Inf Syst Front (2008) 10:461–472
DOI 10.1007/s10796-008-9104-5

Affective e-Learning in residential and pervasive computing environments

Liping Shen · Victor Callaghan · Ruimin Shen

Published online: 17 June 2008

Abstract This article examines how emerging pervasive computing and affective computing technologies might enhance the adoption of ICT in e-Learning which takes place in the home and wider city environment. In support of this vision we describe two cutting edge ICT environments which combine to form a holistic connected future learning environment. The first is the iSpace, a specialized digital-home test-bed that represents the kind of high-tech, context aware home-based learning environment we envisage future learners using, the second a sophisticated pervasive e-Learning platform that typifies the educational delivery platform our research is targeting. After describing these environments we then present our research that explores how emotion evolves during the learning process and how to leverage emotion feedback to provide adaptive e-Learning system. The motivation driving this work is our desire to improve the performance of the educational experience by developing learning systems that recognize and respond appropriately to emotions exhibited by learners. Finally we report on the results about the emotion recognition from physiological signals which achieved a best-case accuracy rate of 86.5% for four types of learning emotion. To the best of our knowledge, this is the first report on emotion detection by data collected from close-to-real-world learning sessions. We also report some finding about emotion evolution during learning, which are still not enough to validate Kort's learning spiral model.

Keywords e-Learning · Affective computing · Pervasive computing · Residential environments · Adoption · ICT

L. Shen (✉) · R. Shen
Computer Science and Engineering Department,
Shanghai Jiao Tong University,
1954 Huashan Road,
200030 Shanghai, China
e-mail: lpshen@sjtu.edu.cn

R. Shen
e-mail: rmshen@sjtu.edu.cn

V. Callaghan
Department of Computing and Electronics System,
University of Essex,
Essex, UK
e-mail: vic@essex.ac.uk

1 Introduction

Pervasive computing and networks is accelerating the adoption and use of information and communication technology (ICT) into our everyday lives. Many homes feature both internal and external networking, making it possible for people to access a huge variety of services from home automation to new types of media-based services. One such service is e-Learning where learners may learn remote from schools, colleges and universities either in the comfort of their own home or whilst moving around the city. Lessons can be delivered on a variety of platforms ranging from conventional PCs, through IP TVs to mobile phones. Technology and information is a key driver in all areas of life from the home, through business to government. Indeed, many describe the modern world as knowledge based society (Clarke and Callaghan 2007). For such a society learning is an important tool.

In this paper, e-Learning means the delivery of a learning, training or education program assisted by ICT. In the past decade, e-Learning has evolved from Computer-Aided Instruction, through Intelligent Tutoring System, to web-based learning, and to blended learning. Today, e-Learning becomes heavily learner-centered, and therefore emphasizes pervasive learning and personalized learning technologies. Also known as ubiquitous or ambient learning, pervasive learning refers to learning that is available anywhere anytime (Thomas 2008). But in these developments, there has been a bias towards the cognitive and relative neglect of the affective. Moore (2007) defined "transactional distance" as a function of two sets of variables, dialog and structure. The neglect of emotions could increase the transactional distance by decreasing the dialogs between teacher and students in both the classroom lecture and distance learning, which would influence the effect of learning negatively. Surveys showed that the lack of affective awareness is a serious problem in e-Learning (Luo et al. 2006). Of course nobody denies the role of 'affect' or emotion in learning. Certainly teachers know that it plays a crucial role in motivation, interest, and attention. Research (Isen 2000) has demonstrated, for example, that a slight positive mood does not just make you feel a little better but also induces a different kind of thinking, characterized by a tendency towards greater creativity and flexibility in problem solving, as well as more efficiency and thoroughness in decision making. These findings underscore the important effects of emotions on learning. Human brain is not just as a purely cognitive information processing system, but as a system in which both affective functions and cognitive functions are inextricably integrated with one another.

The term affective computing was coined by Picard (1995) in the mid 1990s to describe computer methods that are related to, derived from or deliberately designed to influence, emotions. It involves two areas: emotion synthesis used to artificially imitate some of the physical or behavioral expressions associated with affective states, and emotion analysis which is often employed in decision making for interactive systems. Emotion synthesis is useful to develop ways to communicate with humans at a subjective level involving social participation, for example using robots. Emotion analysis could be used to monitor the emotional state of a subject, taking actions based on the type of individual feeling being experienced. Some computing systems are capable of displaying immediate reactions to people's feelings by incorporating a combination of both emotion detection and emotion synthesis (Morishima 2000). As Picard et al. (2004) stated, most research on emotions has not touched upon learning. Therefore, "existing and future affective and cognitive research needs to be adapted and applied to actual learning situations. Thus far, most research on emotions does not bridge the gap to learning."

In this paper we describe a state-of-the-art e-Learning platform based in Shanghai which, when coupled to a state-of-the-art digital home with context aware sensing at Essex (the iSpace) provides a unique holistic platform to evaluate this work. In particular, we will discuss how context sensing in the form of physiological based emotion detection, used initially for intelligent embedded agents, has been integrated with the Shanghai e-Learning system. The goal of this project is to improve the performance of the educational experience by developing learning systems that recognize and respond appropriately to emotions exhibited by learners. In addition to difficulties associated with the complexity of the content, our remote learner's contextual stresses relating to their environments which were ignored in current e-Learning systems are addressed by our work, which we hypothesize should lead to better learning experiences and wider adoption of this form of learning.

This article targets emotion detection as the main vehicle for capturing context and demonstrates the machine's ability to recognize learner emotions from physiological signals. The remainder of the paper is structured as follows. After introducing related work in Section 2, Section 3 introduces our existing e-Learning platform and residential pervasive environment which collectively typify the longer term learning environment our work is targeting. Section 4 describes our research on affective e-Learning model and its theoretical foundations whilst Section 5 presents the preliminary experiments, emotion classification and data analysis. Finally in Section 6 we summarize our findings and describe our future work plans.

