Available online at www.sciencedirect.com
ScienceDirect
Engineering Applications of Artificial Intelligence 20 (2007) 337–345
Engineering Applications of ARTIFICIAL INTELLIGENCE
www.elsevier.com/locate/engappai

A user-independent real-time emotion recognition system for software agents in domestic environments

Enrique Leon*, Graham Clarke, Victor Callaghan, Francisco Sepulveda

Department of Computer Science, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK

Received 1 July 2005; received in revised form 8 October 2005; accepted 2 June 2006
Available online 25 July 2006

Abstract

The mystery surrounding emotions, how they work and how they affect our lives has not yet been unravelled. Scientists still debate the real nature of emotions, whether they are evolutionary, physiological or cognitive are just a few of the different approaches used to explain affective states. Regardless of the various emotional paradigms, neurologists have made progress in demonstrating that emotion is as, or more, important than reason in the process of making decisions and deciding actions. The significance of these findings should not be overlooked in a world that is increasingly reliant on computers to accommodate to user needs. In this paper, a novel approach for recognizing and classifying positive and negative emotional changes in real time using physiological signals is presented. Based on sequential analysis and autoassociative networks, the emotion detection system outlined here is potentially capable of operating on any individual regardless of their physical state and emotional intensity without requiring an arduous adaptation or pre-analysis phase. Results from applying this methodology on real-time data collected from a single subject demonstrated a recognition level of 71.4% which is comparable to the best results achieved by others through off-line analysis. It is suggested that the detection mechanism outlined in this paper has all the characteristics needed to perform emotion recognition in pervasive computing.

Keywords: Emotion detection; Pervasive computing; Software agents

1. Introduction

The aim of this work is to present evidence that user-independent real-time emotion recognition is feasible. From consideration of previous work using off-line analysis, it is suggested that artificial neural networks (ANNs) and sequential analysis could be employed in real-time analysis to guarantee accurate classification. Moreover, ANN generalization capabilities mean that the proposed system could operate on never previously seen data and still provide good performance.

For a long time emotions have been the focus of widespread studies from a variety of scientific areas including medicine and psychology. In 1990, Salovey and Mayer developed a broad framework known as emotional intelligence (EI) to describe how humans perceive and utilize their emotions (Salovey and Mayer, 1990). The initial purpose of EI theorists was to investigate the significance of emotions within the context of intelligence, paying special attention to adaptation and behaviour. However, health, personality, personal ambitions and success have also been analysed from an EI perspective.

The importance of EI in the mechanisms governing human conduct has been highlighted thanks to the work of a group of neurologists who have investigated the idea that, in terms of regulating our decisions and actions, reason is consistently less significant than emotions and, furthermore, that a lack of EI could impair the relationship between humans and their environment (Goleman, 1995). It is evident then that, since emotion recognition is one of the most important components of EI and it has a direct effect on our ability to make optimal decisions (along with the ability to utilize emotions to make decisions), any attempt by computer scientists to model human interaction should, at least in part, be founded on an accurate

*Corresponding author. Tel.: +44 1206 872790; fax: +44 1206 872788.
E-mail addresses: eeleon@essex.ac.uk (E. Leon), graham@essex.ac.uk (G. Clarke), vic@essex.ac.uk (V. Callaghan), fsepulv@essex.ac.uk (F. Sepulveda).

0952-1976/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.engappai.2006.06.001

identification of affective states. It is suggested that by ignoring the emotional component intrinsic to human decision-making, software developers have been missing valuable information that could potentially lead to inadequate interactive models.

One discipline that could particularly benefit from emotion detection is that of pervasive computing. Pervasive or ubiquitous computing involves the integration of computers into the environment allowing the user to interact with them in a more natural way. It is argued that by allowing embedded computers to recognize and use emotional information, software agents of the type used in intelligent inhabited environments (IIE) should be able to use this information to better adapt to what the user wants, increase the accuracy of decisions derived from what the user does, and facilitate mutual interaction. Ultimately actions taken by affective IIE agents could be comparable to intelligent human activity, i.e., would guarantee user's comfort.

