Evaluation of mobile authoring and tutoring in medical issues*

Efthymios Alepis, Maria Virvou

(Department of Informatics, University of Piraeus, Piraeus 18534, Greece)

Abstract: Mobile computing facilities may provide many assets to the educational process. Mobile technology provides software access from anywhere and at any time, as well as computer equipment independence. The need for time and place independence is even greater for medical instructors and medical students. Medical instructors are usually doctors that have to treat patients on top of their tutoring duties. Mobile features are complementary to web-based features for desktop computers to a high extend in asynchronous e-learning environments. Time and place independence is also a considerable potential for medical students with overloaded educational duties. This paper examines the degree of the usefulness of mobile facilities for medical instructors who wish to author and manage their courses by using a mobile authoring tool. Furthermore, this paper investigates how acceptable and useful the mobile features of an e-learning system have been to medical students who have used the system, in comparison with the use of the system through a desktop computer. In addition, it investigates usability issues.

Key words: medical tutoring systems; mobile learning; affective computing

1. Introduction

In the last decade, there has been a growing interest in mobile technology and mobile networks. As a result, a great number of services are offered to the users of mobile phones including education. In the fast pace of modern life, students and instructors would appreciate using constructively some spare time that would otherwise be wasted, for example, when they are traveling or even when they are waiting in queue. Moreover, medical students and their instructors may have to work on lessons at any place, even when away from offices, classrooms and medical labs where computers are usually located. Assets of mobile interaction include device independence as well as more independence with respect to time and place in comparison with web-based education using standard PCs (personal computers).

Mobile education may be quite impersonal since the presence of a human instructor and human co-students may not be available. A remedy to this kind of problem may be given by providing affective interaction based on the users' emotional states. The recognition of emotions can lead to affective user interfaces that take into account the users' feelings and can adapt their behavior according to these feelings. Regardless of the various emotional paradigms, neurologists/psychologists have made progress in demonstrating that emotions play an important role in the process of decision-making and action-deciding (Leon, Clarke, Gallaghan & Sepulveda, 2007, pp. 337-345). Moreover, the way people feel may play an important role in their cognitive processes (Goleman, 1995). Recently, significant research effort has been put in the recognition of users' emotions while they interact with software

* Support for this work was provided by the General Secretariat of Research and Technology, Greece, under the auspices of the PENED-2003 program.

Efthymios Alepis, Ph.D., Department of Informatics, University of Piraeus; research field: informatics.

Maria Virvou, professor, Department of Informatics, University of Piraeus; research field: informatics.

applications. Picard (2003, pp. 55-64) pointed out that one of the major challenges in affective computing is to try to improve the accuracy of recognizing people's emotions. Improving the accuracy of emotion recognition may imply the combination of many modalities in user interfaces. Indeed, human emotions are usually expressed in several ways. Human faces, people's voices or people's actions may all reveal emotions.

Ideally, evidence from many modes of interaction should be combined by a computer system so that it can generate as valid hypotheses as possible about users' emotions (Busso, et al., 2004, pp. 205-211). It is hoped that a multimodal approach may provide not only better performance, but also more robustness (Pantic & Rothkrantz, 2003, pp. 1370-1390). As it is stated by ZENG, Pantic, Roisman and HUANG (2007, pp. 126-133), although the benefit of fusion (i.e., audio-visual fusion, linguistic and paralinguistic fusion, multi-visual-cue fusion from face, head and body gestures) for affecting recognition is expected from engineering and psychological perspectives, the knowledge of how humans achieve this fusion is extremely limited.

In previous work, the authors of this paper have implemented and evaluated with quite satisfactory results emotion recognition systems, incorporated in educational applications (Alepis & Virvou, 2006; Alepis, Virvou & Kabassi, 2009). As the next step, the authors have extended the affective educational system by providing mobile interaction between users and a mobile tutoring system. In many situations, this means that learning may take place at home or some other sites, supervised remotely and asynchronously by a human instructor but away from the settings of a real class.