2 Related works

The rapid evolution of ICT has led to new ways of learning and education. They could enable distance learners at home to receive and interact with educational materials and resources and to engage with teachers and peers in ways that previously may have been impossible. The survey information from the National Center for Education Statistics (NCES 2006) reveal that the weighted estimate of the number of students being homeschooled in the United States in the spring of 2003 was 1,096,000, a figure which represents a 29% increase from the estimated 850,000 students who were being homeschooled in the spring of 1999 with the percentage of the student population being homeschooled rising from 1.7% in 1999 to 2.2% in 2003. Within the homeschooled students in 2003, more than 41% of the students had engaged in some sort of e-Learning. The newest survey from China Internet Network Information Center (CNNIC 2007) collected data from 137.0 million Internet users. The results revealed that about 60% of these were connected through broadband; 55% of all Internet users were in the 18~30 age group which are the promising ages for learning; 76% access the Internet from home; 14.3% users access the Internet for online education as their main purpose, rising from 6.3% in 2005, and, finally, 67% of the users who never took online education were reported to be potential candidates for online education. e-Learning has been promoted by many education institutions and numerous corporations to facilitate better learning either in schools, in homes, or even whilst moving around the city. Products such as WebCT (http://www.WebCT.com) and Blackboard (http://www.Blackboard.com) have been in use for the past few years. Many online colleges such as the UK Open University (http://www.open.ac.uk), the Hong Kong Open University (http://www.ouhk.edu.hk) and the Network Education College Shanghai Jiao Tong University (http://www.nec.sjtu.edu.cn), have developed and deployed their own eLearning platforms and infrastructure to provide adaptive and efficient eLearning services.

However, to date, there has been a bias towards the cognitive and relative neglect of the affective in e-Learning systems. The extension of cognitive theory to explain and exploit the role of affect in learning is, at best, in its infancy (Picard et al. 2004). Kort et al. (2001) proposed a four quadrant learning spiral model in which emotions change while the learner moves through quadrants and up the spiral, yet it has not been empirically validated. He also proposed five sets of emotion that may be relevant to learning, but, no empirical evidence exists to confirm the effects these emotions on learning. The Affective Computing Group at MIT's Media Lab is investigating the interplay of emotion, cognition, and learning as part of its "Learning Companion" project. This project is developing an 'affective companion' prototype that will provide emotional support to students in the learning process, assisting them by helping to alleviate frustration and self-doubt (Burleson et al. 2004). Studies carried out by the AutoTutor Group discovered a link between learning and the affective states of confusion, flow and boredom (Craig et al. 2004). According to Fowler's work, the relationship between learning performance and the arousal is a type of inverted-U curve (Fowler 1977). Emotion can also affect learner motivations (Keller and Suzuki 1988). For user emotion modeling, researchers and developers widely refer to two-dimension 'circumplex model of affect' by Russell (1980), where emotions are seen as combinations of arousal and valence. The OCC (Ortony et al. 1990) model has established itself as the standard appraisal model. This model specifies 22 emotion categories based on emotional reactions to situations constructed either as being goals of relevant events, as actions of an accountable agent, or as attitudes of attractive or unattractive objects. Conati and Zhou (2002) are using the OCC theory explicitly for recognizing user emotions in their educational game Prime Climb. Katsionis and Virvou (2005) adapted OCC theory to model students' emotions when they played an educational game. Emotions are also used to design and model learning content. Papert (1996) conducted a project that he described as 'Instead of trying to make children love the math they hate, make a math they'll love' to design things-to-learn so as to elicit affect in ways that will facilitate learning. Sundström (2005) defined an "affective loop" as "an affective interaction process or cycle where emotion plays an important role in interaction involvement and evolution", which is being evaluated using a mobile messaging service, eMoto.

Emotion recognition is one of the key steps towards affective computing. Many efforts have been taken recently to recognize emotions using facial expressions, speech and physiological signals (Cowie et al. 2001; Healey 2000; Picard et al. 2001). The identification and classification of emotional changes has achieved results ranging from 70~98% on six categories of facial expressions exhibited by actors (Bassili 1979) to 50–87.5% for speech recognition (Nicholson et al. 2000). In physiological emotion detection some of the best results have been achieved by Healey (2000) with 80~90% correct classification for eight emotions, Haag et al. (2004) 90% for three valence states and Picard et al. (2001) with 81% for eight emotions. 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 (Picard et al. 2001).

3 Background

The work reported in this article is based on the integration of a context-aware residential environment based at Essex known as the iSpace with a state-of-the-art e-Learning test bed in Shanghai. The Shanghai test bed consists of a large number of distributed smart classrooms, tens of thousands of enrolled students, and thousands of mobile phone users. Both environments are integrated together via networks to create a unique but representative holistic experimental environment to support the vision for this work.

3.1 The pervasive e-learning platform developed in Shanghai

The pervasive e-Learning platform (Fig. 1) was developed at the online college of Shanghai Jiao Tong University and aims to provide "Learning Anytime, Anywhere". It differs from the previous platforms by being heavily learner-centered, and by the extensive use of wireless computing and pervasive computing technologies. The core of the platform includes a number of "smart classrooms" distributed around Shanghai, the Yangtze River delta, and even in remote western regions of China such as Tibet, Yan'an, Xing Jiang and Nin Xia. They are equipped with numerous smart devices/sensors and specially developed software. In this hi-tech environment, the teacher can move freely, demonstrate his body language, and interact with learners as naturally and easily as in a

