View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Institute of Transport Research:Publications

Continuous affect state annotation using a joystick-based user interface

J. Antony[1], K. Sharma[1*], C. Castellini[1], Egon L. van den Broek[2], C. Borst[1]

[1]Robotics and Mechatronics Center, German Aerospace Center, Wessling, Germany

[2]Department of Information and Computing Sciences, Utrecht University, Utrecht, The Netherlands

[*] *karan.sharma@dlr.de*

Abstract

Ongoing research at the DLR (German Aerospace Center) aims to employ affective computing techniques to ascertain the emotional states of users in motion simulators. In this work, a novel user feedback interface employing a joystick to acquire subjective evaluation of the affective experience is presented. This interface allows the subjects to continuously annotate their affect states, elicited in this scenario by watching video clips. Several physiological parameters (e.g. heart rate, electrodermal activity, respiration rate, etc.) were acquired during the viewing session. A statistical analysis is presented, which shows expected patterns in data that validate the design and methodology of the experiment and lay the groundwork for further experiments to be undertaken at the DLR.

Introduction

The work presented here aims to ascertain the affective experience of a human subject in a motion simulator, namely the DLR-Robotic Motion Simulator (DLR-RMS) [1]. Russell defines *core affect* as *'the neurophysiological state that is consciously accessible as a simple, non-reflective feeling that is an integral blend of hedonic (pleasure-displeasure) and arousal (sleepy-activated) values'* [2]. The study presented here focuses on 4 representative and contrasting emotion classes that can presumably be elicited during the course of a simulation, namely *thrill/amusement, fear, boredom and calmness*. These emotion classes are characterised by the different *core affect* (i.e. valence-arousal levels) states associated with them [1]. With the medium-term goal of quantifying in real-time and in a continuous fashion the affective state of a subject during one of the aforementioned motion simulations, we designed an initial ground-based experiment in which video clips were administered to 30 human subjects; while viewing the clips, each subject was required to annotate his/her affective state using a joystick, and several kinds of biological signals were recorded. Some of the important design aspects of this experiment are explained in the following sections.

Use of Videos for Affect Elicitation

Video stimuli have been used in emotion studies for eliciting specific and intense emotional responses in a relatively short period of time [3,4]. Video stimuli are dynamic, encompassing visual and auditory cues which lead to greater involvement of the participant in the emotion experience [3]. Also, carefully selected video stimuli can be used for the elicitation of the intended affect states (both positive and negative) in sufficient intensity and duration without involving any harm or ethical violation [4]. Although an affective experience on a motion simulator also involves motion, the use of video stimuli offers a similar affective experience to a simulation and is therefore used in this initial investigation.

Use of Biosignals for determining the Affect State

During the course of the experiment, several physiological parameters of users were measured by means of different bio-sensors. These include: electrocardiogram (ECG) and the blood volume pulse (BVP) sensors to measure the cardiac activity, an electrodermal activity (EDA) sensor, a respiration belt sensor to acquire the respiration rate, a skin temperature sensor and three surface electromyography (sEMG) sensors to measure the

Proceedings of Measuring Behavior 2014, (Wageningen, The Netherlands, August 27-29, 2014).
Editors: A.J. Spink, L.W.S. Loijens, M. Woloszynowska-Fraser & L.P.J.J. Noldus. www.measuringbehavior.org

muscular activity in the *Corrugator Supercilii*, the *Zygomaticus Major* and the *Trapezius* muscles. These biosignals are indicative of the underlying affect states [5,6,7]. In the context of affect state estimation, biosignals have the desirable properties of being easily acquired, immune to voluntary manipulation or masking, generally independent of age and gender, etc. [7].

Annotation Methods

Traditional self-reporting/assessment techniques, allowing participants to state their affect/emotional states, are post-stimuli questionnaires such as, e.g. Likert scales [4,8]; however, most of these methods don't allow for continuous self-reporting and therefore the flow of the experiment has to be interrupted to allow the participant to report their experience [9]. To overcome this limitation, self-reporting HMIs using dials [9] or computer mice [9] have been used; along similar lines, we developed a joystick-based user interface (UI) that allows for continuous annotation of affect experiences by positioning a pointer in a 2-D space. The motivation behind using a 2-D annotation space is that according to Lang [9], "all emotions can be located in a 2-D space, as co-ordinates of affective valence and arousal". Therefore, by allowing the user to report their perceived valence and arousal levels instead of discrete emotional states, the interface allows the users to also report mixed emotional experiences. Another advantage that a joystick offers over other methods is that it is intuitive to use and enables the user to annotate his/her affect experience at the same time as they are elicited. This helps in reducing the measurement errors in emotion assessment arising from delays in reporting, fading or distortion of affect state experiences during recall and errors due to distraction of the participant from the emotion stimuli [4,8].

Setup Description: The Annotation Interface

Since this paper focuses on the annotation method used for the experiment, a detailed explanation of the experimental methodology and procedure is not presented here. In the following sub-sections, a concise overview of different aspects of the annotation method is presented.