The main characteristic of the proposed mobile medical tutoring system is that it combines evidence from two modes, namely the mobile device's microphone and keyboard, in order to identify users' emotions. The emotion recognition assumptions from the two modes are combined through a multi-criteria decision-making method. More specifically, the system uses Simple Additive Weighting (SAW) (Fishburn, 1967) for evaluating different emotions, taking into account the input of the two different modes and selecting the one that seems more likely to have been felt by the user. In this respect, emotion recognition is based on several criteria that a human tutor would have used in order to perform emotion recognition of his/her students during the teaching course.

After a thorough investigation in the related scientific literature, it is found that there is a shortage of educational systems that incorporate multi-modal emotion recognition, not to mention affective medical educational systems. Even less are the existing affective educational systems with mobile facilities. In a mobile context-aware (Lim & Aylett, 2007, pp. 317-330), intelligent affective guide is described, that guides visitors touring an outdoor attraction. The authors of this system aim mainly at constructing a mobile guide that generates emotions. On the contrary, the proposed educational system for medicine aims at recognizing users' emotions through their interaction with a mobile device rather than generating emotions.

As a second related approach, the authors found that Yoon, et al (2007, pp. 758-766) proposed a speech emotion recognition agent for mobile communication service. This system tries to recognize five emotional states, namely neutral emotional state, happiness, sadness, anger and annoyance from the speech captured by a cellular phone in real time and then it calculates the degree of affection, such as love, truthfulness, weariness, trick and friendship. In their approach, only data from the mobile device's microphone were taken into consideration, while in this research, the authors investigated a mobile bi-modal emotion recognition approach. Moreover, the proposed system of this research is incorporated in a medical educational application and data pass through linguistic and paralinguistic level of analysis. This derives from the fact that, in an educational application it should take into consideration "how" users speak, using a microphone, or type, using the keyboard (such as low or high voice, slow or quick typing speed), as well as "what" users say or type (such as correct answers or mistakes).

In view of the above, this paper describes a novel mobile tutoring system for medicine that incorporates bi-modal emotion recognition through a multi-criteria theory. In addition, it examines the degree of usefulness of mobile facilities for medical instructors who wish to author and manage their courses using a mobile authoring tool. Furthermore, the authors investigate how acceptable and useful the mobile features of an e-learning system have been to medical students that have used the system, in comparison with the use of the system through a desktop computer. In addition, the authors investigate usability issues. The approach that have been taken for this investigation is through a study that involves medical instructors and students who were asked to use both the mobile and standard desktop facilities of the medical educational system, so that the authors could compare usability issues among the two modes. As a result, most medical instructors and students found mobile facilities useful and easy to use. However, users who had previous computing experience would prefer to combine mobile facilities with standard desktop computing rather than depend on them exclusively. On the other hand, users with no computing experience were happy to use their mobile phones.

2. Overview of the mobile medical tutoring system

The main architecture of the mobile bi-modal emotion recognition system is illustrated in Figure 1. Medical students are able to use their mobile devices and interact with a pre-installed medical tutoring application. Their interaction can be accomplished orally (through the mobile device's microphone) or by using the mobile devices keyboard and of course by combining these two modes of interaction. All data are captured during the interaction of users with the mobile device through the two modes of interaction and then transmitted wirelessly to the main server. This means that the actual emotion recognition process takes place on a main server, not on mobile devices, which have limited capabilities in data processing and storing. When an emotional state is detected, the emotion detection server transmits this information wirelessly back to the mobile device that has initiated this process. Correspondingly, input actions are used as trigger conditions for emotion recognition by the emotion detection server. Finally, the resulting data sets that are comprised of all the detected input actions as well as possible recognized emotional states are stored in the emotion detection server's database.