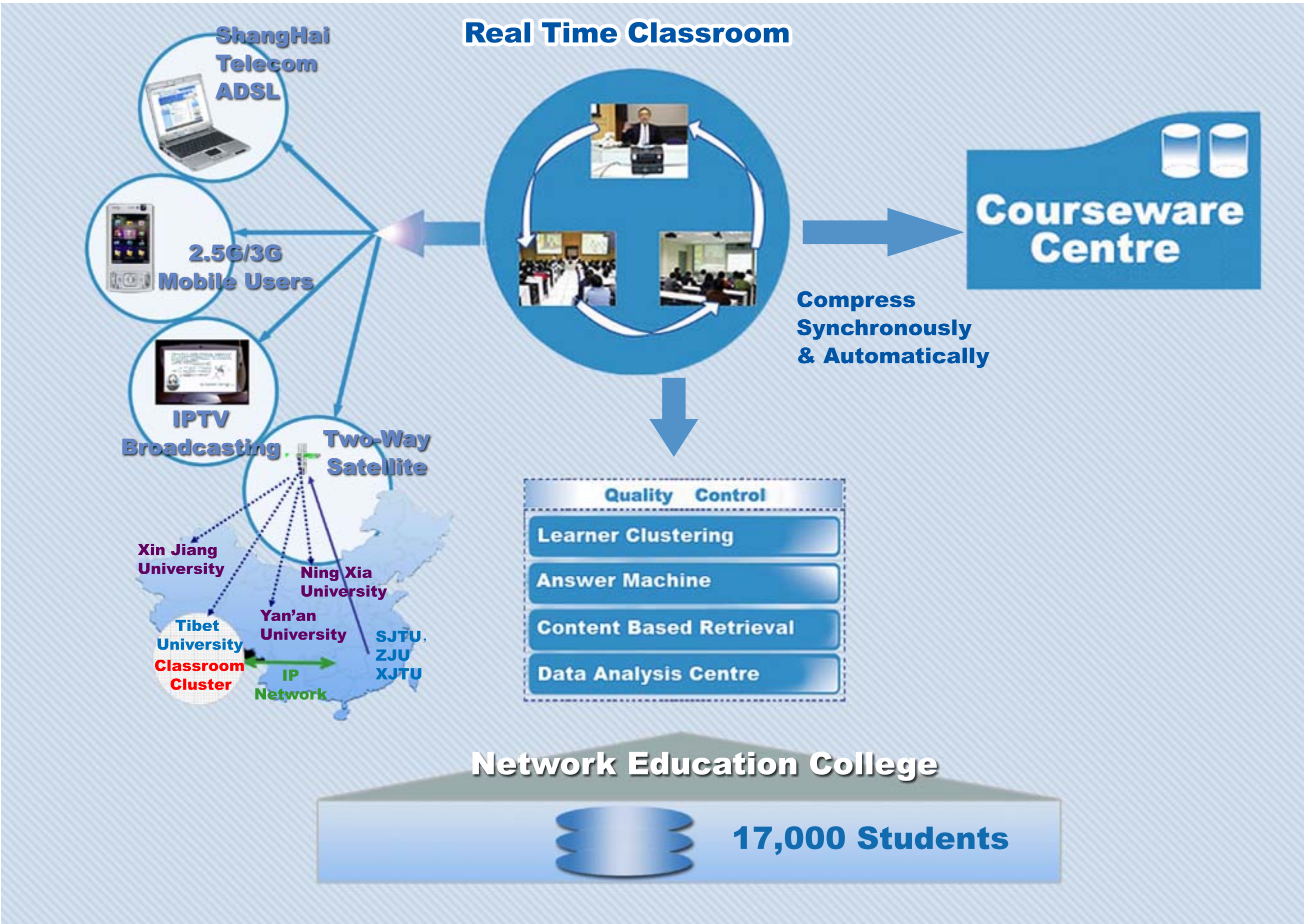

Fig. 1 Pervasive e-Learning platform in Shanghai

traditional face-to-face classroom. The live interactive lectures are digitalized and then delivered to PCs, laptops, PDA, IPTV and mobile phones through various networks such as Shanghai Telecom ADSL, GPRS, IPTV two-way satellite and the Internet. A recording program stores all the media components into courseware including audio, video, handwriting, and any programs or files shown on the computer. Students can access and use these recordings online, or they can download them later for review.

The Shanghai e-Learning platform now serves for about 17,000 enrolled students in the online college. The large number of students and its expansive course delivery systems make it a perfect place to test new and emerging technologies. The hybrid synchronous and asynchronous learning mode, where the learners and teachers are separated in space and time, is the major driving factor of this research for which this eLearning platform offers an ideal test-bed.

3.2 The iSpace pervasive environment test bed in Essex

The intelligent Space (iSpace; Callaghan et al. 2004) is a cutting-edge test bed for domestic context-aware systems based at Essex University. The iSpace is a digital home containing the usual rooms for sleeping, working, eating, washing, and entertaining (Fig. 2). It utilizes many computer networked systems ranging from building utilities

Fig. 2 The iSpace pervasive environment test bed in Essex

to information and media services. To shield users from the technical complexities of programming distributed computer systems, intelligent agents are employed to act on the user's behalf. Occupants of the iSpace utilize a variety of services, including e-Learning (the iSpace is University based, and occupants are frequently learners). Thus iSpace and the Smart Classroom share much in common.

An example of the context aware system developed in the iSpace is a real-time emotion detection system, which achieved an 85.2% correct recognition rate in recognizing three emotional valences (neutral, positive, and negative) from physiological signals (Leon et al. 2007). When this emotional sensing technique was applied to the agents controlling the iSpace systems, the number of times users disagreed with the agents settings dropped by a factor of two. Based on this work we have developed our own emotion sensing technique, described in this paper, specifically targeted at learners. We use the iSpace as a target environment as it represents the type of future home based education environments our work is addressing.

4 Model, rational and R and D strategy

Our affective e-Learning research involves empirically validating theory of emotions that could be used to build an affective e-Learning model. The goal is to understand how learner's emotions evolve during learning process, with the aim of being able to develop learning systems that recognize and respond appropriately to emotions exhibited by learners. The research consists of three main steps:

1. To explore the potential for physiological sensing and emotion evolution for remote learning.
2. To develop an affective learning model to combine emotion with the pervasive e-Learning platform in Shanghai e-Learning test bed.
3. To evaluate and augment the affective learning model that integrates learning and emotion into educational practice.

Step one is the theoretical basis for the second and third steps and forms the primary focus of the work reported in this paper.

Picard et al. (2004) stated, "Theories of affect in learning need to be tested and evolved. However, there is still very little understanding as to which emotions are most important in learning, and how they influence learning. To date there is no comprehensive, empirically validated, theory of emotion that addresses learning". To fill in this tremendous gap between theory and practice, we examined several of the existing emotion theories in learning, so as to help construct our affective e-Learning model. In the experiment reported in this article, we used Russell's 'circumplex model' to describe user's emotion space. We then used the emotion data detected during learning process to explore the affective evolution and empirically validate Kort's 'Learning Spiral Model'. Following is the description of these models and our ongoing research on the affective e-Learning model.

4.1 Russell's circumplex model of affect

In our search for an emotion theory we have focused on dimensional models because they cover the feeling of emotional experience both on a low level and a higher, cognitive level. One well established dimensional model is Psychologist Russell's circumplex model of affect (Russell 1980) where emotions are seen as combinations of arousal and valence (Fig. 3). In Russell's circumplex model of affect, emotions are distributed in a system of coordinates where the *y*-axis is the degree of arousal and the *x*-axis measures the valence, from negative to positive emotions. This model focuses on subjective experiences, which means emotions within these dimensions might not be placed exactly the same for all people. In fact, Fig. 3 shows the author Russell's own dimensional model of emotion.

While Russell provides a comprehensive set of emotions, these are not well matched to our more focused application of learning, and are too many for learning we intend to use as part of our evaluation. Based on Kort et al. (2001) proposed five sets of emotions relevant to learning, we selected four distinctly different emotions: engagement in the quadrant I of Russell two-dimensional affective model (positive valence, high arousal), confusion in quadrant II (negative valence, high arousal), frustration in quadrant III (negative valence, low arousal), and hopefulness in quadrant IV (positive valence, low arousal). These four emotions won't be the optimum set, rather this is a starting point and undoubtedly this emotion set may evolve or take many investigations before it is well established.

