EEG-based Measurement of Subjective Parameters in Evaluations

Daniel Cernea[1,2], Peter-Scott Olech[1], Achim Ebert[1], and Andreas Kerren[2]

[1] University of Kaiserslautern, Department of Computer Science,
Computer Graphics and HCI Group, P.O. Box 3049, D-67653 Kaiserslautern, Germany
{cernea, olech, ebert}@cs.uni-kl.de
[2] Linnaeus University, Computer Science Department,
ISOVIS Group, Vejdes Plats 7, SE-35195 Växjö, Sweden
{andreas.kerren}@lnu.se

Abstract. Evaluating new approaches, be it new interaction techniques, new applications or even new hardware, is an important task, which has to be done to ensure both usability and user satisfaction. The drawback of evaluating subjective parameters is that this can be relatively time consuming, and the outcome is possibly quite imprecise. Considering the recent release of cost-efficient commercial EEG headsets, we propose the utilization of electro-encephalographic (EEG) devices for evaluation purposes. The goal of our research is to evaluate if a commercial EEG headset can provide cutting-edge support during user studies and evaluations. Our results are encouraging and suggest that wireless EEG technology is a viable alternative for measuring subjectivity in evaluation scenarios.

Keywords: Evaluation techniques, Brain-Computer Interface (BCI), Electroencephalographic (EEG) interaction.

1 Introduction

Measuring user satisfaction has been an important factor before introducing market-ready goods and services for decades. Therefore, evaluations have been an essential step for developments in the field of Human-Computer Interaction (HCI). In fact, they are the decisive factor if a solution will really work as intended, based upon usability guidelines and user experiences. Setting up and performing an evaluation can be a time consuming process. Often, however, the real outcome is doubtful. Also, many widely used and accepted techniques have disadvantages. For example, "thinking aloud" encourages subjects to verbalize their thoughts and emotions, thus changing the users' behavior.

In our approach, a consumer market EEG device, the wireless *Emotiv EPOC*[1] headset, is used to add value to standard evaluation methods like questionnaires.

[1] http://www.emotiv.com

In Proceedings of the 14th International Conference on Human-Computer Interaction (HCII '11), poster paper, pages 279-283, volume 174 of CCIS, Orlando, Florida, USA, 2011. Springer.

Instead of focusing on the EEG headset as an interaction device, we want to shift the focus to use the headset as an evaluation support device and even as a standalone device for fast first impressions.

In the following sections, we present an evaluation of the capabilities of the EPOC device itself and highlight the results. We continue by employing the EEG headset in two scenarios for measuring the emotional reactions of test subjects.

2 Related Work

The history of Electroencephalography (EEG) dates back to the late 19th century [8]. Its use has been mainly medical in order to record the electrical activity of the brain, e.g., in the field of neuro-science to detect abnormal brain activity, like to diagnose epilepsy. Brain-Computer Interaction techniques (BCI) have been researched recently in order to provide interaction possibilities, e.g., for physically handicapped people [4]. Grimes et al. [2] advanced the field by investigating brain waves and how to classify working memory load with the help of an EEG devices. Scherer et al. [7] introduced an EEG-controlled Virtual Keyboard. The work of Horlings et al. deals with emotion recognition by using an EEG device [3]. Similarly, Mikhail et al. introduced a feature selection mechanism in [5], which can detect emotions out of noisy EEG data. In the work of Campbell et al. [1], the EPOC headset is used to interact with mobile phones (e.g. dial by using the headset instead of touch). Ranky et al. [6] propose to use the EPOC headset as an interaction device to control a robotic arm. After a training period, they obtained quite satisfying results.

3 The Wireless EPOC EEG Headset

Originally designed as gaming device, the EPOC headset comes with preprogrammed features, which can be quickly employed in evaluation. These offer real-time feedback about the emotional reactions of a user. Using such an affordable EEG device for evaluation purposes means that the results are not externally influenced, except by the brain activity of the subject. In terms of hardware, the EPOC headset is quite non-intrusive, as it is enabled by a wireless connection and very light. It is capable to measure electrical brain activity by means of 14 saline non-dry sensors.