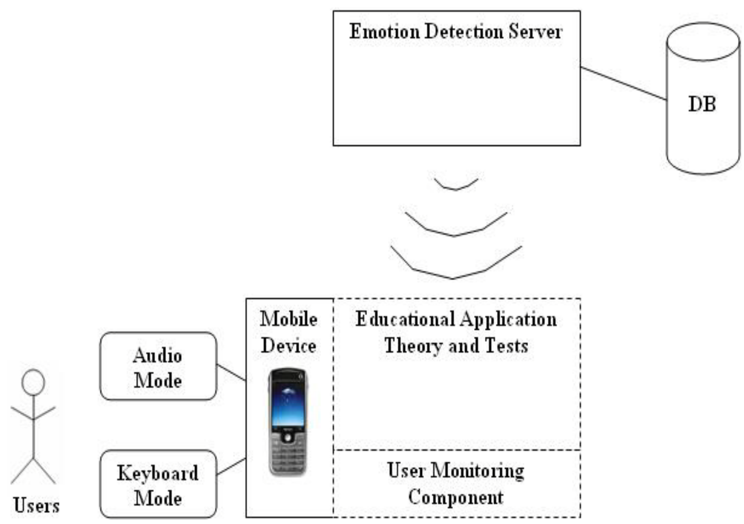

Figure 1 Architecture of the mobile emotion detection system

The discrimination between the medical students and their instructors is done by the application that uses the main server's database, and for each user a personal profile is created and stored in the database. In order to accomplish that, the user's name and password is always required to gain access to the medical tutoring system.

A snapshot of a mobile emulator, operated by a medical student is illustrated in Figure 2. Medical students may write their answers through the mobile device's keyboard, or alternatively give their answers orally, using their mobile device's microphone. In both cases, the data from the two possible modes of interaction are stored in the main system's database (emotion detection server), in order to be processed for emotion recognition purposes. During the short examinations, the system also tries to perform error diagnosis in cases where the answers have been incorrect. Error diagnosis aims at giving an explanation about a mistake taking into account the history record of each user and the particular circumstances where the error has occurred.

Medical instructors can also use the mobile medical tutoring system through the educational process. Medical instructors and their students are not only able to have easy access to the databases of the application, but they can also "communicate" with each other. The communication between them can be realized in many ways. By using a mobile phone (and thereby connecting to the application's mobile pages), medical instructors can send short messages via the short message service (SMS), either directly to their students (If they also have mobile phones) or by e-mail. Alternatively, medical instructors can store a message to the application's database. In this case, they have to declare that the name of the receiver and the tutoring system will use its audio-visual interface to inform her/him as soon as she/he opens the application.

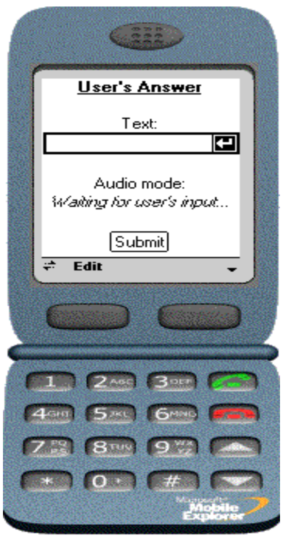

Figure 2 A medical student is answering a question of a test either using the keyboard or orally through the mobile devices' microphone

Medical students can use their mobile phone or palm top PC to interact with the educational application. They can connect to the mobile pages of the application and after the authentication process they may take tests for specific, relevant to medicine courses (Figure 2). User authentication is necessary for the creation and improvement of the user model of each student. Special emphasis should be given to the system's capability to accept written information via the mobile device's keyboard, or audio data via the mobile device's microphone as two alternative ways of bi-modal interaction (Figure 2). This specific bi-modal information is also very important for the system's emotion detection module to make assumptions for each user's possible emotional state.

An example of an instructor monitoring the progress of a student through a mobile phone and sending an SMS to him is illustrated in Figure 3. The application kept in the database completes records of the date and time when students enter the system, as well as the exact duration when they give tests. In this way an instructor can see how much time his/her students spend for the tests, so they can determine whether a test is difficult or not. Both teachers and students are able to send short messages (SMSs) containing remarks and additional information. The body of the SMS is entered in the “enter message” field and the name of the receiver is written in the “enter student name” field.