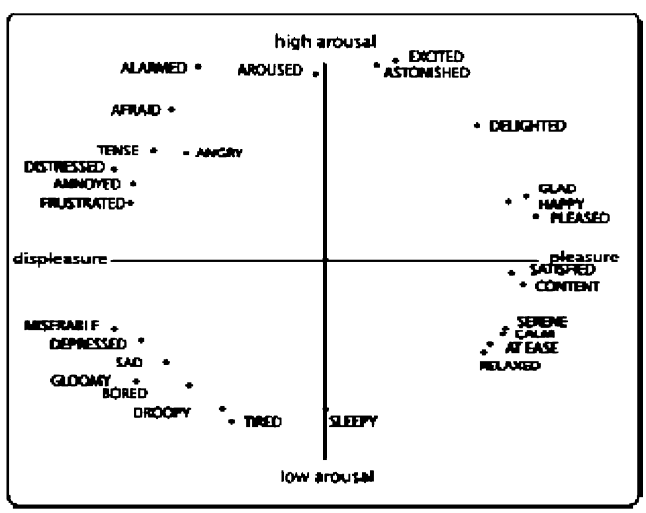

Fig. 3 Russell's circumplex model of affect

4.2 Kort's learning spiral model

Kort et al. (2001) proposed a four quadrant learning spiral model in which emotions change while the learner moves through quadrants and up the spiral (Fig. 4). In quadrant I the learner is experiencing positive affect and constructing knowledge. At this point, the learner is working through the material with ease and has not experienced anything overly puzzling. Once discrepancies start to arise between the information and the learner's knowledge structure, they move to quadrant II, which consists of constructive learning and negative affect. Here they experience affective states such as confusion. As the learner try to sort out the puzzle but fails, he might move into quadrant III. This is the quadrant of unlearning and negative affect, when the learner is experiencing emotions such as frustration. After the misconceptions are discarded, the learner moves into quadrant IV, marked by 'unlearning' and positive affect. While in this quadrant the learner is still not sure exactly how to go forward. However, they do acquire new insights and search for new ideas. Once they develop new ideas, they are propelled back into quadrant I; thus, concluding one cycle around the learning spiral of Kort's model. As learners move up the spiral, cycle after cycle, they become more competent and acquire more domain knowledge.

4.3 Towards the affective e-Learning model

The work described here focuses on how, when we have got the emotion states, we can make sense of and make use of this information to build an affective e-Learning Model, which should support either synchronous classroom education or asynchronous self learning. We have built the following two prototypes to explore affective e-Learning Model.

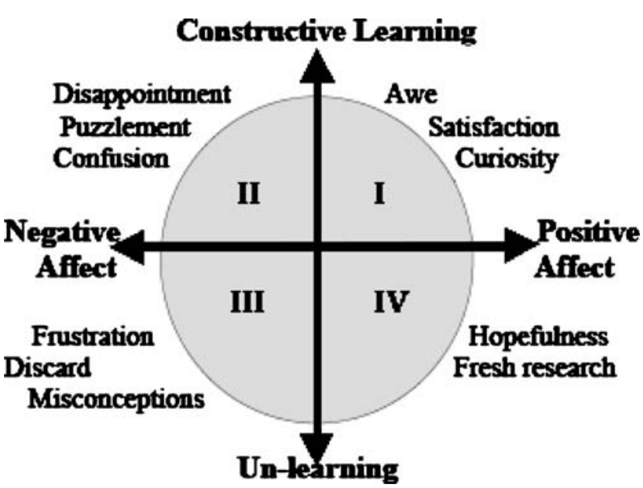

Fig. 4 Kort's learning spiral model

4.3.1 Emotion-aware smart classroom

Smart Classrooms are the core of the pervasive e-Learning platform. Expert teachers are able to recognize the emotional state of their students and respond in ways that positively impact on learning in traditional face-to-face education. But in the e-Learning case, there are large numbers of remote students in distributed classrooms and mobile users; thus the challenge is how a teacher could circumvent this? We provided a solution for such problems via the incorporation of students' emotional information into the pervasive e-Learning platform. Firstly we simply feedback the students emotions back to the lecturer in real-time, that the lecturer would adapt the lecture style, speed and content based on the students' emotional statistics. As emotion plays an important role in interaction involvement and evolution, the teacher should be aware of the students' emotional states and emotional footprint when organizing group discussions so as to enhance the information flow within the group by smoothing the emotion flow. We are collecting data to investigate the computational model of emotion-aware group interaction dynamics.

4.3.2 Emotion-aware adaptive content delivery

Based on our previous work (Shen and Shen 2005), we built a prototype to provide personalized service based on the learner's emotions. The aim of this prototype is to incorporate the learner's emotional states together with the learner's cognitive abilities, and his/her learning goals, to generate appropriate responses to the learner. We are using the prototype to explore the interaction between the user and learning system, to detect user emotional responses to system behavior, and to eventually provide adaptive and personalized service to the learner.

5 Preliminary experiments and results

We used the same method as Picard et al. (2001) that data was gathered from a single subject over many weeks of time, and at two different places (one in UK and another in China), standing in contrast to efforts that examine many subjects over a short recording interval (usually single session on only 1 day). Although this is limited to one subject, the data set is larger than those used in traditional affect recognition studies involving multiple subjects. There are many reasons to focus on one subject in the preliminary experiment. Ekman et al. (1983) acknowledged that even simply labeled emotions like "joy" and "anger" could have different interpretations across individuals within the same culture; so subjects might elicit different physiological patterns for the same emotion. When lots of subjects have been examined over a short amount of time, researchers might have difficulty finding significant physiological patterns in part because physiology can vary subtly with how each individual interprets each emotion. By using one subject, we tried to focus on the same personal interpretation of the emotions, and could learn the affective evolution during the long experiment period. For pervasive/personal computing applications, we desire the machine to learn an individual's patterns, and not just some average response formed across a group, which may not apply to the individual. However, the methodology for gathering and analyzing the data in this paper is not dependent on the subject; the approach described is general.

The subject in our experiments was a healthy female PhD student. The preliminary experiment was firstly carried out in the intelligent inhabited environment, iSpace of Essex University, and later, in the e-Learning lab of Shanghai Jiaotong University. This experiment focused on gathering physiological data for real-time emotion detection, and to explore the affective evolution during learning.

