



Advanced User Interfaces for e-Learning

Luca Bevilacqua¹, Nicola Capuano², Annunziata Cascone², Federico Ceccarini¹, Fabio Corvino¹, Ciro D'Apice², Ivano De Furio¹, Giovanna Scafuro¹, Gianluca Supino¹

¹Engineering.IT, ²Università di Salerno, Dipartimento di Ingegneria dell'Informazione e Matematica Applicata

luca.bevilacqua@eng.it, ncapuano@unisa.it, cascone@diima.unisa.it, federico.ceccarini@eng.it, fabio.corvino@eng.it, dapice@diima.unisa.it, ivano.defurio@eng.it, giovanna.scafuro@eng.it, gianluca.supino@eng.it

Keywords: multimodal, e-learning, attentive use interface, affective user interface, perceptual user interface.

Abstract

This paper outlines an original approach to e-learning systems which integrates the most recent results in the field of human-computer interaction. Notably it will show the applicability of multimodal, attentive, affective and perceptual user interfaces to monitor the students' behavior during their interaction with an e-learning system, in order to measure their attention level and involvement and promptly provide the necessary support to carry out an effective learning process.

1 Introduction

In the last few years all the progress in the field of e-learning has been parallel to technological evolution, with the introduction of new models to take in available innovations. A significant example is the collaborative learning paradigm in which we assist to the affirmation of social networking technologies. Nevertheless, e-learning still suffers from several usage difficulties, both objective and subjective: the long tradition of classroom education, possible negative experiences with first generation products, a background of badly organized self-teaching attempts, the lack of the typical interaction and emotional relationships that can be obtained with a frontal lesson. Especially the last and still unresolved problem represents one of the main reasons of on-line courses drop-out. A significant contribution to its solution and the consequent e-learning development could be the adoption of interfaces that emerges from the studies on human-computer interaction (HCI).

Three paradigms are especially worth noticing:

- the perceptual user interface (PUI) that allows to infer the user's intentions from the observation of his or her explicit and implicit behavior;
- the attentive user interface (AUI), conceived for deducting and managing the user's attention;
- the affective user interface (AFUI), which analyzes the user's emotional state for better adaptation.

The application of PUI, AUI and AFUI in e-learning systems consents not only to reach better results compared to traditional instruments, but also to increase the students' involvement in the whole e-learning process. In fact, as described in (Bangert-Drowns *et al.*, 2001), an involved learner shows a better participation in educational activities from a behavioral, intellectual and emotional point of view.

However, the influence of emotions in learning processes is still underestimated. Only recently we are assisting to a growing number of researches (e.g. Kort *et al.*, 2001 and Currin, 2003) that put emotions at the centre. These studies reveal the importance of the learner's emotional states and, in particular, the relationship between emotions and effective learning (Kort *et al.*, 2001).

Some studies (Isen, 2000) reveal in fact that a positive state of mind stimulates a different way of thinking, characterized with a tendency to greater creativity and flexibility in problem solving and a higher efficiency and precision in decision making.

In conclusion, through the continuous perception of the student's attentive and emotional state it is possible to gather useful information to increase the usability and accessibility of the course. Moreover this makes possible interventions focused both on recovering and maintaining user attention, and on

providing a psychological and didactic support to the student who encounters difficulties during course interaction.

Engineering.IT collaborates with the Dep. Of Information Engineering and Applied Mathematics of the University of Salerno on the national research project WiSe¹, with the goal of investigating new paradigms to improve human-computer interaction. The present contribution describes the activities related to the WiSe project and concerning the development of perceptive-type, attentive-type and affectional-type multimodal interfaces and their applicability in e-learning scenarios.

2 Related Work

Several attempts exist in literature to determine emotional and attentive states through signal processing algorithms that operate on data coming from sensors of different complexity (Picard *et al.*, 2001; Whang *et al.*, 2003). These algorithms provide fairly accurate results, however they require the use of invasive technologies like skin conduction detection, pulse monitoring and cerebral activity measurement.

Even if said technologies can be applied in some restricted domains, they do not meet the requirements of other domains, like e-learning, in which these sensors could distract the users and interfere with their primary activities. For these reasons, researches in the e-learning domain are limited to the use of non intrusive sensors like cameras and microphones.

