, R. W. *Affective Computing*, MIT Press, Cambridge, Massachusetts, 1997.
9. Partala, T., Aula, A., and Surakka, V. Combined Voluntary Gaze Direction and Facial Muscle Activity as a New Pointing Technique, *In Proceedings of INTERACT'01*, IOS Press 2001, 100-107.
10. Partala, T., Jokiniemi, M. and Surakka, V. Pupillary Responses to Emotionally Provocative Stimuli. *Proceedings of ETRA 2000 Eye Tracking Research and Applications Symposium 2000,* ACM Press, 123-129.
11. Surakka, V., and Hietanen, J. K. Facial and emotional reactions to Duchenne and non-Duchenne smiles. *International Journal of Psychophysiology*, 1998, 29, 23-33.