This is a postprint of a paper published in
Human-Computer Systems Interaction: Backgrounds and Applications 3,
Advances in Intelligent Systems and Computing Volume 300,
Springer International Publishing, 2014
DOI: 10.1007/978-3-319-08491-6_5

Emotion recognition and its applications

A. Landowska, M. Szwoch, W. Szwoch, M.R. Wróbel, A. Kołakowska

Faculty of Electronics, Telecommunications and Informatics,

Gdansk University of Technology, Gdańsk, Poland

Abstract. The paper proposes a set of research scenarios to be applied in four domains: software engineering, website customization, education and gaming. The goal of applying the scenarios is to assess the possibility of using emotion recognition methods in these areas. It also points out the problems of defining sets of emotions to be recognized in different applications, representing the defined emotional states, gathering the data and training. Some of the scenarios consider possible reactions of an affect-aware system and its impact on users.

1 Introduction

Inevitably feelings play an important role no only in our relations with other people but also in the way we use computers. Affective computing is a domain that focuses on user emotions while he interacts with computers and applications. As emotional state of a person may influence concentration, task solving and decision making skills, affective computing vision is to make systems able to recognize and influence human emotions in order to enhance productivity and effectiveness of working with computers.

A challenging problem of automatic recognition of human affect has become a research field involving more and more scientists specializing in different areas such as artificial intelligence, computer vision, psychology, physiology etc. Its popularity arises from vast areas of possible applications. In this paper we focus on application of affective computing methods in software engineering, website customization, education and gaming. We propose a number of research scenarios to evaluate the possibility of using emotion recognition methods in these areas.

Section 2 briefly presents the idea and the methods of emotion recognition. Sections 3-6 describe the proposed scenarios to be applied in four different domains, including software engineering, website customization, education and gaming. Section 7 draws some conclusions on challenges, which have to be solved to implement the scenarios in a real environment.

2. Emotion recognition

The goal of human emotion recognition is to automatically classify user's temporal emotional state basing on some input data. There are dozens of definitions of emotions [Picard 2004], and in this paper we adopt the following distinction based on time: *emotion* is a reaction to stimuli that lasts for seconds or minutes, *mood* is an emotional state that lasts for hours or days and *personality* is an inclination to feel certain emotions. We use the term 'emotional state' to indicate current (temporary) state of a person irrespective of its origin (stimuli, mood or personality).

The stated goal may be achieved by one of many types of classifiers developed in the field of pattern recognition. The approach assumes several stages of classifier's construction which are: data acquisition and feature extraction, creation of the training set containing labeled data and classifier's learning.

In our research we assume that emotion recognition will be based on multimodal inputs: physiolgical sensors, video, depth sensors and standard input devices. This approach proved to provide more information to the recognition process as different data channels deliver valuable complementary information eliminating potential drawbacks of any individual input [Kapoor and Picard 2005].

Physiological sensors might be used to measure skin conductance, blood volume pulse, muscle impulses, respiratory signal, temperature or heart rate [Szwoch W 2013]. They are non-invasive but sometimes intrusive or not comfortable for users due to special equipment, which is required. Thus it is not possible in real-life situations when we want to determine emotions of people during usual learning or working processes. But taking specialized measurements of physiological signals can be used in some scenarios as well as for the enhancement or verification of a classifier's accuracy. A great many features may be extracted from physiological signals by calculating their mean, standard deviation, difference, Fourier transform, wavelet transform, high frequency and low frequency power, entropy, thresholding, peak detection etc. [Jerritta et al. 2011].

Video and depth sensors deliver us significant information on facial expression in a non intrusive way by using camera. Face expression may be represented by geometric or appearance features, parameters extracted from transformed images such as eigenfaces, dynamic models and 3D models. The difficulty of this approach is the need of image preprocessing and complex pattern recognition algorithms. One of the main problems with facial expression recognition is that it usually works well only in the case of a posed behavior and proper lightning [Szwoch M 2013]. Depth sensors, which usually use non-visual infrared light technology, are generally resistant to insufficient and uneven lighting conditions. Since introducing Microsoft Kinect depth sensors are also used for recognition of human poses, gestures and movements [Ren et al. 2011].