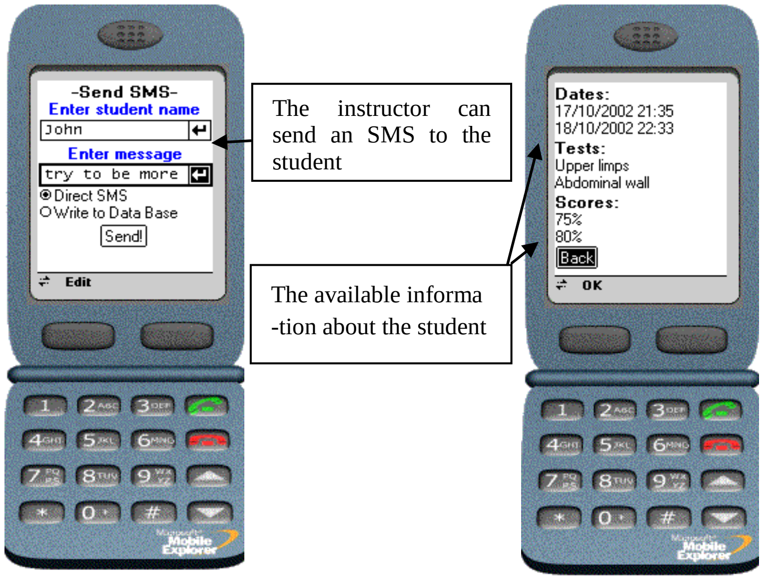

Figure 3 A medical instructor monitors the progress of a student and sends SMS

3. Multi-criteria for emotion recognition

The proposed mobile medical tutoring system incorporates a highly sophisticated emotion recognition module. As the first step in specifying the criteria that would lead to emotion recognition, the authors specified basic input actions through the mobile device’s keyboard and microphone that take place during the interaction of users with the medical tutoring system. These input actions provide information for the emotional states that may occur while a user interacts with an educational system. The overall functionality of this approach also incorporates user stereotypes (recognition of emotions through audio-lingual data and keyboard evidence) and is explicitly described in (Alepis, Virvou & Kabassi, 2007; 2008) for the case of interaction through a personal computer. However, analysing the empirical studies that lead to the specification of the weights for each input action requires more comprehensive writing and is beyond the scopes of this paper.

The basic function of the mobile affective system is to capture data from the available modalities and filter them in terms of some stereotypic variables for improving emotion recognition accuracy. Multi-level stereotypes in conjunction with a multi-criteria theory work as a filter with multiple layers where all captured data are weighted according to the stereotypic user model of each individual user, the users’ input actions that may reveal specific emotions and finally the system’s available modes’ ability in recognizing each emotional state. For example, in a situation where a user uses the mobile device’s microphone and the keyboard while interacting with the educational application, the available data are processed as illustrated in Figure 4. In the first level, the system tries to draw inferences about which the modality of interaction provides more evidence about the specific user’s feelings by taking into account the stereotype where this particular user belongs to. For example, a particular user may be more expressive orally rather than through the keyboard. In that case the microphone should be considered as a more important modality than the keyboard in conveying evidence about the user's emotions. In the early sessions of a new user when the system does not have sufficient information about the preferences of the particular user, user stereotypes provide inferences about which modality is likely to be more important about the user. These inferences are based on personal information about the user, such as gender (female/male), age, computer knowledge level, etc. A more comprehensive representation of the incorporation of user stereotypes into the emotion recognition module can be found in (Alepis, Virvou & Kabassi, 2008, pp. 523-532).

