Evaluating User's Emotional Experience in HCI: The PhysiOBS Approach

Alexandros Liapis, Nikos Karousos, Christos Katsanos, and Michalis Xenos

Hellenic Open University, School of Science and Technology
Parodos Aristotelous 18, Patra Greece, 26 335
{aliapis,karousos,ckatsanos,xenos}@eap.gr

Abstract. As computing is changing parameters, apart from effectiveness and efficiency in human-computer interaction, such as emotion have become more relevant than before. In this paper, a new tool-based evaluation approach of user's emotional experience during human-computer interaction is presented. The proposed approach combines user's physiological signals, observation data and self-reported data in an innovative tool (PhysiOBS) that allows continuous and multiple emotional states analysis. To the best of our knowledge, such an approach that effectively combines all these user-generated data in the context of user's emotional experience evaluation does not exist. Results from a preliminary evaluation study of the tool were rather encouraging revealing that the proposed approach can provide valuable insights to user experience practitioners.

Keywords: User Emotional Experience, Human Computer Interaction, Evaluation, Physiological Signals, Emotions.

1 Introduction

People's daily interaction with technology, including personal computers, tablets, and mobile phones, has increased the need for usability. Although the available technology is rather mature, interaction with it can still be frustrating [1-2]. Thus, evaluating and designing for user emotional experience (UEX) is growing in importance.

So far usability evaluation studies are mainly focusing on task-related metrics, such as task success rate and time-on-task. Such metrics are an important indicator of users' performance, but lack in qualitative insight about other factors of user experience [3], such as emotions. Questionnaires, interviews and video analysis can provide such qualitative data, but these methods are time consuming and prone to subjectivity [4-6]. In an attempt to address these limitations, new and innovative approaches such as facial expression recognition [7] and speech tone and keystroke analysis [8-9] have been introduced. Towards the same direction, collecting and analyzing data from users' physiology (e.g. heart rate, respiration, skin conductance) is also a promising approach. Physiological signals are directly connected with emotions [10] and their study could result in the establishment of new user-centered design techniques.

M. Kurosu (Ed.): Human-Computer Interaction, Part II, HCII 2014, LNCS 8511, pp. 758–767, 2014.

Emotions recognition and analysis (Fig. 1) are gaining interest in the human-computer interaction (HCI) domain, and especially in usability evaluation studies [11]. Existing approaches to interpret physiological signals in terms of emotions suffer from two important limitations. First, the recognition success of existing physiological signals datasets [12-13] used for emotion analysis relies on contexts that induce intense emotions, such as watching a scary movie, listening to a favorite song, major hardware failures and gaming. However, identification of subtle emotions is of more interest in typical HCI tasks and remains an open research topic. Additional, existing approaches in the HCI domain, attempt to recognize one or two emotions [14-17], thus ignoring any other emotions that users may have felt during an interaction session, which might lead in serious misunderstandings of UEX. Mandryk and Atkins [18] have proposed a psychophysiological method that can continuously monitor and also recognize more than one emotional state. However, it targets a specific domain (i.e. gaming and entertainment) and its application remains rather challenging for a practitioner.

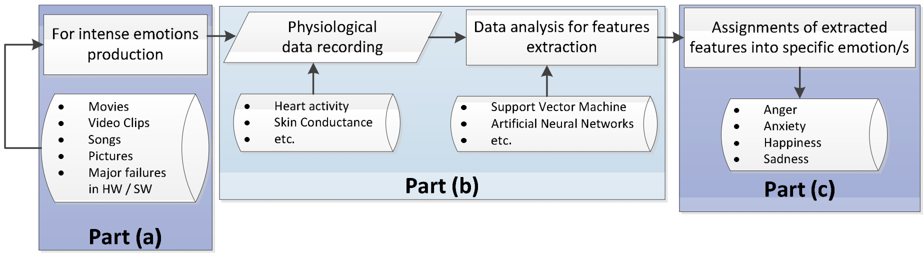

Fig. 1. Interpretation of physiological signals into emotions. (a) emotions induction, (b) physiological data recording and analysis, (c) interpretation of extracted features into emotions.