5.1 Experimental method/gathering affective data

Collecting good affective data is crucial to the results of the experiment; however, this is usually not as easy as in the computer vision or speech recognition. Cameras and microphones are reliable and easy to use, but there are factors influence the reliability of the bio-sensors. For example, whether or not the subject just washed her hands, how much gel she applied under an electrode, how tight the electrodes were placed, and even the humidity could affect the readings. What's more, emotion is subjective reaction to the environment, and people might not be totally aware of their feelings that the ground truth of the data is uncertain at some time.

We carefully design the experiment to obtain high quality physiological data for affect analysis. The subject sit in her quiet comfortable lab, learning as usual, the emotion was elicited naturally according to the situation, and the subject reported her own emotion by selecting one of the four emotions whenever she felt any change, which was used to label the data. The materials she read or watch was on her own selection with the only requirement that the difficulty level should be moderate. This set-up is more natural and closer to the real world, contrasting to those using guided imaginary technique (Clynes 1977) and classified emotion picture set (CSEA 2001) in many experiments, where emotions were subject-elicited, and might be just external expressions instead of internal feelings, and the presented emotions might be different from expected emotions.

The first data set was collected in iSpace where the subject lived and worked for 4 months. During the experiment, she wore the X-Vest (Leon et al. 2007) which provided the valence value and data from three raw biosensors (Skin Resistance (SR), Heart Rate (HR) and Blood Volume Pressure (BVP)). An UPnP control point was used to collect data from the X-Vest UPnP Device every 2 s (the sample rate was adjustable by the experimenter via a menu; the base rate that the X-Vest sent data was once a second). The subject was asked to conduct the experiment twice a day for 5 days, wearing the X-Vest while she was learning. Each session lasted at least 40 min. Ten learning sessions were collected. All the raw data, valence, and self reports were recorded together with time tag in a data file for further study and analysis.

The second data set was collected in a Shanghai e-Learning lab where the subject was studying for her PhD. Data were gathered from three sensors: a skin conductance (SC) sensor measuring electrodermal activity from the middle of the three segments of the index and ring fingers on the palm side of the left hand, a photoplethysmyograph measuring blood volume pressure (BVP) placed on the tip of the middle finger of the left hand, and a pre-amplified electroencephalograph sensor measuring EEG activity from the brain whose electrodes were placed on PC_z, A_1 and A_2 according to the *10–20 International System of Electrode Placement*. In our case, three EEG electrodes were sufficient (Lévesque 2006). Sensors and sampling were provided by the Thought Technologies ProComp5 suite, chosen because the suite was small enough to attach to a wearable computer. Signals were sampled at 256 Hz. The ProComp5 could automatically compute the heart rate (HR) as a function of the inter-beat intervals of the blood volume pressure, BVP and could separate different kinds of brainwaves into δ, θ, α, low and highβwith filters. The frequencies and relationships with emotion of each brainwave were listed in Table 1. Totally 18 40-min sessions were conducted within two weeks experiment. Each sample of data set II comprised three raw data (SC, BVP, EEG), the HR from BVP, five brainwaves from EEG, five power percentages of the brainwaves and the label with time tag.

5.2 Data preprocessing and feature extraction

For the data collected, data set I had totally 12,000 samples and data set II had totally 11,059,200 samples. For data set II such big data set would make data training and data classification

Table 1 Brainwaves and their relationship with emotion

Wave type	Frequency	When wave is dominant
δ Delta	0–4 Hz	Deep sleep
θ Theta	4–8 Hz	Creativity, dream sleep drifting thoughts
α Alpha	8–13 Hz	Relaxation, calmness, abstract thinking
Lowβ Beta	15–20 Hz	Relaxed focus. high alertness, mental activity. agitation,
Highβ Beta	20–40 Hz	Anxiety

very time-consuming. According to the fact that emotion won't change so frequently as much as 256 Hz, we fuse n samples into one sample to make it more efficient. We used very simple fusion algorithm that we computed the mean of the non-oscillating signals (SC, BVP, HR, the power percentages of the five brainwaves) and the FFT (Fast Fourier Transform Algorithm) of the oscillating signals (EEG, five brainwaves from EEG) as the corresponding values of the resulting sample. Finally we got one sample every 1 s When n=256 and we got one sample every 8 s when n=2,048.

Because the signals involved have different and complex sources, and different value ranges, we explored a feature-based approach to classification at the same time. The third kinds of samples were the features we extracted over 60-s raw samples with same label. Let X_n represent the value of the nth sample of the raw signal, $\widetilde{X}_n$ refer to the normalized signal (zero mean, unit variance):

$$\widetilde{X}_n = \frac{X_n - \mu_x}{\sigma_x}$$

Where μ_x and σ_x are the means and standard deviations of X as explained below. Following are six statistical features we investigated:

1. the means of the raw signals

$$\mu_x = \frac{1}{N}\sum_{n=1}^{N} X_n \tag{1}$$

Where N is the number of the raw samples within 60 min

2. the standard deviations of the raw signals

$$\sigma_x = \left(\frac{1}{N-1}\sum_{n=1}^{N}(X_n - \mu_x)^2\right)^{1/2} \tag{2}$$

3. the means of the absolute values of the first differences of the raw signals

$$\delta_X = \frac{1}{N-1}\sum_{n=1}^{N-1}|X_{n+1} - X_n| \tag{3}$$

4. the means of the absolute values of the first difference of the normalized signals

$$\widetilde{\delta}_x = \frac{1}{N-1}\sum_{n=1}^{N-1}\left|\widetilde{X}_{n+1} - \widetilde{X}_n\right| \tag{4}$$

5. the means of the absolute values of the second differences of the raw signals

$$\gamma_X = \frac{1}{N-2}\sum_{n=1}^{N-2}|X_{n+2} - X_n| \tag{5}$$

6. the means of the absolute values of the second differences of the normalized signals

$$\widetilde{\gamma}_X = \frac{1}{N-2}\sum_{n=1}^{N-2}\left|\widetilde{X}_{n+2} - \widetilde{X}_n\right| \tag{6}$$

The features (1)~(6) were chosen to cover and extend a range of typically measured statistics in the emotion physiology literature. One advantage of this features is that they can easily be computed in an online way (Vyzas and Picard 1999), which makes them advantageous for real-time recognition systems. However, the statistical features do not exploit knowledge we have about the physical sources of the signals. Factors such as hand washing, gel application, and sensor placement can easily affect the statistics. These influences combine with the subject's daily mood and with other cognitive and bodily influences in presently unknown ways, making them hard to model. In an effort to compensate for some of the non-emotion-related variations of the signals, let hX refer to the smoothed sample set after applying a Hanning window h (Oppenheim et al. 1998), hX=X×h, we also compute another set of three physiology-dependent features:

7. the means of the hX

$$\mu_{hX} = \frac{1}{N}\sum_{n=1}^{N} hX_n \tag{7}$$

8. the means of the first difference of the hX

$$\mu_{(hX_{n+1}-hX_n)} = \frac{1}{N-1}(hX_N - hX_1) \tag{8}$$

9. a form of contrast normalization of the hX

$$\overrightarrow{hX} = \frac{hX \times \min(hX)}{\max(hX) - \min(hX)} \tag{9}$$

We computed features (1)~(6) over the eight non-oscillating signals (SC, BVP, HR, the power percentages of the five brainwaves) and computed features (7)~(8) over SC and HR, and computed features (9) over SC, finally we got 53 features every 60 s.