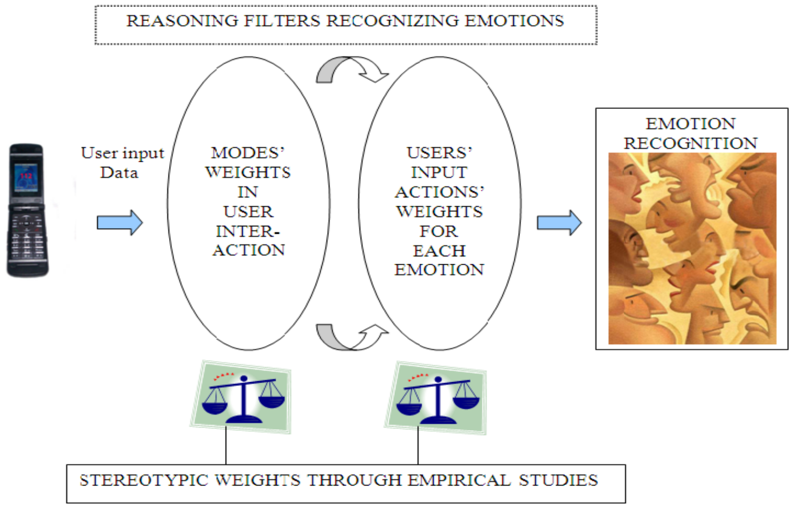

Figure 4 Recognition of emotions through stereotypic weights

Bi-modal user input actions are considered as criteria for evaluating all different emotions and selecting the one that seems more confident. There are six emotional states for the system to recognize, namely happiness, anger, sadness, disgust, surprise, as well as the neutral emotional state. More specifically, each emotion is evaluated first using only the criteria (input actions) from the keyboard and then only the criteria (input actions) from the microphone. According to the proposed approach, in cases where both modalities (keyboard and microphone) recognize the same emotion, the probability that this emotion has occurred is increased significantly. Otherwise, the mean of the values for each emotion is calculated and the one with the higher mean is selected.

Considering the mobile device's keyboard, there are the following categories of user actions: (1) user types normally; (2) user types quickly (speed higher than the usual speed of the particular user); (3) user types slowly (speed lower than the usual speed of the particular user); (4) user uses the "delete" key of his/her mobile device often; (5) user presses unrelated keys on the keyboard; and (6) user does not use the keyboard. These actions are considered as criteria for the evaluation of emotion with respect to the user actions through the keyboard.

Considering the users' basic input actions through the mobile device's microphone, there are 7 cases: (1) user speaks using strong language; (2) user uses exclamations; (3) user speaks with a high voice volume (higher than the average recorded level); (4) user speaks with a low voice volume (lower than the average recorded level); (5) user speaks in a normal voice volume; (6) user speaks words from a specific list of words showing an emotion; and (7) user does not say anything. These seven actions are considered as criteria for the evaluation of emotion with respect to what users say.

For the evaluation of each alternative emotion, the system uses the Simple Additive Weighting (SAW) method. However, the overall functionality of this approach exceeds the scopes of this paper, since the authors' main aim is to focus on the evaluation of the resulting system. For a thorough examination of the incorporation of an emotion recognition module into a sophisticated mobile system, it may refer to past work of the authors (Alepis, Virvou & Kabassi, 2008; 2009).

4. Evaluation

Software that is meant to help the educational process can be considered successful if it is approved by human instructors and is educationally beneficial to students. Otherwise, it may not even be included in the educational process and may not be accepted by its targeted users. Thus, evaluation of this kind of software is an important phase that has to follow development at all times.

The basic differences between a desktop and the authors' mobile educational system considering the modalities of interaction are mainly in their user interfaces. Desktop applications may offer a variety of multimedia tools such as speech engines, speech recognizers and animations with the additional "companionship" of animated agents whose behavior may be programmatically controlled. As these features are not currently supported by mobile phones, the graphical user interface of the application's mobile pages is much simpler. The mobile interaction may lack in user-friendliness, but it can also be very "handy" as medical students can give tests at any time and any place, simply by connecting to the main application through their mobile devices. They do not need to own a desktop computer, while a mobile phone is much cheaper and more frequently used.

For the evaluation of the resulting mobile educational system, 10 medical instructors as well as 100 medical students participated. All of the instructors who participated in the experiment were familiar with the use of computers. In addition, they had been trained for the use of the mobile medical system before the experiment.