In this paper, a new tool-based evaluation approach for evaluating UEX during HCI is presented. The proposed approach combines user's physiological signals (e.g. heart rate, blood volume pressure, skin conductance), observation data (e.g. users' face recording, screen recording) and self-reported data (e.g. responses in questionnaires, interviews) in an innovative tool (PhysiOBS) that allows continuous and multiple emotional states analysis. To the best of our knowledge, such an approach that effectively combines all these user-generated data in the context of UEX evaluation does not exist. PhysiOBS supports continuous and in-depth evaluation of UEX in a straightforward and easy way. It also combines multiple data sources for both subtle and intense emotions recognition.

The rest of the paper is structured as follows. First, a brief background on physiological signals and emotion analysis is provided. Next, the proposed tool-based approach is delineated along with a presentation of research papers that mention the need for such approach in the HCI domain. In addition, results from a preliminary study in which practitioners used PhysiOBS to evaluate UEX are also presented. The paper concludes with a discussion of the implications of the presented work and directions for future research.

2 Background: Physiological Signals and Emotions

Changes in both the external and internal of a user's body can be measured through physiological signals. Physiological measurements along with other traditional metrics such as questionnaires and interviews have been used in many HCI studies [15-17] offering a new perspective in usability evaluation.

This section describes the advantages and disadvantages of physiological signals along with their relation with emotions and emotion structure theories.

2.1 Advantages and Disadvantages of Physiological Signals

Physiological signals are derived from vital organs, such as the heart and brain. Some of the most-widely used physiological signals are the following:

1. **Heart rate:** measures the electrical activity of the heart;
2. **Skin conductivity** (Sweat)**:** measures the resistance of the skin and it is one of the most well-studied physiological signals in the literature;
3. **Muscle tension:** measures the electrical activity generated by muscles;
4. **Respiration rate:** measures the stretch amount of a person's chest. It is a metric that needs to be treated carefully because it can be affected by cardiac function.

Special and sophisticated sensors systems (e.g. Electroencephalograph, Galvanometer and Electrocardiograph) have been developed in order to support researchers and practitioners in both data acquisition and analysis. In addition characteristics such as objectivity, multidimensionality, unobtrusiveness and continuity [19-22] reinforced the use of physiological signals and made them a valuable asset for usability studies.

However, physiological measurements have some limitations [23]. First, data acquisition depends on specialized and costly devices. Second, physiological measurements can be noisy because of various external factors such as fluctuations in room temperature, user's general health condition and environment humidity levels. Application of filters can alleviate such issues, but only to a certain extent. Finally, the experimental conditions along with sensors attachment can also affect users' physiological signals.

2.2 Emotion Theories and Physiological Signals

The human body is a complex system that reacts to various external stimuli. These stimuli can affect a person's psychology causing a variety of emotional states such as happiness, enthusiasm, frustration or boredom.

William James and Carle Lange theory, known as the "James-Lange theory of emotion", refers to emotions as an interpretation of a psychological state which can be identified through physiological signals [24]. According to this theory, an external stimulus leads to a physiological reaction. The psychological reaction depends on how one interprets these physical reactions. For instance, a walk in the woods and an unexpected encounter with a wild animal can increase one's heart beats and trigger a body tremble reaction. In James-Lange theory, interpretation of physical reactions would conclude that the person is afraid "I am trembling, therefore I am afraid". To this direction [25-26] were the first who studied the relations between physiological signals and emotions, concluding to four types of relations:

1. **one-to-one relation:** one physiological signal is capable to define a unique emotion;
2. **many-to-one relation:** more than one physiological signals are needed in order to define an emotion;
3. **one-to-many relation:** a physiological signal is associated with more than one emotions;
4. **many-to-many relation:** several physiological signals are associated with several emotions.

So far, the last relation is the most plausible and has been adopted by several scientific domains such as HCI.

James-Lange theory of emotion was questioned by [27], who proposed the "two-factor theory of emotion". According to the latter, emotions are neither purely physical nor purely cognitive reactions, but a combination of both. The theory posits that physical reactions must be interpreted along with the situation that someone is facing. Therefore, a fast pounding heart could be interpreted as anxiety, if someone is taking part in exams and as fear if someone encounters a wild animal. To the same direction, our tool-based approach offers to evaluators four views: user's screen capture, user's face recording, user's physiological signals and user's self-reported data. Having available all these perspectives at the same time, UEX evaluation may be more reliable than considering only physical reactions.