Design of the UI: The interface design is in essence based on the 2-D valence-arousal affect model (see Figure 1) [12]. One difference is the addition of Self-Assessment Manikin (SAM) [12,13] to the co-ordinate axis of the interface. The manikin figures depict different valence (on X-axis) and arousal (on Y-axis) levels and are arranged on the co-ordinate axes in an ascending order of these levels. A red pointer, seen at the position (5,5) in Figure 1, represents the rest position of the joystick and is defined as the neutral affect position. The UI is displayed on the upper right corner of the LCD monitor (see Figure 2). This allows the user to view the video-clip and annotate his affect state simultaneously.

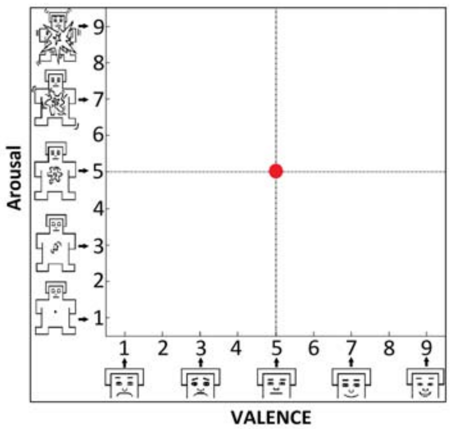

Figure 1. The annotation UI.

Figure 2. The annotation UI embedded in a video.

Annotation Procedure: While viewing a video-clip, the user undergoes an affective experience resulting in a change in his valence and arousal levels. The user annotates his perceived experience by positioning the red pointer in the appropriate region of the UI. When the elicited affect changes, the user is supposed to then position the pointer in the region that best characterizes his/her perceived affect state.

Training Procedure: A few days before the test, the subjects were provided a document with an overview of the annotation interface. On the day of the test, before commencing the session, they were again given a short re-introduction to the valence-arousal model and the annotation procedure. They were instructed not to annotate the affective content of the video, but rather provide ratings based on their perception of their affect state. After this, they were given a short demo of the UI and were allowed to get a feel of the system. This was followed by a practice session containing 5 video-clips of 1 minute duration. During this session, the users annotated their affect states while watching the videos and were allowed to ask questions (pertaining to the annotation process) at the end of each video-clip. After the training session, the main experiment was started.

Data Analysis and Results

Figure 3 shows the affect state annotation values of a typical subject for the entire duration of the test. The *x* and *y* co-ordinate values of joystick data (see Figure 1) provide the valence and arousal levels, respectively. The data are colour coded (and labelled) based on the emotion that the video clip was expected to elicit. The emotion label associated with each video was determined through an initial evaluation, undertaken by a different set of participants. Except the first two videos in the sequence (marked as ‘Calmness’ and ‘Neutral’ in Figure 3) and the last video (marked as ‘Neutral’ in the Figure 3), the order of all the other videos was randomly shuffled and every participant viewed a different sequence. Also, no two videos eliciting the same type of emotional response would ever be shown one after another, and each video was followed by a blue screen (duration: 2 minutes, marked at ‘Neutral’ in Figure 3) to isolate the impact of each video on the affect state of the participant. The total number of videos in a sequence was 18, i.e. 8 videos for emotion elicitation (2 for each emotion label), 9 blue screens and 1 ‘calming’ video at the start of the sequence. After the test, the users provided feedback on the usability of the annotation system by filling out a System Usability Scale (SUS) survey form. An analysis of the results of this survey is not presented here as it’s beyond the scope of this publication.

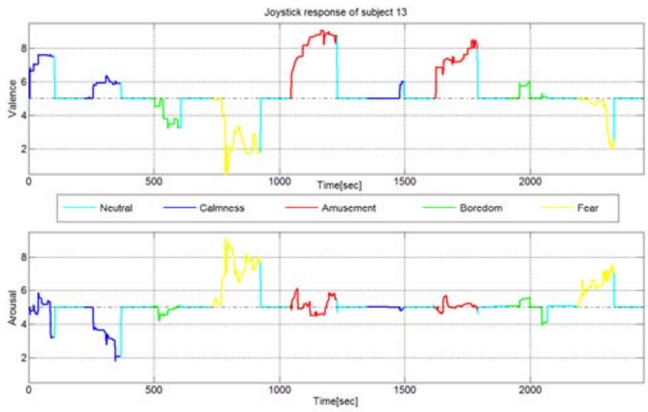

Figure 3. Annotation values for subject 13.

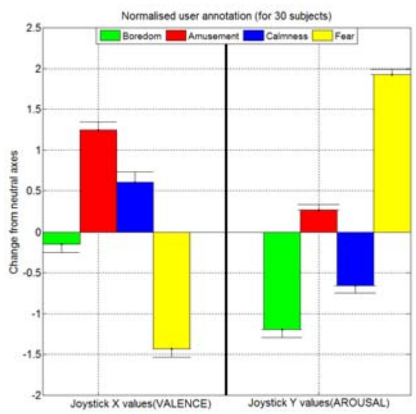

Figure 4. Normalised annotation values for all subjects.