When interviewed, all of the instructors confirmed that the resulting system had a user-friendly interface and that the mobile facilities were either useful or very useful. More specifically, 9 of them stated that they found the mobile facilities either useful or very useful both for the creation and the maintenance of their courses, whereas only 1 of them said that s/he had not used the mobile features at all during the creation of the course, but s/he found them useful during the maintenance of the course. The exact answers of instructors to questions about the mobile features of the authoring tool are illustrated in Figures 5 and 6. As expected, all of the 9 instructors who found useful the mobile features of the application for both phases, made clear that they had used the mobile facilities in a complementary way with a desktop computer, since the authoring process involves inserting a lot of data. Thus it would have been difficult for anyone to develop the whole course using a mobile phone.

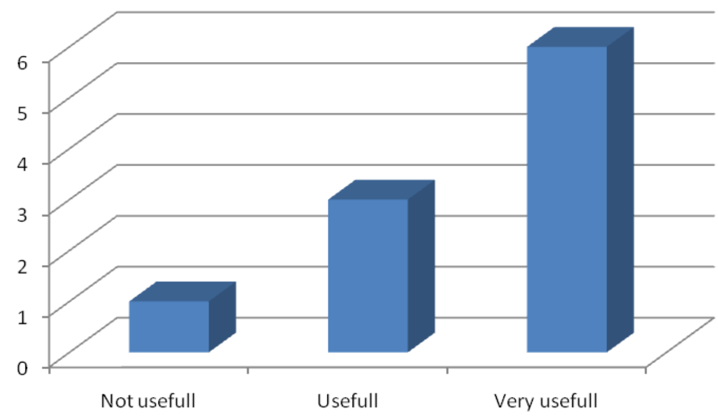

Figure 5 Instructors' opinions about the usefulness of mobile creation of courses

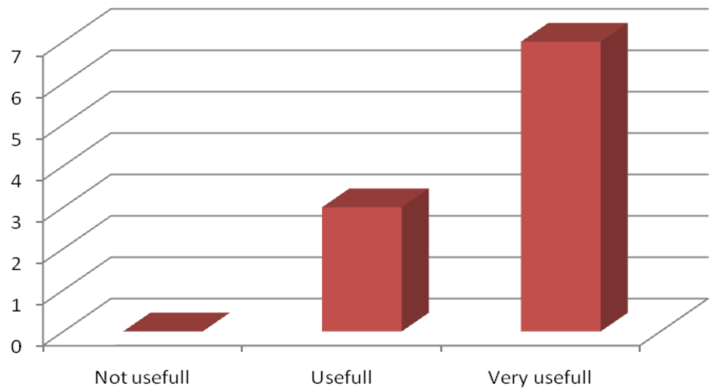

Figure 6 Instructors' opinions about the usefulness of mobile maintenance of courses

Additionally, the 100 medical student participants were asked to score the degree of usefulness of the resulting system after having used it. As a result of this evaluation study, medical students have appreciated both the desktop and mobile features of the e-learning system for different reasons. The affective desktop facilities were considered very user-friendly by students who had previous computing experience. However, one very important finding came up from the students who were not familiar with computers. Mobile facilities were preferred by these students. Most of the students, who do not have much experience in using computers, own a mobile phone and therefore know how to use it. These reasons make the mobile interaction "more attractive" and "accepted" by the majority of medical students. Figure 7 illustrates the medical students' scores, taking their values from 0 indicating that the resulting system was not useful at all, up to 5 indicating that the resulting system was found to be very useful to them.

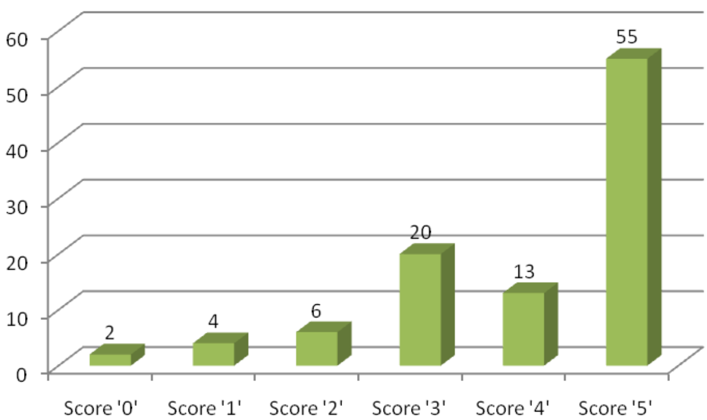

Figure 7 Medical students' scores for the degree of usefulness of the mobile tutoring system

5. Conclusions

This paper has described a mobile medical tutoring system that incorporates mobile technology in order to help students and instructors with many learning and training obligations. The system has been designed to provide the relatively new mobile facilities, while retaining high quality of the educational application with respect to high interactivity, personalization and user-friendliness. This work also shows how mobile devices may be used constructively in medical education by combining existing technologies of educational software with mobile features.