Standard input devices, such as keyboard and mouse, enable a completely unobtrusive way of collecting data, because no special hardware is needed and moreover it may be done during users' usual computer activities. Features extracted from keystrokes may be divided into timing and frequency parameters. Mouse characteristics include both clicking and cursor movement measurements [Kołakowska 2013].

The collected data has to be labeled, i.e. an emotion has to be assigned to each data sample. Most of the emotion recognition algorithms use discrete or dimensional models for affective state representation. The best recognized discrete model is Ekman six basic emotions, including: joy, sadness, fear, anger, disgust and surprise, whereas a three-dimensional PAD model proposes to represent an emotional state as a combination of valence, arousal and dominance values. Usually the labels for the training data are assigned according to specially designed questionnaires given to the users or on the base of independent observers' evaluations. However such labeling may not be objective, which in turn may result in poor accuracy of the trained system. To avoid this situation we are going to validate the labels provided by humans with the labels assigned on the base of physiological measurements.

A great many machine learning algorithms have been already applied in the task of emotion recognition, e.g. SVM, decision trees, linear discriminant analysis, Bayesian networks, naive Bayes, neural networks [Zeng et al. 2009]. An ideal emotion recognition method in the proposed real-life applications would be a combination of adaptive classifiers which could cope with high number of features of different types and would be able to improve its effectiveness with increasing amounts of training data continuously recorded during users' typical activities.

3. Software engineering

The main purpose of this part of our research is to use emotion recognition based on multimodal inputs to improve some aspects of software engineering process and to overcome the limitations of usability questionnaires [Kołakowska et al. 2013]. We focus on two areas of application in software engineering: usability testing and development process improvement.

3.1. Extended software usability testing

There is a lot of evidence, that human emotions influence interactions with software products. There is also a record of investigation on how products can influence human feelings [Hill 2009] and those feelings make people buy or not. Therefore investigating emotions induced by products is an object of interest of designers, investors, producers and customers, as well.

Software usability depends on multiple quality factors, such as functionality, reliability, interface design, performance and so on. All of the quality indicators can be improved indefinitely, but there is a point to stop optimizing – it is a customer satisfaction. Measuring this satisfaction with questionnaires may be misleading. We propose to extend usability testing with emotion recognition based on multimodal inputs. We have defined the following test scenarios with required emotional state distinctions depending on the scenario.

Scenario 1. First Impression test

First impression is a state that is evoked mainly by visual input in human-systems interaction and is created in a very short time (approximately 5 seconds). Research shows, that in web page design first impression is a good predictor of 10-minute usability opinion [Linggaard et al. 2006]. Many of the websites will not have any more time to make an impression than these 5 seconds – the first impression makes the users stay or quit. In first impression testing the most important distinction is to differentiate user's *interest* (*excitement*) from *boredom* or *disgust*. This scenario is especially dedicated to web page usability testing.

Scenario 2. Task-based usability test

The second usability scenario proposed uses cognitive walkthrough method [Blackmon et al. 2002], which is a task-based approach. Software usability evaluation in this method usually involves identification of typical tasks (which may be extracted from use case models) and the optimal processes for performing them (possibly derived from dynamic models, user stories or user instructions). The representative user group performs the tasks in a controlled environment with camera recording, biometric sensors and keystroke analysis tools. Registered channels are then a subject to further analysis of usability and emotional state fluctuation. This scenario is dedicated rather to applications designed to help the user to perform specific tasks and not for entertainment or content access systems. The purpose of emotion recognition in task–based usability testing is to differentiate *frustration* from *empowerment*.

Scenario 3. Free interaction test

The third usability scenario proposed is based on free interaction with application, which is supposed to evaluate overall user experience. There are no pre-defined tasks to be performed by representative user group; instead they are asked to freely interact with application under examination. This scenario is dedicated for entertainment and content access systems, but other applications may also benefit. The objective of emotion recognition in this scenario is the distinction of *engagement* from *discouragement*.

Scenario 4. Comparative test

Comparative scenario is a selection or combination of methods used in previously defined scenarios performed on two software versions or on the application and the main competitive software product.

3.2. Development process improvement

In each segment of the job market the most valuable employees are those who are highly productive and deliver high-quality products or services. A similar situation is with respected software developers [Wróbel 2013]. Employers require high work efficiency and high quality code. Unfortunately, these two requirements are often in conflict, as a computer program developed under time pressure is usually of low quality [MacCormack et al. 2003].