One of the most used methods to identify the user's emotional state is the analysis of facial expressions through image processing. With the application of these techniques in (Neji and Ammar, 2007), for example, it is possible to detect emotions like joy, sadness, anger, fear, loathing and surprise. Some researches, instead, are based on posture analysis (Mota and Picard, 2003). For example, in (D'Mello *et al.*, 2007) a sensor was used to measure pressure in different points of the chair back and seat to determine the emotional and attentive state of users interacting with an e-learning system. Analyzing the pressure maps in output it has been possible to train the system to recognize five different emotions (boredom, confusion, involvement, joy and frustration) with an accuracy of 70%.

The AutoTutor group of the University of Memphis identified in (Craig *et al.*, 2004) some links between learning objectives and emotions, theorizing that learning is always associated with an emotional episode.

In (Kort *et al.*, 2001) a learning model is proposed in which emotions are

¹ The WiSe "Wireless Services: an User Centric Approach" project was founded by the Italian Ministry of Instruction, University and Research (D.Lgs. N. 297/1999). Started in June 2006 with the purpose of conducting a systematic program of studies and experiments to investigate the industrial potential of new paradigms for the improvement of human-machine interaction (with particular attention for e-learning), it will end in May 2010.

represented on a Cartesian coordinate system and the learner moves between adjacent emotions following a spiral-shaped path. This way the learning level and the emotional level vary according to the quadrant the student is in at a certain moment.

MIT’s Affective Computing Group is investigating, instead, on the correlations between emotion, knowledge and learning (Burlison *et al.*, 2004). Other researches like (Fowler and Mayes, 1999) describe the relationship between learning performance and excitement as an inverted U-shaped curve: people learn better when their emotions are at a moderate level.

Another widespread model to evaluate emotions in didactics is the one known as OCC (Ortony *et al.*, 1990). This model specifies 22 emotion categories based on the emotive reaction of the users to pre-constructed situations and it is extensively used to recognize and map user emotions during interactions with educational games.

3 WiSe and MAAPUI for e-Learning

The study of PUI, AUI and AFUI in the context of the WiSe project has led to the definition of a unified model encompassing these three paradigms, called MAAPUI: Multimodal Attentive Affective Perceptive User Interface (Fig. 1). MAAPUI is a general purpose model that can be applied in different applicative domains. Among these, e-learning (subject of this work) is of particular interest for WiSe project.

The followed approach consists in the analysis of students’ behavioural patterns, captured during an e-learning session by the observation of different interaction means (both implicit and explicit).

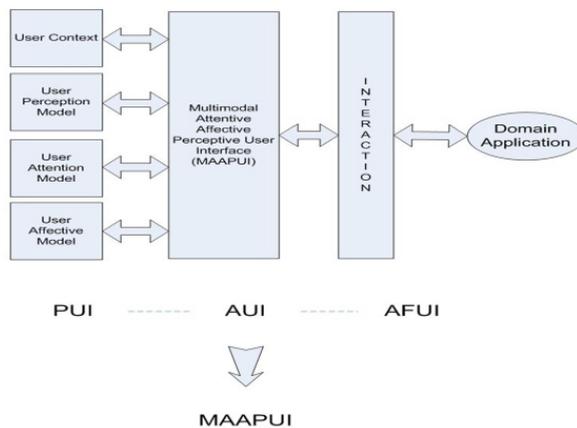


Fig. 1 Multimodal Attentive Affective Perceptual User Interface.

Attentive and affective user's state is inferred from:

- analysis of features extracted from video streams; from them it is possible to deduce information about posture and facial expression;
- analysis of user's interaction with his work station (tracking of the activity on the PC, mouse and keyboard usage).

This information is integrated with the one related to student profile (preferences, competences, personal traits and motivation level) in order to feed an overall user cognitive model (Fig. 2).

The advantages of this approach are numerous and they allow to make an e-learning experience similar to a frontal learning one. For example, the real time perception of a student bored or inattentive posture may allow the system to perform some actions to recover user attention. Furthermore, continuous monitoring of the way a student answers to the questions of an examination test, provides more information than a mere delayed analysis of the answers' correctness. In fact, without caring about the correctness of an answer, the way a student has reached his solution (through a casual selection, a long time meditation, an afterthought), if related to student's skill, can provide information about the need to improve both educational content and test presentation form.

Moreover, detecting that the student is searching the Internet on course related topics may suggest his/her need for more details or additional information. Satisfying this need can improve student satisfaction level or reduce his/her sense of frustration.