5.3 Classification and results

We have got one sample sets for data set I (one sample every 2 s) and three sample sets for data set II: one sample every 1 s, one sample every 8 s, and one 53-feature sample every 60 s. The 53 features was reduced and selected with Fisher Projection (Duda and Hart 1973). And then two pattern classification methods were tested: support vector machine (SVM) and K-nearest neighbor (KNN).

SVM maps training vectors into a higher dimensional space and then finds a linear separating hyperplane with the maximal margin in this higher dimensional space. The mapping function is called the kernel function. We selected to use the radial basis function (RBF) kernel:

$$K\left(X_i - X_j\right) = e^{-\gamma\left\|X_i - X_j\right\|^2}, \quad \gamma > 0 \tag{10}$$

The problem of SVM then requires the solution of the following optimization problem:

$$\min_{w,b,\xi} \quad \frac{1}{2} w^T w + C \sum_{i=1}^{l} \xi_i \tag{11}$$
$$\text{subject to} \quad y_i(w^T \phi(x_i) + b \geq 1 - \xi_i, \xi_i \geq 0)$$

Where $K\left(X_i, X_j\right) \equiv \phi(X_i)^T \phi\left(X_j\right)$, l is the number of samples is in the training set, x_i is the attributes, y_i is the label.

By using the software LIBSVM (Chang and Lin 2007), we firstly searched the best parameter C and γ with cross-validation, then used the best parameter C and γ to train the whole training set, and finally tested over the testing set. KNN was tested with k=1~10 with native MATLAB functions.

5.3.1 Results of the classification

The results we obtained by applying the methods described above on data set I and data set II are listed on Tables 2 and 3. From Table 2 we could see that when we added the valence just used the three raw SC, BVP and HR signals, the recognition rates by LibSVM and KNN were 67.4% and 58.2% respectively, but when we added the valence information input from the X-Vest which achieved 85.2% correct rate, the rates rise to 81.8% and 77.3% respectively. This is no surprising because valence was also computed from SC, BVP and HR, using a UK patent classification technology that adding it to the final classifier formed a two-layer classifier. For data set II, when we just used SC, BVP and HR attributes, the rates are 68.6% and 59% by SVM and KNN, and when we just used the power percentages of five EEG brainwaves, the rates are 66.1% and 60.3%; but when we used these two groups of attributes together, the rates are as high as 86.5% and 76.0% respectively. From Table 3 we found that the brainwave power percentages contributed more than the sheer EEG powers from FFT, and the raw data were better than feature extraction in this case.

Table 2 Recognition rates of four learning emotions for data set I

Attribute space	LibSVM	KNN
SC, BVP, HR	67.4%	58.2%
SC, BVP, HR, valence	82.5%	77.3%

Table 3 Recognition rates of four learning emotions for data set II

Attribute space	LibSVM	KNN
SC, BVP, HR	68.6% (1 s)	59% (8 s)
EEG power % for brainwaves	66.1% (1 s)	60.3% (8 s)
SC, BVP, HR EEG power% for brainwaves	86.5% (1 s)	76.0% (8 s)
SC, BVP, HR EEG FFTs	63% (8 s)	59% (8 s)
52 features subsets (60 s)	80.5%	73.7%

5.3.2 Emotion evolution results

Kort et al. (2001) suggested that learning behavior would manifest itself in a spiral-like form i.e. a series of linked cycles separated in time. In order to learn how emotion evolves during learning the subject's emotion was displayed, in real time on a colored four quadrant diagram. Colors were often used to express arousal, where red represents emotions with high arousal and blue is calm and peaceful (Fagerberg et al. 2004; Fig. 5). From the emotion distribution (Table 4) for all the 55,200 samples (data set I+B1 sample every one second data set II) of 28 learning sessions, engagement and confusion were the most important and frequently occurred emotions in learning, and frustration was the least. We believe that the emotion distribution is related to the learning content. If the learning content is too difficult for the subject, then there should be more confusions and frustrations. This distribution result is reasonable because the learning content was selected by the subject herself and the difficulty level was moderate. The transition distribution (Table 5) showed that there were a lot of turns between engagement and confusion in both directions, and then a lot from confusion to frustration,

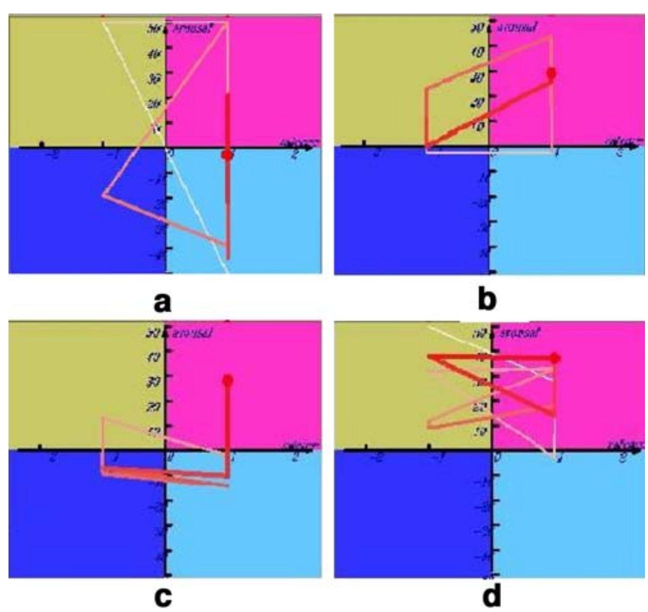

Fig. 5 Affective loop during learning process

Table 4 The emotion distribution for all the 55200 samples

Emotions	Sample numbers	Percentage
Engagement	19,892	36.0%
Confusion	18,321	33.2%
Frustration	5,769	10.5%
Hopefulness	11,218	20.3%
Total	55,200	100%

hopefulness to engagement, from confusion to hopefulness, and from frustration to hopefulness in frequency order. There occurred one loop during single session and two loops spanning two successive sessions. But there were more quasi-loops running through three quadrants within minutes (Fig. 5). Kort et al. (2001) didn't provide any information about the loop duration that whether it is within seconds or within hours. From our experiment, there were shorter loops within several minutes and longer loop more than 40 min. Of all the 28 learning sessions there were only three loops and some other quasi-loops which couldn't prove the Kort's learning spiral model. However, we hope these initial results will prove encouraging to others who have speculated on this relationship and hopefully will motivate more detailed work on this aspect.