2.3 Emotion Structure Approaches

Two main theories-approaches of emotion structure have been established in the literature. A discrete approach supported by Ekman [28] and a dimensional approach supported by [29]. Ekman's approach uses six discrete emotional states: anger, fear, sadness, enjoyment, disgust and surprise. These emotional states can be recognized in all cultures and are gender-independent. By contrast, in the approach proposed in [29], emotions can be characterized using two dimensions: Valence and Arousal. Facial muscular activity and unsolved tasks (high arousal – low valence) have been found to be negatively correlated [30], and this fostered the use of the Valence-Arousal space in emotion analysis. Finally, it should be noted that both approaches are still used by researchers and practitioners, forming two schools of thought. Thus, our tool-based approach takes into consideration both emotion structures theories, offering appropriate supportive mechanisms to evaluators.

3 The PhysiOBS Approach

In this section, a new tool-based evaluation approach of UEX during human-computer interaction is presented. A researcher can combine user's physiological signals (e.g. heart rate, blood volume pressure, skin conductance), observation data (e.g. users' face recording, screen recording) and self-reported data (e.g. questionnaires, interviews) through an innovative tool (PhysiOBS) that allows continuous and multiple emotional states analysis. PhysiOBS supports continuous and in-depth evaluation of UEX in a straightforward and easy way. It also combines multiple data sources, for both subtle and intense emotions recognition.

3.1 The Need for an Emotion Oriented Evaluation Tool

Approaching physiological signals from the perspective of an additional evaluation parameter, Wilson and Sasse [16] used them in order to assess video and audio quality of multimedia conferencing (MMC) systems. Their evaluation approach used three dimensions a) stress (user cost), b) satisfaction and c) performance. The weightiness of each dimension depends on the purpose of MMC use. Physiological measurements analysis revealed that physiological responses can be detected even in degradation of both video and audio quality.

Lin et al. [14] used physiological signals for stress detection and correlated them with traditional usability metrics. Experiment participants' were instructed to complete three stages of a video game as quickly as possible and with a minimum number of mistakes. Each stage offered a different difficulty level. Data analysis revealed a positive correlation between physiological signals and game difficulty.

Ward and Marden [15] examined whether physiological measurements can be used instead of traditional metrics in web usability studies. In a between-subject study, two groups of users performed a website navigation scenario. The first group navigated in a well-designed website and the second one in an extremely bad-designed website. During their navigation, three physiological signals (skin conductivity, blood volume pressure, and heart rate) were recorded. Results didn't reveal significant differences between the two groups, but did reveal differences between individuals.

Along the same direction, [31] related physiological signals to traditional metrics used in web usability evaluation. In a within-subject study, a group of 42 subjects took part in a website navigation scenario. In one trial users had at their disposal navigation help from an artificial face in cases of navigation problems, whereas in the other trial they weren't provided with this help. Physiological measurements from participants that used the artificial face didn't reveal any significant variations.

A common characteristic of all these studies is the use of physiological signals in combination with other methods, such as questionnaires, interviews and video analysis, to collect data about the user experience. To this end, a holistic approach that can combine and support analysis of these user-generated data in an effective way would be of great value for practitioners.

In a different perspective, [32] used physiological signals in a domain called "Affective computing". Affective computing can create new ways of communication between humans and machines, by enabling machines to respond to users' emotions. To this end, users' emotion recognition is a prerequisite. Piccard et al. [32] achieved to recognize eight emotions with high levels of accuracy using physiological signals from one actor in a twenty days experiment.

Schreirer et al. [17] used a hacked computer mouse (random delays were introduced) with the aim to evoke intensive frustration episodes to participants. While participants used this hacked mouse, two physiological signals were collected (skin conductivity and blood volume pressure). Using a Hidden Markov Model as a feature extraction technique, they succeeded to automatically detect frustration events. This method gave a 50% accuracy level in frustration detection for 21 out of 24 users.

Mandryk et al. [18] used physiological signals in order to detect users' emotions while engaged with entertainment technologies. Participants played a video game against the computer, a friend and a stranger. In all three conditions, their physiological signals were continuously recorded and a fuzzy logic system converted them to a Valence-Arousal space. Then, a new fuzzy logic system was used in order to convert the Valence-Arousal space into discrete emotions.

The above studies based their emotion detection success on contexts that induce intense emotions, such as hardware failures and gaming. However, identification of subtle emotions is also of interest in typical HCI tasks, and remains an open research topic.

