Usability Indicators - In Your Face

Pedro Branco
IMEDIA
400 Westminster St.
Providence, RI 02903 USA
+1 401 383 1900
pb@imedia-labs.com

ABSTRACT
This paper discusses the method of observing facial expressions as a potential technique to infer usability problems perceived by the users. We present some of the topics of discussion that this approach might raise and the research issues associated.

Keywords
Facial Expressions, Usability, Affective Computing.

INTRODUCTION
Consider the following scenario: "You are running short on time, the software does not seem to share your urgency: you need to insert a few pictures aligned vertically next to the column of text and you still need to review everything before submitting the final version. The text layout keeps shifting all over the page as you insert the pictures and resize them. You grim, you start to get agitated, it is obvious that the stress is taking over you."

If we were observing that moment, we would most likely be able to recognize the tension in our user, maybe through the agitated movements, maybe through changes in our user's facial expressions. Noticeably the computer would neither. The nonverbal behavior, and in particular the grim face, would be important clues to understand the emotional distress and very much indicative of an incident with the interaction. In this paper we present the recognition of facial expressions as a method to identify adverse-event occurrences at the user interface level and discuss the related challenges.

FACE
The face is the most visible and expressive of all the channels for communication of emotions. From the early childhood the face is one of the first links of communication between the mother and the infant. From the thousands of possible different facial expressions, we learn to observe in others surprise, happiness, love, sadness, approval or disapproval, ranging from subtle to more expressive facial expression.

For a long time though the face was considered an unreliable source of emotional information. Facial expressions were thought to depend exclusively on the social environment. Reports of some cultures smiling in grief seemed to support those who argued that facial expressions are socially learned. Eventually those assumptions were challenged when the methodology of observing facial expressions modulated by the cultural settings and by social rules was questioned. Ekman set of studies [3] showed there are set of innate universal expressions that occur spontaneously to emotional stimuli.

The study of spontaneous facial expressions in human computer interaction (HCI) provides a channel to detect users' difficulties. Evidence of that fact, is the video recording from the user some usability labs do during the testing session. That video portraits the user posture and/or face, being useful to convey the user difficulties and providing a stronger argument for the potential recommendations.

We argue that an automated system capable of recognizing facial expressions of distress could in the same manner be helpful to identify problematic features of the interaction. We have two specific applications in mind: as a tool in a usability lab, it could be used to identify the segments of the recording where the users experienced most difficulties; as deployed software it would monitor the daily use of the application and record events where users experienced difficulties. The logs collected over a large pool of users would constitute a remote collection of incidents associated to task in real-world scenarios. However, the envisioned applications raise a set of research questions that need to be addressed to comprehend the feasibility of such approach. We discuss next some of the pertaining issues that need to be addressed or further investigated.

A first and fundamental question is to what extent facial expressions are at all relevant to measure computer related task difficulty? Work with physiological sensors has explored a relation between usability problems and physiological signals using measures such as skin

conductance and heart rate [7]. But the work with facial expressions has not been investigated to a great extent, probably due to the difficulties in measuring it. Two processes are available to record facial expressions: electromyogram (EMG) sensors that measure muscular activiy, and visual observation. The latter can be based on protocols such as the Facial Action Encoding System (FACS) [2], and requires either a manual, time-consuming process or alternatively automatic computer systems, which are yet not fully mature. Electromyogram sensors can be placed over key facial muscles to detect certain expressions. This approach has been taken by a variety of work, studying facial expressions of emotions and constitutes a validated measure. Since it requires to physically place sensors over the face, it is acceptable for a study but not adequate for applications outside the lab. Video-based computer systems are less intrusive in the sense they don't require contact with the user, but remain difficult to implement for all the users and all the different appearances, as well as tend to require high processing time to yield good results. Nevertheless, the increasing presence of web cameras in desktop computing environments and the emergence of commercial software capable of facial expression analysis [5] make this the most promising approach. It remains to be seen what the users' acceptance will be of the presence of video cameras for monitoring emotions.

There are some instances of work linking facial EMG recordings and usability assessment. In [4] the overall usability of a website is inferred from the corrugator facial muscle activity that occurred above a threshold level during the interaction. In our own study [1] with a word processing application, EMG sensors quantified the users' facial expressions. We analyzed task by task the facial expressions of the user and in parallel ranked the degree of difficulty they experienced executing that task. The results established a statistically significant relationship between an increase in the task difficulty experienced by the user and an increase in facial muscle activity.

The EMG sensors applied in those studies record muscular activity that is not necessarily translated in visible facial expressions. A next step requires validation of those same results using video-based facial expression recognition techniques. To my knowledge, Ward [9] provided so far the only study on usability using video-based facial expression monitoring. The reliability of this technique, however, has to improve before it can be fully exploited for evaluation purposes.

Besides the technology constraints and despite the evidence shown above, there remain fundamental questions regarding the viability of the approach. For example, can the face reveal anything about the severity of the problem being perceived by the user? Facial expressions occur with different intensities. Individual differences or mood fluctuations could modulate the intensities of the expression, rather or in parallel with the intensity of the stimuli. The FACS describes the appearance changes that occur with the different intensities of the expression. The ability to recognize those changes in intensity might be just out of reach for now in respect to video-based facial expression recognition. Other metrics such as the duration of the expression or the number of occurrences within a period of time could eventually provided a better measure of the degree of the intensity.