A ‘scary’ video is expected to elicit ‘fear’ in the participants, which is characterized by high arousal and low valence in the valence-arousal model of affect. From the annotation data for the two horror clips (labelled as ‘fear’ and represented by the yellow segments in Figure 3 & 4) used in this test, it can be seen that the response of subject 13 to these videos is categorized by low valence and high arousal values. In the valence-arousal affect model, the other type of video segments eliciting boredom, amusement and calmness are classified by: low valence-low arousal, high valence-high arousal and intermediate valence-low arousal levels, respectively. The response of subject 13 (see Figure 3) is along the same lines.

Figure 4 shows the mean valence and arousal values and the standard error of means (SEM) for all 30 subjects, across different video labels; valence and arousal values from every participant were first averaged across every video clip, then these averaged values were normalised. From the normalised dataset, the mean and the SEM values for each video type, across all 30 subjects were calculated. The *x*-axis in this figure is equivalent to neutral (5,5) position in the UI (see Figure 1). Therefore, negative values in this figure signify low valence and arousal values (from 1 to 5) in the UI. From this plot, it can be observed, e.g. that ‘fear’ eliciting video clips are annotated as high arousal and low valence events by all users. Similarly, ‘calm’ videos are represented by low arousal and intermediate valence.

Conclusion

Using a joystick based annotation UI offers several advantages over conventional self-reporting methods. For example, through the use of a joystick based UI, the video clips are continuously annotated. Therefore, in the case of ‘fear’ eliciting videos, we can also explore the ‘build up’ phase for an affect state, which is of interest [cf. 14,15]. Secondly, any errors in labelling originating from temporal latency are also avoided. Based on the results presented above, we state that the proposed annotation system is a viable and favourable alternative to the conventional self-reporting techniques.

References

1. Bellmann, T., Heindl, J., Hellener, M., Kuchar, R., Sharma, K., Hirzinger, G. (2011). The DLR Robot Motion Simulator Part I: Design and Setup. In *Proceedings of the 2011 IEEE International Conference on Robotics and Automation (ICRA 2011)* (Shanghai, China, May 9-13), p. 4694 –4701.

2. Russell, J.A. (2003). Core affect and the psychological construction of emotion. *Psychological review **110**(1)***,** 145-172.

3. Gross, J.J., Levenson, R. W. (1995). Emotion elicitation using films. *Cognition and Emotion **9**(1)***,** 87-108.

4. Rottenberg, J., Ray, R.D., Gross, J.J. (2007). Emotion elicitation using films. In J.A. Coan, J.J.B. Allen (Eds.), *Handbook of emotion elicitation and assessment, Chapter 1*, p. 9-28. New York, NY, USA: Oxford University Press, Inc.

5. Picard, R.W. (1995). *Affective computing*. Technical Report 321, M.I.T Media Laboratory Perceptual Computing Section, Cambridge, MA, USA.

6. Liu, C., Rani, P., Sarkar, N. (2006). Human-robot interaction using affective cues. In *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN06)* (Hatfield, UK, September 6-8), p. 285–290.

7. Van den Broek, E.L. (2011). *Affective Signal Processing (ASP): Unravelling the mystery of emotions.* CTIT Ph.D.-thesis series No. 11-204, Human Media Interaction (HMI), Faculty of Electrical Engineering, Mathematics, and Computer Science, University of Twente, Enschede, The Netherlands.

8. Levenson, R.W. (1988). Emotion and the autonomic nervous system: a prospectus for research on autonomic specificity. In H.L. Wagner (Ed.), *Social psychophysiology and emotion: Theory and clinical applications, Chapter 2*, 17-42. Chichester, West Sussex, Great Britain: John Wiley & Sons Ltd.

9. Ruef, A.M. and Levenson, R.W. (2007). Continuous measurement of emotion: The affect rating dial. In J.A. Coan, J.J.B. Allen (Eds.), *Handbook of emotion elicitation and assessment, Chapter 17*, p. 286-297. New York, NY, USA: Oxford University Press, Inc.

10. Nagel, F., Kopiez, R., Grewe, O., Altenmüller, E. (2007). Emujoy: Software for continuous measurement of perceived emotions in music. *Behavior Research Methods, **39**(2)*, 283-290.

11. Lang, P.J. (1995). The emotion probe: Studies of motivation and attention. *American Psychologist, **50**(5)***,** 372-385.

12. Van den Broek, E.L. (2013). Ubiquitous emotion-aware computing. *Personal and Ubiquitous Computing, **17**(1)*, 53-67.

13. Bradley, M.M. and Lang, P.J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, **25**(1)***,** 49-59.

14. Van den Broek, E.L., Westerink, J.H.D.M. (2009). Considerations for emotion-aware consumer products. *Applied Ergonomics, **40**(6)*, 1055-1064.

15. Van den Broek, E.L., Lisý, V., Janssen, J.H., Westerink, J.H.D.M. , Schut, M.H., Tuinenbreijer, K. (2010). Affective Man-Machine Interface: Unveiling human emotions through biosignals. In A. Fred, J. Filipe, H. Gamboa (Eds.), *Biomedical Engineering Systems and Technologies,* p. 21-47., Berlin/Heidelberg, Germany: Springer-Verlag.