The purpose of this study is to verify the hypothesis that emotions have significant impact on software quality and developers' productivity. The aim is to answer the question on correlation between employee's emotional state and his work efficiency as well as quality of the developed software. The study will also determine the emotional states of IT professionals that support their work.

We have defined four research scenarios to explore multiple factors of the relationship between programmers' emotional states and their work, including the work environment, personal productivity and quality of developed code.

Scenario 5. IDE usability comparison

This scenario is an adoption of task-based usability test described in scenario 2. Integrated development environments (IDE) are one of the essential tools used by developers. Their advantages and disadvantages can significantly affect the emotional state of the users. Research will be conducted in a laboratory environment with biometric sensors. User group will be represented by both novice programmers - ICT students and ICT stuff with years of experience. The object of the research will be a set of popular IDEs. A developer will perform a series of programming tasks, such as compiling, debugging, refactoring, on three randomly selected environments, excluding those he uses the most frequently. Tests will evaluate the quality of those IDEs. However, the collected data will be used to investigate the individually differentiated impact of problems encountered in an IDE on developers' emotions. The goal of this scenario is to distinguish between the *frustration* and *empowerment*.

Scenario 6. Productivity and emotions

This scenario is designed to answer the question of whether and how emotional state affects the productivity of the programmer. The research will be conducted in a laboratory environment. Behavior of the maliciously prepared environment will evoke developers' emotions that may affect their productivity, measured for example by the number of lines of code per hour of work. In the first place *stress* associated with time pressure and *boredom* will be induced. The analysis of the collected data will determine the optimal emotional space for developer productivity.

Scenario 7. Code quality and emotions

This scenario, despite similarities to the previous one, should not be conducted in laboratory environment. It is hard to accurately evaluate the quality of the code developed in a short test. Therefore, to provide the reliable results, this scenario requires continuous monitoring of the emotional state of the programmer and the collection of incremental versions of the source code. Only the cross-examination of emotional states and source codes may lead to the designation of the correlation between quality and emotions. In this scenario, it is essential to detect emotions such as *empowerment*, *frustration* and *stress*.

Scenario 8. Tele and office working comparison

The last scenario is designed to detect whether there are differences in emotional states of programmers when working in office or at home. The number of telecommuters is growing rapidly in recent years. This research should be conducted in real work environments. This will be possible only after the development of reliable, non-intrusive methods of user emotional states recognition. The objective emotion recognition is to detect the whole range of emotions, particularly all those identified in the previous scenarios.

The scenarios 5 and 6 can be conducted in a laboratory environment. In this research, it is possible to use a biometric sensor to detect emotions of programmers. This will deliver more accurate recognition of emotions than with the previously developed non-intrusive methods. However, the implementation of the other two scenarios (7 and 8) will be possible only using non-intrusive methods for detecting the emotions of computer users.

As the computer is the primary working environment of programmer, the implementation of emotions recognition mechanisms in human-computer interface is a natural choice. However research, as well as proposed scenarios (except scenario 5), are sufficiently universal to be applied to many professions.

4. Education and e-education

There is lots of evidence, that some emotional states support learning processes and other suppress them [Hudlicka 2003, Picard 2003, Sheng et al. 2010]. The distinction of the two groups of emotional states in some cases is not obvious, for example such positive mood as hilarity is not good for learning processes, while slightly negative emotional states foster critical thinking and are appropriate for analytical tasks [Landowska 2013]. Automatic emotion recognition algorithms can help to explore this phenomena by making assessments of learner emotional states more objective than typical questionnaire-based investigations.

Scenario 1. Emotional templates of educational tasks.
The purpose of this scenario is investigation on emotional states that occur during different types of educational tasks. This investigation aims at identification of emotional templates of educational tasks, that can be defined as distinguishable sets of effective and counter-productive emotional states for solving specific task types. To perform this investigation representative set of educational tasks should be prepared and both learners' performance in task execution and his/her emotional state must be measured. Analysis of the correlation between performance and emotional states would enable to justify statements on effective and counter-productive emotions for specific task types, however a significant number of respondents should be engaged in order to make the thesis reliable. Information on effective emotional states can be then used in educational problems diagnosis, educational tool design or in further exploration of educational processes.