6 Conclusion

The motivation driving this work is our desire to improve the performance of the educational experience by developing learning systems that recognize and respond appropriately to emotions exhibited by learners. This paper described the first steps towards realizing this vision by making use of physiological signals to sense emotion evolution during learning. As part of this work we described a cutting-edge digital home test-bed and a state of the art pervasive e-Learning platform which collectively act as a holistic system to evaluate our ideas. We gathered physiological data from one subject in two different places over many weeks and got a best-case classification rate of 86.5% which was yielded by SVM from raw data. This opens up a number of possibilities, such as providing emotional feedback to teachers or e-Learning systems for remote learners. As far as we know, this is the first report on emotion detection by data collected from close-to-real-world learning sessions. The correction rate is even better than those in lab-setting ones, for example, 81% for eight emotions (Picard et al. 2001). When brainwave signals were added to the other peripheral physiological data, the classification rate rose significantly from 68.6% to 86.5%, this suggested that there were close relationships between brainwaves and emotions during learning. From our experiment, we couldn't find sufficient empirical evidence of Kort's affective learning spirals. Other noteworthy observations included that engagement and confusion were the most important and frequently occurred emotions in learning, and that using the power percentages of brainwaves yielded better results than using the sheer FFT powers of brainwaves.

Table 5 The transition distribution for all the 147 emotion transitions

Transitions from	Transitions to			
	Engagement	Confusion	Frustration	Hopefulness
Engagement	×	52	2	3
Confusion	32	×	13	10
Frustration	6	5	×	7
Hopefulness	12	1	3	×

The results reported in this paper stem from first-stage studies of a much longer term research program. Whilst such results were very encouraging, and might have proved some basic principles, they still need further refinement. In particular we flag the following issues for future research:

- In this paper we were simply concerned about proving the underlying principle, but to take this analysis further, the learning material need to be more formally designed and be more diverse and representative.
- The multi-modal pattern analysis of signals from face, voice, body and the surrounding situation is likely to achieve better emotion recognition results which we will investigate in the next step.
- We built two prototypes to leverage the emotion detected, but there still need further investigations to establish an affective e-Learning model combining emotion feedback with the existing pervasive e-Learning platform.
- There are factors, other than learning, that could influence emotion; for example, who you are learning with; what you are learning; how are you learning; where you are learning; why you are learning and so on. It may be that combining these variables at the right degree is the key to a better affective learning model.

Our current experiments were based only on one participant; clearly, to make the results more reliable, we would need to have a bigger and more controlled sample. However, we contend, the value of our work is to demonstrate that the general principles involved are feasible, and hopefully via these encouraging initial results, motivate a more detailed study. The results we obtained with this data may not be the same for other subjects. However, the methodology for gathering and analyzing the data in this paper is not dependent on the subject; the approach described in this paper is general.

As is clear from our discussion, in this paper we are reporting results from the first phase of a much longer term research program. Our immediate aims are to design structured learning material, and gather data from more participants. Next we plan to research on and develop the affective learning model. And consider wider issues, such as the role of psychological state in learning. Finally we aim to deploy this model in the Shanghai e-Learning platform and evaluate it with real learners. Our hope is that this work should lead to better learning experiences and wider adoption of this technology. We will look forward to reporting on this work as it moves from research to real deployment over the coming years.

Acknowledgements Affective e-Learning is a collaborative research project conducted between the Digital Lifestyles Center, University of Essex, UK (http://digital-lifestyles.essex.ac.uk) and e-Learning Center, Shanghai Jiaotong University, China (http://www.dlc.sjtu.edu.cn). During the formation of the research ideas and models, we have enjoyed many discussions with our colleagues whose intellectual contributions we are most grateful for. We would especially like to thank Michael Gardner for his support and encouragement which made the visit to Essex possible, Chris Fowler for the discussion of learning psychology, Enrique Leon for his excellent valence recognition work, Malcolm Lear for the design of the early emotion sensing prototypes and Bernard Horan of Sun Microsystems for use of the wearable computing Sun-SPOT technology. This work was carried out as part of the 'Research on Context-aware Multi-modal Interaction Model and Key Technologies' Project Supported by the High-tech Research and Development Program of China (National 863 project) under Grant No. 2007AA01Z157.

References

Bassili, J. N. (1979). Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face. *Journal of Personality and Social Psychology*, *37*, 2049–2058.

Burleson, W., Picard, R. W., Perlin, K., & Lippincott, J. (2004). *A platform for affective agent research, 3rd International Conference on Autonomous Agents and Multi Agent Systems*. New York: ACM.

Callaghan, V., Clark, G., Colley, M., Hagras, H., Chin, J. S. Y., & Doctor, F. (2004). Inhabited intelligent environments. *BT Technology Journal*, *22*(3), 233–247, (215).

CSEA (2001). The international affective picture system: Digitized photographs: CENTER for the Study of Emotion and Attention, University of Florida.

Chang, C. C., & Lin, C. J. (2007). LIBSVM—a library for support vector machines.

Clarke, G., & Callaghan, V. (2007). Ubiquitous computing, informatization, urban structures and density. *Built Environment Journal*, *33*(2), 196–212.

Clynes, M. (1977). *Sentics: The touch of the emotions*. New York: Prism Pr Ltd.

CNNIC (2007). Statistical Survey Report on The Internet Development in China: China Internet Network Information Center.

Conati, C., & Zhou, X. (2002). *Modeling students' emotions from cognitive appraisal in educational games, the 6th International Conference on Intelligent Tutoring Systems*. Biarritz, France: Springer.

Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., et al. (2001). Emotion recognition in human–computer interaction. *IEEE Sinal Process Magazine*, *18*, 32–80.

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 AutoTutor. *Journal of Educational Media*, *29*(3), 241–250.

Duda, R. O., & Hart, P. E. (1973). *Pattern classification and scene analysis*. New York: Wiley.