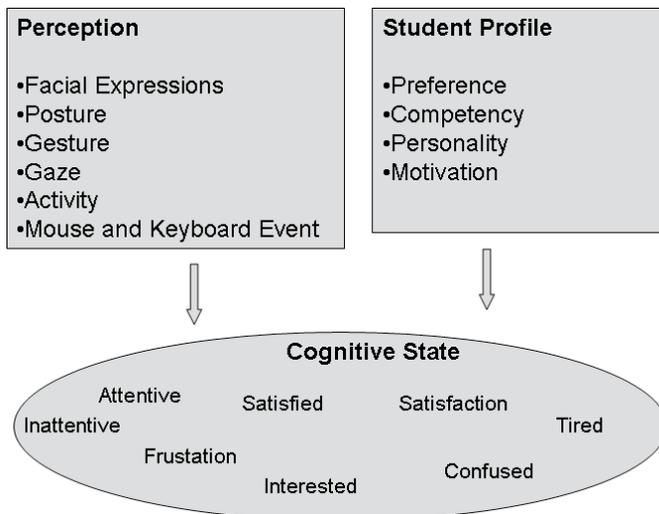


Fig. 2 Student attentive-emotional model.

As explained by the spiral learning model (Kort *et al.*, 2001), during the learning process the student experiences several emotions and mental states. The purpose of our system is to monitor student's behaviour in order to alleviate negative moods (boredom, learning difficulties, etc.) that can hinder the learning process.

The implementation of such a system requires many issues to be solved, related to signal acquisition and to the most significant features extracted, classified and used in a cognitive model. This model has to allow the inference of the student's attentive and affective state with acceptable reliability.

Our approach in signals acquisition and features extraction, according to the considerations contained in section 2, has been to focus only on largely available and unintrusive devices, such as webcam and microphones.

3.1 Posture analysis

Posture, that is the position everyone assumes with his body, is a non verbal involuntary signal, less controllable than face expression and voice tone and it represents a useful clue about emotive and attentive state of a subject.

Body can assume a lot of different postures, but the range is significantly restricted if we consider only sitting postures assumed by a student in front of an e-learning workstation. So it is possible to perform a statistical analysis with the aim of correlating the sitting posture to the attentive and emotive state, as shown in (Mota e Picard, 2003). Starting from the results of these authors, we are designing and prototyping a posture perception module that processes video stream captured by a frontal webcam.

Issues in revealing posture are usually faced by techniques of object shape automatic acquisition. These techniques represent the best solution when it is necessary to distinguish postures having strongly different form factors, such as seated, standing and kneeling postures. On the contrary, in the case of a student involved in an e-learning session, being the student always sit, features like head orientation and its distance from the screen assume more importance than form factors. For this reason, we chose to infer posture through a head tracking module.

There are lots of ways to face the issue of head tracking. The followed approach is tracking-based and provides an effective trade-off between accuracy and computational load. Our algorithm, as tracking-based methods, works following relative head movement across successive frames in a video sequence.

First, during an initialization phase, the face is identified in a specific frame. Then, the research of salient points and main face components (nose, mouth and eyes) is performed. Once these points have been found, a tracking algorithm

(called Optical Flow) is used in order to follow their position across the frames. The knowledge of the points position in every image and the knowledge of their typical position on a face, make it possible to estimate head orientation.

3.2 Gaze tracking

The way a subject visually examines a scene, the time he spends analysing it completely, the duration of fixation on a particular element, some uncontrolled eye movements, represent features related to person cognitive state. User eye movements can be used to study visual attention and exploration.

User interactions with surrounding environment and in particular with graphical user interface on a computer screen represent usual situations in which a visual exploration is performed. According to this assumption, an ocular movement study can give objective and quantitative contributes in human behaviour analysis and it may give interesting clues about the attention level of a person involved in a video terminal activity.

Our gaze tracking module does not require expensive hardware, nor invasive ones and it uses a webcam with a visible light spectrum. It is also based on an Active Appearance Model (AAM) technique to identify face characteristic points in photos and video streams.

The experimentation carried out has been addressed to pupil position identification with respect to eye left and right boundaries, in order to determine gaze direction of a user sitting in front of a computer screen. We have also inspected the video stream looking for typical ocular reading patterns.

3.3 Facial expressions analysis

Without great effort, expert teachers are able to be aware of student needs, to understand their sensations and difficulties simply meeting their gaze or reading their expressions. Consequently, they perform the right actions to recover students attention or reduce their sense of frustration.

For this reason, in the WiSe project, great care has been given to the analysis and interpretation of human facial expressions. The system proposed for this aim combines three approaches:

- holistic analysis of single images that uses Principal Component Analysis techniques to extract features and reduce computational complexity to acceptable levels; a Support Vector Machine (SVM) is used as a classifier;
- a spatial temporal analysis, based on Motion Energy Maps, which also uses a SVM to classify energy maps for different emotions;
- expressions analysis through salient facial points individuation, based

on AAM, and their classification in terms of a Facial Action Coding System.