Scenario 2. Emotional stereotypes of learners.
Emotionality is one of the elements of human personality and may differ significantly based on in-born temper, previous experience and socialization process. However some emotional reactions are common for people living in one culture or having the same experience and similar characteristics is expected in educational processes. Learner affective stereotype is a definition of typical emotional states that might be observed in educational settings. It is expected, that novice learners will more frequently show symptoms of frustration, while more experienced ones could feel boredom. To support that thesis with evidence, emotional states of different (novice/experienced) students will be measured and recognized while they perform the same tasks set of growing difficulty. Learners' stereotypes can be then used in e-educational environments to adapt learning paths and/or interaction models, when no individual information on user is available.

Scenario 3. Evaluation of educational resources.

The goal of this scenario is evaluation of educational resources, especially those prepared for self-learning. In distance and electronic education one of the critical success factors is learner discipline in following provided learning path. When one fails to deal with fluctuation of motivation and attention, learning processes are paused or even abandoned. One of the frequently launched cause for course resignation is: “Boring resources”. In this scenario observation of student’s interaction with resources is combined with monitoring his/her emotional state in order to identify parts of resources that cause boredom. That information may be then used to remove weak points and improve overall resource quality. A set of different types of educational resources including recorded lectures, screencasts and interactive materials will be investigated. This scenario might be also used for quality evaluation of resources provided in virtual universities and other distance learning environments.

Scenario 4. Usability testing of educational tools.

In this scenario usability of educational tools is evaluated. Software usability tests are usually based on eye-tracking techniques and we propose to extend it with user emotion recognition, which can be a valuable information while evaluating user experience [Kołakowska et al. 2013]. Typical tasks performed with educational tools include: educational tool access, resource search, resource launch, performing interactive tasks, viewing results or feedback information, communication with teachers or class mates and more. More specific task description for the scenario will be performed using cognitive walkthrough method [Blackmon et al. 2002]. Then representative group of students will be asked to perform tasks in controlled environment that will additionally record and recognize their emotional states. Information on affect and its fluctuations (especially identification of frustration) can help to improve software products that are designed to support learning processes.

5. Enhanced websites customization

With the grow of the Internet, service providers collect more and more information about their users. Based on these data, content, layout and ads are displayed according to the user's profile. Adding information about the emotions of users could provide more accurate personality models of the users.

We have defined two scenarios, the first to explore how emotions of web surfers influence their behavior on websites, the second to determine what emotions are triggered by different types of on-line ads.

Scenario 1. Affective customization

The purpose of the scenario is to determine how emotions affect the way users consume information on the websites. The main expected outcome of this investigation is a list of factors that, in conjunction with a specific mood, increase the time spent on the website. The study will examine the impact of the following factors: page layout, content and photo sets. Users are intended to review a set of prepared web pages (with different values of factors). Based on biometric sensors and cameras their mood will be recognized. These data will be aggregated with information about the time spent on each site. Analysis of the results will help to determine for example which website layout would interest bored person and which is the best for angry one.

Scenario 2. Advertisement reaction model

The revenue model for a significant number of websites is based on the on-line advertising [Dubosson-Torbay et al. 2002]. However users describe them as uninformative, ineffective and very intrusive [McCoy et al. 2007].

The aim of this scenario is to find the most eye-catching and interesting advertisement types for different groups of Internet users. In the laboratory environment, the emotional reaction will be measured in response to various formats of on-line ads. Additionally, using eye-tracking tool, information about the user focus on advert will be collected. Set of information about time elapsed before noticing the ad's, user emotional response and duration of focused attention will allow to choose the appropriate type of advertising, depending on the target audience.

6. Video games

Video games belong to the wide area of entertainment applications. Thus, assuming the existence of human emotions and in fact basing on them, they attempt to make the player to become emotionally attached with them. As the primary goal of a video game is to entertain the player [Adams 2009], each video game try to allow the player to fulfill his or her “dream”. Standard video games try to do it in different ways depending on their genre and involving such elements as good gameplay, immersing storytelling, novelty, graphics and so on [Adams 2009].