The EPOC device has convenient features in its framework that enable the detection of a set of facial expressions and emotional states. Sadly, the algorithms for this are proprietary. As such, owners of an EPOC device have to rely on the manufacturer's encoding without much proof for the correctness of the detections. To overcome this and validate the output of the EEG headset, we have compared the results of the EPOC headset against the results from commonly accepted evaluation methods during two evaluations involving the detection of facial expressions and emotions.

The following tests involved 12 subjects with a basic level of knowledge in computer usage. In terms of distribution, the test group contained four women and eight men, aging from 21 to 52, and with an average age of 29.75 years. The users have diverse ethnicity and varied cultural background.

3.1 Detection of Facial Expressions

Facial expressions, especially when executed subconsciously, have the ability to reflect the persons' inner feelings. Usually, video logs cannot be analyzed in real-time and need to be interpreted after the completion of the subject's task. To circumvent this problem, we considered capturing the facial expressions of the test subjects via the EPOC device. As the framework and coding of the detection of the facial motions is not accessible to us, we started by validating the results offered by the headset against more common evaluation methods—in this case, video log analysis.

The subjects were equipped with the EPOC headset and positioned in front of a monitor and a webcam. They were given a sequence of words on the screen that represented facial expressions (e.g., smile, blink, etc.). Then, they were asked to execute them while the text was displayed and for as long and as often as they considered. The facial expression texts were given to the users randomly with each expression appearing at least 3 times. After the task was completed and the EPOC output data was available, the video logs and the information given by the headset were compared. The results show that the correct detection of the facial expressions varies between 70-100%, depending on the particular facial expression.

3.2 Detection of Emotional Reactions

After checking the correctness of the results for facial expressions detection, we turned our attention to evaluating its capacities in correctly assessing the emotional state of the user. The subjects were given tasks that should provoke emotional reactions. These emotional responses measured by the EPOC were compared with the results of a questionnaire, posed to the subjects after the tasks were completed. If the results from the EEG headset and the questionnaire are close to identical, we can argue that the EPOC device is a viable alternative or supportive method in evaluation.

The emotional states we tested for were *calmness*, *meditation*, *engagement in dexterity and mental tasks*, and *excitement*. The tasks the subjects had to complete involved watching informative videos, listening to music, and playing dexterity and memory games. Each task would generate one particular emotional response, and both the questionnaire and the EPOC output would focus only on reading that emotion. Afterwards, the users confirmed that the tasks were appropriate for generating the expected emotional response.

During the tasks, the device returns a constant stream of values for each of the emotions mentioned above, at a rate of approximately 4 Hz. As these values fluctuate, we can deduce the changes that affect the user. The recognition of an emotion is triggered in two ways: by computing the angle of increasing slopes of the values during the task and by computing the difference between the maximum and minimum during a given task.

An emotion is considered as triggered if we have a slope bigger than 30 or 60 degrees as well as if the max-min difference is more than 30% or 60%. These four thresholds, together with the neutral state for an emotion, depict a set of five possible levels for the values of an emotion as returned and interpreted by the EPOC device. Similarly, the questionnaire that assesses the emotional reaction of each user at the end of the task inquires about a particular emotion, and the subject has to answer in a 5-level point system: strongly disagree, disagree, neutral, agree and strongly agree (represented by 0, 0.25, 0.5, 0.75, and 1).

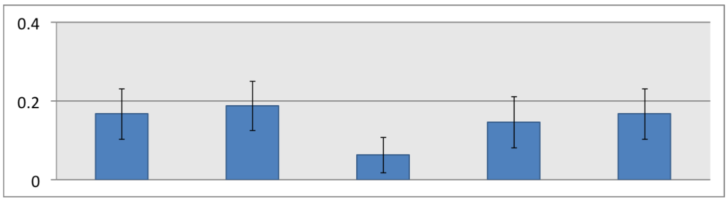

Fig. 1. Average difference between the EPOC device output and the questionnaire results for the emotional evaluation (left to right: calmness, meditation, engagement on dexterity task, engagement on metal task, excitement), with both on the same 5-point system: 0, 0.25, 0.5, 0.75, and 1. Black vertical lines represent the standard deviation for each average difference.

