Affective Computing

Maja Pantic
Delft University of Technology, The Netherlands

INTRODUCTION

We seem to be entering an era of enhanced digital connectivity. Computers and the Internet have become so embedded in the daily fabric of people's lives that they simply cannot live without them (Hoffman et al., 2004). We use this technology to work, to communicate, to shop, to seek out new information, and to entertain ourselves. With this ever-increasing diffusion of computers in society, human-computer interaction (HCI) is becoming increasingly essential to our daily lives.

HCI design was dominated first by direct manipulation and then delegation. The tacit assumption of both styles of interaction has been that the human will be explicit, unambiguous, and fully attentive while controlling the information and command flow. Boredom, preoccupation, and stress are unthinkable, even though they are very human behaviors. This insensitivity of current HCI designs is fine for well-codified tasks. It works for making plane reservations, buying and selling stocks, and, as a matter of fact, almost everything we do with computers today. But this kind of categorical computing is inappropriate for design, debate, and deliberation. In fact, it is the major impediment to having flexible machines capable of adapting to their users and their level of attention, preferences, moods, and intentions.

The ability to detect and understand affective states of a person with whom we are communicating is the core of emotional intelligence. Emotional intelligence (EQ) is a facet of human intelligence that has been argued to be indispensable and even the most important for a successful social life (Goleman, 1995). When it comes to computers, however, not all of them will need emotional intelligence, and none will need all of the related skills that we need. Yet man-machine interactive systems capable of sensing stress, inattention, and heedfulness, and capable of adapting and responding appropriately to these affective states of the user are likely to be perceived as more natural, more efficacious and more trustworthy. The research area of machine analysis and employment of human affective states to build more natural, flexible HCI goes by a general name of affective computing, introduced first by Picard (1997).

BACKGROUND: RESEARCH MOTIVATION

Besides the research on natural, flexible HCI, various research areas and technologies would benefit from efforts to model human perception of affective feedback computationally. For instance, automatic recognition of human affective states is an important research topic for video surveillance as well. Automatic assessment of boredom, inattention, and stress will be highly valuable in situations where firm attention to a crucial but perhaps tedious task is essential, such as aircraft control, air traffic control, nuclear power plant surveillance, or simply driving a ground vehicle like a truck, train, or car. An automated tool could provide prompts for better performance, based on the sensed user's affective states.

Another area that would benefit from efforts toward computer analysis of human affective feedback is the automatic affect-based indexing of digital visual material. A mechanism for detecting scenes or frames that contain expressions of pain, rage, and fear could provide a valuable tool for violent-content-based indexing of movies, video material, and digital libraries.

Other areas where machine tools for analysis of human affective feedback could expand and enhance research and applications include specialized areas in professional and scientific sectors. Monitoring and interpreting affective behavioral cues are important to lawyers, police, and security agents who are often interested in issues concerning deception and attitude. Machine analysis of human affec-

Table 1. The main problem areas in the research on affective computing

- *What is an affective state?* This question is related to psychological issues pertaining to the nature of affective states and the way affective states are to be described by an automatic analyzer of human affective states.
- *What kinds of evidence warrant conclusions about affective states?* In other words, which human communicative signals convey messages about an affective arousal? This issue shapes the choice of different modalities to be integrated into an automatic analyzer of affective feedback.
- *How can various kinds of evidence be combined to generate conclusions about affective states?* This question is related to neurological issues of human sensory-information fusion, which shape the way multi-sensory data is to be combined within an automatic analyzer of affective states.

tive states could be of considerable value in these situations where only informal interpretations are now used. It would also facilitate research in areas such as behavioral science (in studies on emotion and cognition), anthropology (in studies on cross-cultural perception and production of affective states), neurology (in studies on dependence between emotional abilities impairments and brain lesions), and psychiatry (in studies on schizophrenia) in which reliability, sensitivity, and precision are persisting problems.

BACKGROUND: THE PROBLEM DOMAIN

While all agree that machine sensing and interpretation of human affective information would be quite beneficial for manifold research and application areas, addressing these problems is not an easy task. The main problem areas are listed in Table 1.

On one hand, classic psychological research follows from the work of Darwin and claims the existence of six basic expressions of emotions that are universally displayed and recognized: happiness, anger, sadness, surprise, disgust, and fear (Lewis & Haviland-Jones, 2000). In other words, all non-verbal communicative signals (i.e., facial expression, vocal intonations, and physiological reactions) involved in these basic emotions are displayed and recognized cross-culturally. On the other hand, there is now a growing body of psychological research that strongly challenges the classical theory on emotion. Russell (1994) argues that emotion in general can best be characterized in terms of a multi-dimensional affect space, rather than in terms of a small number of emotion categories. Social constructivists argue that emotions are socially constructed ways of interpreting and responding to particular classes of situations and that they do not explain the genuine feeling (affect). Also, there is no consensus on how affective displays should be labeled (Wierzbicka, 1993). The main issue here is that of culture dependency; the comprehension of a given emotion label and the expression of the related emotion seem to be culture dependent (Matsumoto, 1990). In summary, it is not certain that each of us will express a particular affective state by modulating the same communicative signals in the same way, nor is it certain that a particular modulation of interactive cues will be interpreted always in the same way independent of the situation and the observer. The immediate implication is that pragmatic choices (e.g., application- and user-profiled choices) must be made regarding the selection of affective states to be recognized by an automatic analyzer of human affective feedback.

Affective arousal modulates all verbal and non-verbal communicative signals (Ekman & Friesen, 1969). Hence, one could expect that automated human-affect analyzers should include all human interactive modalities (sight, sound, and touch) and should analyze all non-verbal interactive signals (facial expressions, vocal expressions, body gestures, and physiological reactions). Yet the reported research does not confirm this assumption. The visual channel carrying facial expressions and the auditory channel carrying vocal intonations are widely thought of as most important in the human recognition of affective feedback. According to Mehrabian

Related Content