Although video games belong to applications in which emotions naturally play an important role, only few of them try to incorporate their players’ affective state into the gameplay. Such games can be referred as affective or more properly affect-aware games. The importance of affect in delivering engaging experiences in entertainment and educational games is well recognized. Potential profits for affect-aware video games are not to be underestimated. Unfortunately, this “affect-awareness” is usually “statically” built-in the game at its development stage basing on the assumed model of

so called representative player [Adams 2009]. There are two problems with such attitude. Firstly, each player differs in some way from that averaged model. Secondly, and more important, player's affect state can change even radically from session to session making almost impossible to predict the current user emotions at the development stage.

There are a lot of reasons that can influence upon human's behavior during playing the game. They could be divided into factors connected with the game, such as increasing monotony or becoming accustomed player, and to game independent factors which are connected with current physical and mental condition of the player. The first group of reasons may be in some extent predicted or estimated by the game designer but that is impossible to the reasons of the other group. That is why the real-time recognition of player's affect may become such important for video games industry in the nearest future. Video games that are able to dynamically react to the recognized current player's emotions we can call a truly affect aware video games.

Scenario 1. Emotional templates in video games

Though generally people are able to express wide spectrum of emotions, not all of them are observed while playing video games. Expressed emotions depend on many factors such as game's genre, player's genre and experience, and on many other predictable or even unpredictable factors [Adams 2009]. Moreover, emotional reaction in the situation highly depends on individual personality and even current mood or unpredictable external factors.

The goal of this scenario is to verify which emotions are common while playing video game and which are rare. The scenario will also allow verifying the hypothesis that for specific game genres and specific group of game players some expressions are more common then others.

A typical test in this scenario will cover recording of player emotions while playing several different video games with especially prepared scenes. Additional questionnaire will allow to classify the player's age, gender and gaming experience.

Scenario 2. Stimuli adaptation

Games try to attract the player's attention in the game to tie it as long as possible with the it. Game developers use interesting story, high quality graphics and an intense arousal to keep the attention of the player.

The goal of this scenario is to verify, whether long, frequent, repeated stimulation cause a negative reaction of the player due to stimuli adaptation. After a specified time, the player can stop to react on stimuli, and it is possible he or she will be expected more and more powerful experience. Is it necessary to keep the attention of the player at the highest level, or perhaps we should interlace periods of intense emotional arousal with periods of silence, which will allow the player's senses to rest.

Scenario assumes that study group of players will be exposed to intense stimuli during the game. The time after which the player's reaction to stimuli disappears and also the moment of weariness of the game will be tested. During the experiment, the observer and the player will determine the moment at the time or time period in which there was a loss of response to stimuli.

Scenario 3 Player reaction to external signals on different levels of immersion

Sometimes computer games players so strongly entered into the virtual world, they stop to notice the real world around them. It is important to monitor the depth of players' engagement to detect when they stop responding to external stimuli. This can help for example to protect players against addiction. When a player is too absorbed in the play, affect-aware game can reduce its attractiveness, causing return to real world.

During the experiment player's reaction to external stimuli will be investigated. The problem may be to get the appropriate involvement in the game. Player reaction (rapid / quiet, fast / slow) will determine the degree of his engagement. During the experiment various types of "disturbing" the player during the game - for example, verbal expression, noise, etc. will be used.

Scenario 4 The influence of affect feedback on player's satisfaction

Sometimes a video game can become boring or too stressing for the player. Detecting such situations would allow giving a proper feedback to the player changing the current gameplay character. On one hand such capability could improve satisfaction of experienced players and on the other it could protect novice or young players from excessive violence. Adapting feedback to stimulate the player can also be used in therapy in suppressing negative emotions by proper stimulation.

The purpose of the scenario is to check whether the affect-aware games can increase players' satisfaction from playing. For this scenario a specially developed affect aware video game will be used. The scenario assumes the use of a questionnaire, in which the player will determine the level of satisfaction within the game when increasing and decreasing difficulty according to the detected emotions.

7. Conclusions

The paper has presented a number of possible applications of emotion recognition methods in different areas such as software engineering, website customization, education and gaming. Although some of the presented research scenarios are ready to be applied, in the case of most of them a few challenging problems still have to be solved. First of all the limitations of emotion representation models not always let us precisely describe the actual feelings of a user. We often do not realize the real number

of possible emotional states, which should be considered in an application. Even if we are able to define a set of emotions, the fuzzy nature of emotional states and their instability along time entail subsequent difficulties. Moreover the quality of the training data and the way emotion labels are assigned strongly influence the the results of the training algorithm. Finally the accuracy of recognition process often does not fulfill the requirements of a system working in a real environment. All this does not let us ignore high uncertainty of emotion recognition methods, especially when combining separate models. This is an open research problem requiring investigation.