Figure 1 presents the results of our validation. Note that the average difference between the answers from the questionnaire and the EPOC output is between 0.185 and 0.068. To put this into context, an average difference of 0.25 means that on average the distance between the user's answer in the questionnaire and the EPOC output was one unit out of five.

To statistically validate our results, we also computed the standard deviation for the average differences. The results are promising as the maximum standard deviation is 0.071. Also, we executed a paired sample t-test for the data obtained from each task (questionnaire results and EPOC output). In each of the five cases, there is no significant difference between the paired sets, suggesting that the EPOC device can offer quite accurate interpretations of human emotional states.

4 Evaluation Scenario

Once we had confirmation for the correct results obtained via the EPOC headset in terms of facial expressions and emotional states, we continued with an evaluation scenario. We tested the emotional effects a spot-the-difference task would have on users. For this mostly mental/visual task, we presented the users three pairs of slightly different images sequentially.

The emotional states that were incorporated are *engagement*, *excitement*, *satisfaction* and *frustration*. In detecting the levels of satisfaction and frustration, the emotional output of excitement and various facial expressions—like smiling or clenching—were considered. The highest average difference obtained between the EPOC outputs and the questionnaire answers was 0.33 for the excitement level. The other average differences are situated between 0.2-0.25, similar to the previously described validation. The standard deviation was computed for the differences, resulting in a maximum value of 0.12. A paired sample t-test for the data obtained from each emotion showed that all paired sets present no significant difference, except for the excitement emotion. A possible reason for this is that emotional excitement in an intrinsically mental task is hard to define.

5 Conclusion

We tested the detection of facial expressions and emotional states with the Emotiv EPOC device and obtained promising results. Building on these results, we employed the EEG headset in an evaluation scenario producing encouraging outcomes in terms of using the headset as an evaluation device. While not knowing the exact purpose of wearing the EPOC device, one user even mentioned that he “would use this device in market research”, further suggesting that an evaluation approach based on a mobile EEG could open the door towards real-time efficient subjectivity measurement.

Acknowledgments. This work was supported by the German Research Foundation (DFG, grant number 1131) as part of the International Graduate School (IRTG) in Kaiserslautern on “Visualization of Large and Unstructured Data Sets”.

References

1. Campbell, A., Choudhury, T., Hu, S., Lu, H., Mukerjee, M.K., Rabbi, M., Raizada, R.D.S.: NeuroPhone: Brain-Mobile Phone Interface using a Wireless EEG Headset. In: Proc. of MobiHeld '10, ACM Press (2010), 3-8.
2. Grimes, D., Tan, D.S., Hudson, S.E., Shenoy, P., Rao, R.P.N.: Feasibility and Pragmatics of Classifying Working Memory Load with an Electroencephalograph. In: Proc. of CHI '08, ACM Press (2008), 835-844).
3. Horlings, R., Datcu, D., Rothkrantz, L.J.M.: Emotion Recognition using Brain Activitiy. In: Proc. of CompSysTech '08, ACM Press (2008), Article 6.
4. Leeb, R., Friedmann, D., Mueller-Putz, G.R., Scherer, R., Slater, M., Pfurtscheller, G.: Self-paced (Asynchronous) BCI Control of a Wheelchair in Virtual Environments: a case study with a tetraplegic. In: Intell. Neuroscience, Hindawi Publishing Corp. (2007), 1-12.
5. Mikhail, M., El-Ayat, K., El Kaliouby, R., Coan, J., Allen, J.J.B.: Emotion Detection using Noisy EEG Data. In: Proc. of AH '10, ACM Press (2010), 1-7.
6. Ranky, G.N., Adamovich, S.: Analysis of a Commercial EEG Device to Control a Robot Arm. In: Proc. of the 2010 IEEE 36th Annual Northeast (2010), 1-2.
7. Scherer, R., Mueller, G.R., Neuper, C., Graimann, B., Pfurtscheller, G.: An Asynchronously Controlled EEG-based Virtual Keyboard: Improvement of the Spelling Rate. In: IEEE Transactions on Biomedical Engineering, Volume 51, Number 6 (June 2004), 979-984.
8. Swartz, B.E., Goldensohn, E.S.: Timeline of the History of EEG and Associated Fields. In: Electroencephalography and clinical Neurophysiology, Vol. 106, No. 2. (February 1998), 173-176.