All these approaches use face detection techniques and a pupil identification algorithm allowing an improvement in image alignment and normalisation. Preliminary results are promising and we believe that combined, simultaneous use of these systems may lead to a good reliability level for facial expression recognition.

3.4 Activity Analysis

A student involved in e-learning session plays out a sequence of implicit interactions which can be gathered through a detailed study of actions performed by the user on the work station. This information is communicated in an unintentionally way, but it may help in inferring the student's attentive and emotional state. Reactivity, smartness, tiredness are all states influencing the way a user deals with the task to carry out.

As a consequence of these considerations, we realised a mouse and keyboard input tracking subsystem that allows the annotation of interesting details about user activity. This module does not alter in any way the user experience and it is software platform independent. We are now performing controlled experiments out to establish which reasonable attentive inferences can be derived from activity analysis.

Anyway the reliability of user activity data is influenced by environmental factors altering user attention in a direct way, as in presence of noise, or indirectly, as in the case of excessive room temperature, that may cause a state of stress and weariness.

Conclusions

In parallel with prototypes of perceptual subsystems, Engineering.IT and the Department of Information Engineering and Applied Mathematics of the University of Salerno are setting up a laboratory to perform an extensive evaluation of prototypes and behavioural assumptions.

Every single e-learning workstation will be able to record streams of the sitting users, and log all systems events (including the ones performed by keyboard and mouse).

Expert teachers will examine and annotate the collected data in order to create a kind of ground truth. Then, analysing the real user behaviour, it will be possible to fine tune and validate the behavioral models proposed.

BIBLIOGRAPHY

- Bangert-Drowns R.L., Pyke C. (2001), *A taxonomy of student engagement with educational software: An exploration of literate thinking with electronic text*. Journal of Educational Computing Research, 24(3), 213-234.
- Bull E. P. (1987), *Posture and Gesture*, Pergamon Press.
- Burleson W., Picard R. W., Perlin K., Lippincott J. (2004), *A platform for affective agent research*, Proceedings of the 7th international Conference on Multimodal Interfaces, New York, ACM Press, 108-115.
- Craig S. D., Graesser A. C., Sullins J., Gholson B. (2004), *Affect and learning: An exploratory look into the role of affect in learning with Auto Tutor*, Journal of Educational Media, 29 (3), 241-250.
- Currin L. (2003), *Feelin 'groovy: Experts now believe that e-learning must elicit positive emotions to succeed*, eLearn Magazine October 2003.
- D'Mello S. K., Picard R., Graesser A. C. (2007), *Toward an affect-sensitive Auto Tutor*, IEEE Intelligent Systems, 22(4), 53-61.
- Fowler C. J. H., Mayes J. T. (1999), *Learning relationships from theory to design*, Association for Learning Technology Journal, 7 (3), 6-16.
- Kort B., Reilly R., Picard R.W. (2001), *An Affective Model of Interplay Between Emotions and Learning: Reengineering Educational Pedagogy-Building a Learning Companion*, Proceedings of International Conference on Advanced Learning Technologies (ICALT 2001), Madison Wisconsin, August 2001.
- Isen, A. M. (2000), *Positive affect and decision making*, In M. Lewis & J. Haviland (Eds.), Handbook of emotions (pp. 720), Guilford, New York: The Guilford Press.
- Mota S., Picard R. W. (2003), *Automated Posture Analysis for Detecting Learner's Interest Level*, Workshop on Computer Vision and Pattern Recognition for Human-Computer Interaction, CVPR HCI.
- Neji M., Ben Ammar M. (2007), *Agent-based Collaborative Affective e-Learning Framework*, The Electronic Journal of e-Learning 2 (5), 123-134.
- Ortony A., Clore G. L., Collins A. (1990), *The cognitive structure of emotions*, Cambridge, Cambridge University Press.
- Picard R.W., Vyzas E., Healey J. (2001), *Toward Machine Emotional Intelligence: Analysis of Affective Physiological State*, IEEE Transactions Pattern Analysis and Machine Intelligence 23 (10), 1175-1191.
- Whang M. C., Lim J. S., Boucsein W. (2003), *Preparing Computers for Affective Communication: A Psychophysiological Concept and Preliminary Results*, Human Factors, 45 (4), 623-634.
- Wilson, G. F., Russell, C. A. (2003b), *Real-time assessment of mental workload using psychophysiological measures and artificial neural networks*. Human Factors, 45, 635-643.