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Automatic nervous system activity distinguishes among emotions. *Science*, *221*, 1208–1210.

Fagerberg, P., Ståhl, A., & Höök, K. (2004). eMoto: Emotionally engaging interaction. *Personal and Ubiquitous Computing*, *8*(5), 377–381.

Fowler, C. J. H. (1977). *The role of arousal in memory and attention*. London: University of London.

Haag, A., Goronzy, S., Schaich, P., & Williams, J. (2004). Emotion Recognition Using Bio-Sensors: First Step Towards an Automatic System. Paper presented at the Affective Dialogue Systems, Tutorial and ResearchWorkshop, Kloster Irsee, Germany.

Healey, J. (2000). *Wearable and automotive systems for affect recognition from physiology*. Cambridge, MA: MIT.

Isen, A. M. (2000). Positive affect and decision making. In M. Lewis, & J. Haviland (Eds.), *Handbook of emotions* (p. 720). Guilford, New York: Guilford.

Katsionis, G., & Virvou, M. (2005). *Adapting OCC theory for affect perception in educational software, the International Conference on Human–computer Interaction 2005* (*HCI 2005*). Las Vegas, Nevada USA.

Keller, J. M., & Suzuki, K. (1988). Use of the ARCS motivation model in courseware design. In D. H. Jonassen (Ed.), *Instructional designs for microcomputer courseware*. Hillsdale, NJ: Erlbaum.

Kort, B., Reilly, R., & Picard, R. W. (2001). An affective model of interplay between emotions and learning: Reengineering educational pedagogy-building a learning companion, IEEE International Conference on Advanced Learning Technologies. Madison, WI, USA.

Lévesque, J. (2006). *BCIA-EEG*. Université de Montréal, Canada.

Leon, E., Clarke, G., & Callaghan, V. (2007). A user-independent real-time emotion recognition system for software agents in domestic environments. *International Journal of Intelligent Real-time Automation*, *20*(3), 337–345.

Luo, Q., Wan, L. Y., & Wu, Y. W. (2006). The application of affective computing in e-Learning systems. *Open Education Research*, *12*(3), 80–84, (in Chinese).

Moore, M. G. (2007). The Theory of Transactional Distance. In M. G. Moore (Ed.), *The Handbook of Distance Education.* (pp. 89–108, 2nd ed.). Mahwah, NJ: Erlbaum.

Morishima, S. (2000). Real-time face Analysis and Synthesis using Neural Networks, *IEEE Workshop on Neural Networks for Signal Processing*.

NCES. (2006). *Homeschooling in the United States: 2003* (No. NCES 2006-042): National Center for Education Statistics.

Nicholson, J., Takahashi, K., & Nakatsu, R. (2000). Emotion recognition in speech using neural networks. *Neural Computing and Applications*, *9*(4), 290–296.

Oppenheim, A. V., Schafer, R. W., & Buck, J. R. (1998). *Discrete-time signal processing* (*2nd edn*). New Jersey: Prentice-Hall.

Ortony, A., Clore, G. L., & Collins, A. (1990). *The cognitive structure of emotions*. Cambridge: Cambridge University Press.

Papert, S. (1996). An exploration in the space of mathematics educations. *International Journal of Computers for Mathematical Learning*, *1*(1), 95–123.

Picard, R. W. (1995). *Affective computing*. Cambridge, MA: M.I.T.

Picard, R. W., Papert, S., Bender, W., Blumberg, B., Breazeal, C., Cavallo, D., et al. (2004). Affective learning—a manifesto. *BT Technology Journal*, *22*(4), 253–269.

Picard, R. W., Vyzas, E., & Healey, J. (2001). Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *23*(10), 1175–1191.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178.

Shen, L. P., & Shen, R. M. (2005). Ontology-based intelligent learning content recommendation service. *International Journal of Continuing Engineering Education and Life-Long Learning*, *15*(3–6), 308–317.

Sundström, P. (2005). *Exploring the affective loop*. Stockholm: Stockholm University.

Thomas, S. (2008). Pervasive scale: A model of pervasive, ubiquitous, and ambient learning. *IEEE Pervasive Computing*, *7*(1), 85–88.

Vyzas, E., & Picard, R. W. (1999). Offline and online recognition of emotion expression from physiological data, *The Third International Conference on Autonomous Agents*. Seattle, WA.

Liping Shen is a lecture of Computer Science and Engineering Dept. at Shanghai Jiaotong University (SJTU), and is also a PhD candidate for Computer Science. Her work and research focus on learning technology standards, personalized learning technologies and context-aware ubiquitous learning. When she visited Essex University in UK 2006, the affective e-Learning collaboration work was reported by New Scientist http://www.newscientisttech.com/article/dn10894-emotionaware-teaching-software-tracks-student-attention.html and EuroNews (broadcast on 13th February 2007). She has peer-reviewed articles published in journals such as International Journal of Continuing Engineering Education and Life-Long Learning, Lectures Notes in Computer Science, INFORMATICA, etc.

Victor Callaghan is professor of computer science at Essex University, director of the Digital Lifestyles Centre and leader of the Inhabited Intelligent Environments group at the Essex University. He graduated from the University of Sheffield with a B. Eng in Electronics and a PhD in Computing. His main expertise concerns intelligent embedded-agents, affective computing and end-user programming, especially applied to the creation of digital homes, intelligent buildings and other pervasive computing environments. Professor Callaghan has attracted in excess of £2 million of research funding, published over 100 papers and was responsible for setting up the University's Brooker Laboratory, Robot Arena and Intelligent Dormitory (iSpace) research facilities. He is a key member of the organisational team for major pervasive computing conferences (e.g. Pervasive Computing and Applications, Ubiquitous Intelligence and Computing, and Intelligent Environments) and journals (e.g. Pervasive Computing, Communications, Ubiquitous Computing and Intelligence and Intelligent Buildings).

Ruimin Shen is a distinguished professor of Computer Science and Engineering at Shanghai Jiaotong University (SJTU). He is also the founder and director of the world-renowned E-Learning Lab, and the founding dean of SJTU's Network Education College (15,000 students). Winner of several national awards in innovations in distance education and e-Learning, Shen's research and development specialty focus on cutting-edge technologies for "available anywhere and updatable anytime" distance education, including mobile learning, standard natural classrooms, knowledge discovery and data mining, data encoding and information retrieval. He has peer-reviewed articles published in top-tier journals such as International Journal of Distance Education Technologies, Lecture Notes in Artificial Intelligence, Lectures Notes in Computer Science, Computers and Education, etc.