3.2 PhysiOBS Interface and a Typical Usage Scenario

PhysiOBS is a Windows-based tool and has been developed in C#. The aim of the tool is to support researchers and practitioners in the demanding task of UEX data analysis. PhysiOBS is meant to be used as a tool for post-study analysis, and thus requires, as a prerequisite, all users' data sources appropriately synchronized. PhysiOBS will be soon available for download at http://quality.eap.gr/en. In the following, the main interface of the tool (Fig. 2) along with a typical UEX evaluation scenario, are presented.

First, the evaluator adds at least one video (user's or screen recording). If both user's video and screen recording are available, the evaluator can watch them concurrently (Fig. 2, part a). In the example presented in Fig 2 (part a), the screen recording video also includes eye fixations and saccadic movements as captured by an eye-tracker. Thereafter, the evaluator can perform a typical video observation analysis supported by the tool functionality. For instance, the user of PhysiOBS can define tasks/subtasks and assign them to specific time periods (Fig. 2, part b). Extra information such as duration, result (successful/unsuccessful) and general evaluator's comments can be added for each user's task.

More importantly, user's physiological signals can also be inserted into PhysiOBS (Fig. 2, part c). Embedded semiautomatic processes, such as signal normalization and statistical analyses reported in [33], can be applied to provide a general overview of each physiological signal, basic characteristics and identified areas of potential emotional interest. Research-based guidelines [34] to complementary support emotional state identification are also provided to the evaluator. In specific, the evaluator assigns an emotion to a user-defined time period from a list of basic emotions [28] with specific associated characteristics, such as facial expressions and body movements [35-36]. To this direction an extra report, produced by user's answers analysis about their emotional state, is also provided to evaluators.

The result of the analysis process is represented as a series of emotional periods and emotion transitions (Fig. 2, part d). Color coding denotes different identified emotions (e.g. red: anger, coral: anxiety) and can be adjusted through the tool interface. In addition, all emotional periods can be optionally enriched with user's self-reported data in the form of comments. The evaluators' sense-making of the available data is supported by navigation controls (Fig. 2, part e), which are synchronized across all available views. An also important functionality provided by the tool is the save/load project option. The evaluator can save each participant's evaluation in order to edit it later.

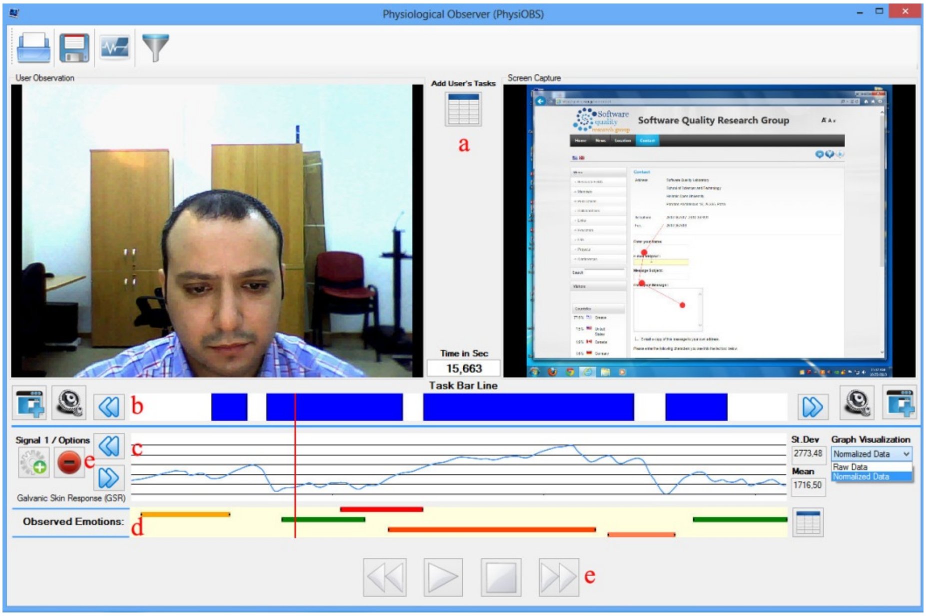

Fig. 2. PhysiOBS: a tool that combines physiological measurements, observation and self-reported data. (a) concurrent view of user's video and screen recordings, (b) task/subtasks view, (c) physiological signal(s) view, (d) observed emotions view with adjustable color coding, (e) navigation controls, synchronized across all available views.