How to interpret the meaning of a facial expressions? The recognition of emotions, in general focus on the set of universal emotions devised from the Ekman studies, eventually with slight variations. Other authors might question the set of discrete emotions and rather prefer to talk about a continuum of emotional states. Regardless of the open discussions in the psychology field, research can proceed focusing on simplified models recognizing just the valence of the emotion: positive and negative. In the particular case of facial expressions there are thousands of different possible facial expressions to be observed. Despite that variability, we are able to recognize in others to some degree facial expression of emotions, and certainly identify positive or negative emotional states.

Computer vision systems prototypes have been shown capable to categorize different categories of emotion though under more restrictive conditions.

So far we have discussed facial expressions solely as a display of an emotional reaction. It is clear that the face is not only a vehicle for the emotional display but can also be attributed to conversational signals or cognitive processes. For example, in the context of a conversation, raising the eyebrows is used to lay emphasis while speaking. If people find it difficult in certain moments to interpret the expression in others, even more so the computer.

On the other hand, interpreting the meaning of facial expressions in a HCI might be less ambiguous in comparison to a social setting, since the context is more restricted. The facial expressions in this specific domain will reflect mostly a cognitive process, (e.g. the effort on planning the next steps to execute a task) or an emotional reaction (e.g. not finding an option in the menu). While the class of the facial expression might be less ambiguous, the same cannot be said about the causes: the behavior might refer to a thought (related or not with the task in hand), to the degree of the user concentration, or to the emotional reaction to an event.

The discrimination between the different categories of expression could possibly be inferred from the task context and the timing, for example if the facial expression and an interface event occurred within few seconds. More complex interactions, e.g., where the user attention is divided between multiple tasks or between the computer and other persons, require other forms of disambiguation, namely head pose, or an indication of whether the user is speaking

or not. To my knowledge the research on these topics has yet to be done.

Another relevant point that might be raised, especially by those who oppose the idea of having a computer monitoring the user, or the computer taking initiatives without user intervention, is: Why not rather ask the user to explicitly communicate when they feel frustrated; would it not solve some of the intrinsic problems being discussed? The disadvantage of such approach is the need for the user to disengage from the task at hand to report the frustration. Depending on the severity of the problem, that extra step might not deserve the consideration from the user when more urgent matters call for attention. Also the extra interface might be a cause of frustration in itself. Notwithstanding there are mechanisms through which users could naturally and intentionally communicate frustration, for example a pressure sensitive mouse. Those modalities are e.g., discussed by Reynolds [6]. Along this line of argumentation we can also think of facial expressions as a channel for intentional communication. Ward [8] suggests this approach could circumvent some of the ambiguity associated with the interpretation of spontaneous facial expressions. Certainly it seems an interesting perspective to be further investigated. Spontaneous expressions might be accompanied by change of posture or other gestures, (e.g. covering the mouth with the hand) which cause difficulties on the recognition; a deliberate expression might simplify the task. On the other hand, given the innate nature of emotional reactions, we might discover users resort to that intentional expression but after a spontaneous reaction.

Finally, what category of usability problems might be detected by monitoring facial expressions? The monitoring of the nonverbal language and in particular facial expressions is an attempt to measure the perception of the user on the task at hand. Identifying the problems the user recognizes is just one step of the usability analysis. Other levels of analysis are necessary to identify the source of the problems and so the approach suggested is just a complement to all the other existing protocol for usability analysis.

CONCLUSION

Certainly recognizing non-verbal behavior is not an easy task for a machine. Spending the effort to provide it with some basic abilities might prove though quite effective in shortening the language gap between users and computers. While the technology to recognize facial expressions starts to emerge, an understanding of how to best utilize it is still in its infancy.

We presented above some of the necessary challenges related to the users facial expressions language.

REFERENCES

1. Branco, P., Firth, P., Encarnação, L. M., Bonato, P. *Faces of Emotion in Human-Computer Interaction*. Proc.of CHI 2005 Extended Abstracts, ACM Press, 1236-1239.
2. Ekman, P., Friesen, W.V., & Hager, J.C. *The Facial Action Coding System*. Second edition (2002). Salt Lake City: Research Nexus eBook.
3. Ekman, P. Universals and cultural differences in facial expressions of emotion. In J. Cole (ed.), *Nebraska symposium on motivation*, 1971. Lincoln, Neb.: University of Nebraska Press, 1972. Pp. 207- 283.
4. Hazlett, R. Measurement of User Frustration: A Biologic Approach. Ext. Abstracts CHI 2003, ACM Press (2003), 734-735.
5. Neven Vision. (n.d.). Retrieved Jan. 16, 2006, from: http://www.nevenvision.com/devtools.html
6. Reynolds, C. *The Sensing and Measurement of Frustration with Computers*. Master's thesis, MIT, 2001.
7. Ward R. D., Marsden P H, Cahill B and Johnson C. *Using skin conductivity to detect emotionally significant events in human-computer interaction*. Proc. IHM-HCI, 2001.
8. Ward, R. D., Marsden, P. H., *Affective computing: problems, reactions and intentions*. Interacting with Computers 16 (2004), 707-713.
9. Ward R. D. *An Analysis of Facial Movement Tracking in Ordinary Human-Computer Interaction*. Interacting with Computers 16 (2005), 879-896.