1.1. Related research

In the last two decades there has been a considerable amount of research into methods for recognizing emotions using various physical parameters. Facial gestures, vocal utterances, and more recently bodily measures stemming from the autonomic nervous system and the brain have been used to classify affective states. It is suggested however that, because physiological measures are more difficult to conceal or manipulate than facial expressions and vocal utterances, and potentially less intrusive to detect and measure, they are a more reliable representation of inner feelings and remain the most promising way for detecting emotions in computer science.

Hitherto, the identification and classification of emotional changes has achieved mixed results ranging from 60% to 95.5% detection accuracy for facial recognition (Avent et al., 1994; Rosenblum et al., 1994; Sun et al., 2004; Bartlett et al., 2003; Anderson and McOwan, 2003; De Silva and Hui, 2003) to 50–87.5% for speech recognition (Nicholson et al., 2000; Tsuyoshi and Shinji, 1999), and 72% in bimodal recognition (face and speech) (De Silva and Ng, 1999). In physiological emotion detection of some of the best results have been achieved by Kim et al. (2002) with 61.2% correct classification for 4 emotions, Nasoz et al. (2003) 50–90% for 5 emotions and Picard et al. (2001) with 81% for 8 emotions. Some of the recognition techniques employed in the above approaches include neural networks (Avent et al., 1994; Rosenblum et al., 1994; Sun et al., 2004; Bartlett et al., 2003; Anderson and McOwan, 2003; De Silva and Hui, 2003; Nicholson et al., 2000) and advanced statistical mechanisms (Tsuyoshi and Shinji, 1999; De Silva and Ng, 1999; Kim et al., 2002; Nasoz et al., 2003; Picard et al., 2001).

Some applications using emotion recognition for practical purposes include the "Affective DJ"—implemented as an affective wearable (Healey et al., 1998), a driver stress detector (Healey, 2000), a frustration level estimator for poorly interactive computer applications (Fernandez, 1997) and an emotion-based auto-adjustable computer environment (Ark et al., 1999). Emotion detection and estimation has also been used to recognize emotional content in voicemail messages (Inanoglu and Caneel, 2005) and provide immediate medical feedback in tele-home health care systems (Lisetti et al., 2003).

It is important to mention, that one of the main omissions made by researchers investigating emotional states is that they repeatedly forget to verify whether subjects actually experienced the targeted emotional states at the measured level of intensity during experimentation. In many experiments, researchers often assume that participants react to emotional stimuli in a similar manner neglecting personal interpretation that sometimes means intended emotional states never occur. The supposition that individual responses to emotional stimuli are similar for all the participants and that the simple presentation of such stimuli suffices to elicit emotional states could lead to the acquisition of flawed or biased experimental data. A way of correcting this problem is by the use of self-reports collected after a given stimulus has been provided or by selecting experimental subjects whose emotional responsiveness would make it more likely that targeted emotions actually took place.

Another important assumption often found in emotional experimentation and in particular in physiological emotion detection is that situations causing physical arousal that are not linked to emotional episodes, e.g. exercise, can be ignored during data analysis and pattern recognition. The introduction of noise or the potential upheaval of physiological signals due to unforeseen stresses like exercise could make emotion detection more difficult and thus skew the results.

Finally, it should be said that the results achieved in the physiological methods described thus far are based on off-line analysis that requires collection of substantial amounts of data in order to estimate statistical features. As a consequence, on-line dynamic operation is difficult and even in those cases in which real-time implementation is possible (Nasoz et al., 2003; Lisetti et al., 2003), little attention is paid to issues such as robustness and user independence.

2. Background

In Leon et al. (2004a), the combination of autoassociative neural networks (AANNs) and sequential analysis, namely the sequential probability ratio test (SPRT), proved to be effective in detecting changes in 4 physiological signals, more specifically the electromyogram obtained from the masseter, blood volume pressure, skin conductance and respiration rate associated with emotional states from a single individual. The recognition rate on that occasion was 100%. This methodology is based on the idea that the detection of emotional changes using physiological