3.3 PhysiOBS Preliminary Evaluation Results

Results from a preliminary evaluation study that was conducted in the Software Quality Research Laboratory (http://quality.eap.gr/en/lab) involving five HCI experts and two software practitioners were rather encouraging. Participants had at their disposal a set of user-generated data collected through a previous usability study and were asked to perform an analysis using PhysiOBS. The study indicated that PhysiOBS use can decrease time and perceived effort required to evaluate UEX from user-generated data. Furthermore, practitioners using PhysiOBS reported that the tool enabled a more in-depth UEX evaluation. In specific, participants confirmed that the simultaneous use of all available data sources can contribute different insights in the context of UEX evaluation. Study participants also provided feedback on the tool. For instance, two participants argued that for emotion selection they would prefer a wizard-like functionality with embedded help, instead of a simple menu.

4 Conclusions and Future Work

For many usability evaluation studies, emotions are no longer a supplementary parameter but a must. This paper presented PhysiOBS, an innovative tool-based evaluation approach of user's emotional experience during human-computer interaction. PhysiOBS combines multiple data sources in order to support continuous and in-depth evaluation of UEX in a straightforward and easy way. Results from a preliminary study that involved five practitioners revealed that the proposed tool-based approach could provide valuable help in such evaluations, offering an in-depth analysis of UEX.

One of our future aims is to provide additional automation in the emotion identification process based on physiological measurements of participants involved in typical HCI tasks. To this end, we are already planning studies to produce such emotionally-labeled datasets. Towards a more rigorous evaluation of our proposed tool-based approach, one future aim is to conduct a between-subjects study in which one group of evaluators will be provided with PhysiOBS to analyze user study data while the other will follow its working practices, and then compare findings between the two groups.

Acknowledgments. This paper has been co-financed by the European Union (European Social Fund – ESF) and Greek National funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) (Funding Program: "Hellenic Open University").

References

1. Lazar, J., Jones, A., Shneiderman, B.: Workplace user frustration with computers: An exploratory investigation of the causes and severity. Behaviour & Information Technology 25(2007), 239–251 (2007)
2. Igbaria, M., Chakrabarti, A.: Computer anxiety and attitudes towards microcomputer use. Behaviour & Information Technology 9, 229–241 (1990)
3. Lazar, J., Feng, J.H., Hochheise, H.: Research Methods in Human-Computer-Interaction. John Wiley & Sons Ltd. (2010)
4. Sanderson, P.M., Fisher, C.: Exploratory Sequential Data Analysis: Foundations. Human-Computer Interaction 9, 251–317 (1994)
5. Carter, K., Anderson, B.: Can video research escape the technology? Some reflections on the problems and possibilities of A.V. research. In: SIGCHI Bulletin, vol. 27, pp. 112–114 (1989)