References

[Adams 2009] Adams E (2009) Fundamentals of Game Design. New Riders Publishing

[Blackmon et al. 2002] Blackmon M, Polson P, Kitajima M, Lewis C (2002) Cognitive Walkthrough for the Web, In Human Factors in Computing Systems: Proceedings of CHI 2002. ACM, New York: 463–470

[Dubosson-Torbay et al. 2002] Dubosson-Torbay M, Osterwalder A, Pigneur Y (2002) E-business model design, classification, and measurements. Thunderbird International Business Review 44(1): 5-23

[Hill 2009] Hill D (2009) Emotionomics: Winning Hearts and Minds, Kogan Page Publishers, USA

[Hudlicka 2003] Hudlicka E (2003) To feel or not to feel: The role of affect in human-computer interaction. International Journal of Human-Computer Studies 59(1): 1-32

[Jerritta et al. 2011] Jerritta S, Murugappan M, Nagarajan R, Wan K (2011) Physiological signals based human emotion recognition: a review. In: Proc. IEEE 7th International Colloquium on Signal Processing and its Applications

[Kapoor and Picard 2005] Kapoor A, Picard RW (2005) Multimodal affect recognition in learning environments. In: Proceedings of the 13th annual ACM International Conference on Multimedia, pp 677 - 682

[Kołakowska et al. 2013] Kołakowska A, Landowska A, Szwoch M, Szwoch W, Wróbel M (2013) Emotion Recognition and its Application in Software Engineering. In: Proc 6th International Conference on Human Systems Interaction, pp 532-539

[Kołakowska 2013] Kołakowska A (2013) A review of emotion recognition methods based on keystroke dynamics and mouse movements. In: Proc 6th International Conference on Human Systems Interaction, pp 548-555

[Landowska 2013] Landowska A (2013) Affective computing and affective learning – methods, tools and prospects, EduAction. Electronic education magazine (in printing)

[Linggaard et al. 2006] Linggaard G, Fernandes G, Dudek C, Brown J (2006) Attention web designers: You have 50 miliseconds to make a good first impression, Behaviour and Information Technology, 25(2): 115-126

[MacCormack et al. 2003] MacCormack A, Kemerer CF, Cusumano M, Crandall B (2003). Trade-offs between productivity and quality in selecting software development practices. Software, IEEE, 20(5), 78-85

[McCoy et al. 2007] McCoy S, Everard A, Polak P, Galletta DF (2007) The effects of online advertising. Communications of the ACM 50(3): 84-88

[Picard 2003] Picard R (2003) Affective computing: challenges. International Journal of Human-Computer Studies 59(1): 55-64

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

[Ren et al. 2011] Ren Z, Meng J, Yuan J, Zhang Z (2011) Robust hand gesture recognition with kinect sensor. In: Proc 19th ACM international conference on Multimedia, pp 759-760

[Sheng et al. 2010] Sheng Z, Zhu-ying L, Wan-xin D (2010) The model of e-learning based on affective computing. In: Proc 3rd IEEE International Conference on Advanced Computer Theory and Engineering (ICACTE), V3-269

[Szwoch M 2013] Szwoch M (2013*)* FEEDB: a Multimodal Database of Facial Expressions and Emotions In: Proc 6th International Conference on Human Systems Interaction, pp 524-531

[Szwoch W 2013] Szwoch W (2013*)* Using Physiological Signals for Emotion. In: Proc 6th International Conference on Human Systems Interaction, pp 556-561

[Wróbel 2013] Wróbel MR (2013) Emotions in the software development . In: Proc 6th International Conference on Human Systems Interaction, pp 518-523

[Zeng et al. 2009] Zeng Z, Pantic M, Roisman G, Huang T (2009) A survey of affect recognition methods: Audio, visual, and spontaneous expressions. IEEE Transactions on Pattern Analysis and Machine Intelligence, 31(1):39-58