The research conducted, has resulted in the development of a sophisticated mobile medical tutoring system that has been evaluated among medical instructors and their students. The evaluation results were very encouraging and showed that the contribution of mobile software features to education can be appreciated by both instructors and students.

References:

Alepis, E. & Virvou, M. (2006). Emotional intelligence: Constructing user stereotypes for affective bi-modal interaction. *Lecture Notes in Computer Science, 4251*(LNAI-I), 435-442.

Alepis, E., Virvou, M. & Kabassi, K. (2007). Development process of an affective bi-modal Intelligent Tutoring System. *Intelligent Decision Technologies, 1*, 1-10.

Alepis, E., Virvou, M. & Kabassi, K. (2008). Affective reasoning based on “bi-modal interaction and user stereotypes”. *Studies in Computational Intelligence, 142*, 523-532.

Alepis, E., Virvou, M. & Kabassi, K. (2008). Knowledge engineering for affective bi-modal interaction in mobile devices, frontiers in artificial intelligence and applications. Proceedings of the *Eighth Joint Conference on Knowledge-Based Software Engineering, 180*, 305-314.

Alepis, E., Virvou, M. & Kabassi, K. (2009). Mobile education: Towards affective bi-modal interaction for adaptivity. *International Journal of Interactive Mobile Technologies, 3*, 40-45.

Alepis, E., Virvou, M. & Kabassi, K. (2009). Knowledge engineering aspects of affective bi-modal educational applications. *Communications in Computer and Information Science, 23*, 328-339.

Busso, C., DENG, Z., Yildirim, S., Bulut, M., Lee, C. M., Kazemzadeh, A., Lee, S., Neumann, U. & Narayanan, S. (2004, October 13-15). Analysis of emotion recognition using facial expressions, speech and multimodal information. Proceedings of the *6th International Conference on Multimodal Interfaces*, State College, PA, USA. New York: ICMI ’04, ACM, NY, 205-211.

Fishburn, P. C. (1967). Additive utilities with incomplete product set: Applications to priorities and assignments. *Operations Research*.

Goleman, D. (1995). *Emotional intelligence*. New York: Bantam Books.

Leon, E., Clarke, G., Gallaghan, V. & Sepulveda, F. (2007). A user-independent real-time emotion recognition system for software agents in domestic environments. *Engineering Applications of Artificial Intelligence, 20*(3), 337-345.

Lim, M. Y. & Aylett, R. (2007). Feel the difference: A guide with attitude! *Lecture Notes in Computer Science, 4722*(LNCS), 317-330.

Picard, R. W. (2003). Affective computing: Challenges. *Journal of Human-Computer Studies, 59*(1-2), 55-64.

Pantic, M. & Rothkrantz, L. J. M. (2003). Toward an affect-sensitive multimodal human-computer interaction (Vol. 91). Proceedings of the *IEEE*, 1370-1390.

Yoon, W. J., Cho, Y. H. & Park, K. S. (2007). A study of speech emotion recognition and its application to mobile services. *Lecture Notes in Computer Science, 4611*(LNCS), 758-766.

ZENG, Z., Pantic, M., Roisman, G. I. & HUANG, T. S. (2007, November 12-15). A survey of affect recognition methods: Audio, visual and spontaneous expressions. Proceedings of the *9th International Conference on Multimodal Interfaces*, Nagoya, Aichi, Japan. New York: ICMI ’07, ACM, NY, 126-133.

(Edited by Nicole and Lily)