6. Ambady, N., Rosenthal, R.: Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. Psychological Bulletin 111, 256–274 (1992)
7. Bartneck, C., Lyons, M.J.: HCI and the face: Towards an art of the soluble. In: Jacko, J.A. (ed.) Human-Computer Interaction, Part I, HCII 2007. LNCS, vol. 4550, pp. 20–29. Springer, Heidelberg (2007)
8. Kim, J.-H., André, E.: Emotion recognition using physiological and speech signal in short-term observation. In: André, E., Dybkjær, L., Minker, W., Neumann, H., Weber, M. (eds.) PIT 2006. LNCS (LNAI), vol. 4021, pp. 53–64. Springer, Heidelberg (2006)
9. Epp, C., Lippold, M., Mandryk, R.L.: Identifying emotional states using keystroke dynamics. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2011), pp. 715–724. ACM, New York (2011)
10. Wagner, J., Jonghwa, K., Andre, E.: From Physiological Signals to Emotions: Implementing and Comparing Selected Methods for Feature Extraction and Classification. In: IEEE International Conference on Multimedia and Expo, pp. 940–943. IEEE press (2005)
11. Experience Dynamics, `http://dev.experiencedynamics.com/blog`
12. DEAP DataSet, `http://www.eecs.qmul.ac.uk/mmv/datasets/deap/`
13. MIT Media Lab, `http://affect.media.mit.edu/share-data.php`
14. Lin, T., Hu, W.: Do Physiological Data Relate To Traditional Usability Indexes? In: 17th Australia conference on Computer-Human Interaction: Citizens Online: Considerations for Today and the Future, pp. 1–10 (2005)
15. Ward, R.D., Marsden, P.M.: Physiological responses to different WEB page designs. International Journal of Human Computer Studies 59, 199–212 (2003)
16. Wilson, G., Sasse, M.A.: Do users always know what's good for them utilising physiological responses to assess media quality. In: McDonald, S., Waern, Y., Cockton, G. (eds.) People and Computers XIV-Usability or Else, pp. 327–339. Springer, London Limited (2000)
17. Scheirer, J., Fernandez, R., Klein, J., Picard, R.W.: Frustrating the user on purpose: A step toward building an affective computer. Interacting with Computers 14, 93–118 (2002)
18. Mandryk, R.L., Atkins, M.S.: A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. Human-Computer Studies 65, 329–347 (2007)
19. Kivikangas, J.M., Ekman, I., Chanel, G., Järvelä, S., Salminen, M., Cowley, B., Henttonen, P., Ravaja, N.: Review on Psychophysiological Methods in Game Research. In: Proceedings of 1st Nordic DiGRA. DiGRA (2010)
20. Kramer, A.F.: Physiological Metrics of Mental Workload: A Review of Recent Progress. In: Damos, D.L. (ed.) Multiple-Task Performance, pp. 279–328. Taylor and Francis, London (1991)
21. Ikehara, C.S., Crosby, M.: Physiological Measures Used for Identification of Cognitive States and Continuous Authentication, Atlanta, USA (2010)
22. Bellur, S., Sundar, S.S.: Psychophysiological responses to media interfaces. In: ACM SIGCHI, Atlanta, GA,USA (2010)
23. Calhoun, B.H., Lach, J., Stankovic, J., Wentzloff, D.D., Whitehouse, K., Barth, A.T., Brown, J.K., Li, Q., Oh, S., Roberts, N.E., Zhang, Y.: Body sensor networks: A holistic approach from silicon to users. In: Proceedings of the IEEE, pp. 91–106 (2012)
24. James, W.: Discussion: The physical basis of emotion. Psychological Review 1, 516–529 (1894)
25. Cacioppo, J.T., Tassinary, L.G.: Inferring psychological significance from physiological signals. American Psychologist 45, 16–28 (1990)

26. Cacioppo, J.T., Tassinary, L.G., Berntson, G.G.: Psychophysiological science. In: Handbook of Psychophysiology. Cambridge University Press, Cambridge (2000)
27. Schachter, S.: The Interaction of Cognitive and Physiological Determinants of Emotional State. Advances in Experimental Psychology 1, 49–80 (1964)
28. Ekman, P.: An argument for basic emotions. Cognition and Emotion 6, 169–200 (1993)
29. Posner, J., Russell, J.A., Peterson, B.S.: The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. Development and Psychopathololy 17, 715–734 (2005)
30. Branco, P., Firth, P., Encarnac¸ao, L.M., Bonato, P.: Faces of emotion in human–computer interaction. In: CHI 2005 Extended Abstracts on Human Factors in Computing Systems, pp. 1236–1239. ACM Press, New York (2005)
31. Foglia, P., Prete, C.A., Zanda, M.: Relating GSR Signals to traditional Usability Metrics: A Case Study with an anthropomorphic Web Assistant. In: IEEE International Instrumentation & Measurement Technology Conference, Victoria, Canada, pp. 1814–1819 (2008)
32. Picard, R.W., Vyzas, E., Healey, J.: Toward Machine Emotional Intelligence: Analysis of Affective Physiological State. IEEE Transactions on Pattern Analysis And Machine Intelligence 23, 1175–1191 (2001)
33. Vyzas, E., Picard, R.W.: Affective Pattern Classification, In: Symposium, Emotional and Intelligent: The Tangled Knot of Cognition (October 1998)
34. Gunes, H., Piccardi, M.: Bi-modal emotion recognition from expressive face and body gestures. Journal of Network and Computer Applications 30, 1334–1345 (2007)
35. Fridlund, A.J., Izard, C.E.: Electromyographic Studies of Facial Expressions of Emotions and Patterns of Emotions. In: Cacioppo, J.T., Petty, R.E. (eds.) Social Psychophysiology: A Sourcebook. Guilford Press, New York (1983)
36. Bartneck, C., Lyons, M.J.: HCI and the face: Towards an art of the soluble. In: Jacko, J.A. (ed.) Human-Computer Interaction, Part I, HCII 2007. LNCS, vol. 4550, pp. 20–29. Springer, Heidelberg (2007)