A VR Game-based System for Multimodal Emotion Data Collection

Chiara Bassano, Giorgio Ballestin, Eleonora Ceccaldi, Fanny Isabelle Larradet, Maurizio Mancini, Erica Volta, Radoslaw Niewiadomski*

ABSTRACT

The rising popularity of learning techniques in data analysis has recently led to an increased need of large-scale datasets. In this study, we propose a system consisting of a VR game and a software platform designed to collect the player's multimodal data, synchronized with the VR content, with the aim of creating a dataset for emotion detection and recognition. The game was implemented ad-hoc in order to elicit joy and frustration, following the emotion elicitation process described by Roseman's appraisal theory. In this preliminary study, 5 participants played our VR game along with pre-existing ones and self-reported experienced emotions.

CCS CONCEPTS

• **Human-centered computing** → *Human computer interaction (HCI)*; *Virtual reality*.

KEYWORDS

VR interaction, emotion elicitation, appraisal theory, mutimodal

ACM Reference Format:

Chiara Bassano, Giorgio Ballestin, Eleonora Ceccaldi, Fanny Isabelle Larradet, Maurizio Mancini, Erica Volta, Radoslaw Niewiadomski. 2018. A VR Game-based System for Multimodal Emotion Data Collection. In *Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY*. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

One of the greatest challenges in Affective Computing is the creation of ecological multimodal datasets for emotion detection and recognition. Ideally, such datasets would contain affective expressions recorded in real-life settings. Unfortunately, this poses many challenges, such as maintaining synchronization across different data streams. Nevertheless, the attempt to create emotional states in the laboratory for scientific aims has a longstanding tradition [1]. The attention towards video games [2] and VR games [3] is increasing in affective computing. In fact, not only they provide controlled, replicable environments, along with the possibility of manipulating stimuli for emotion elicitation, but also they allow the assessment of emotions by collecting expressive behaviors and physiological measurements [4].

In this paper, we present an approach consisting of a VR game that aims at eliciting different emotions, and a system for recording synchronized multimodal data streams. The players' task is to assemble an exoskeleton suit within a limited amount of time. Features of the game-play were based on the appraisal variables that, according to the appraisal theory of emotions proposed by Ira Roseman [5] would elicit joy and frustration. In the VR game background, our system is able to record physiological (heart rate, muscle contraction), kinematic (acceleration), visual (video of the user and the seen VR environment) and auditory (respiration) data. As previous works demonstrated [6, 7] that all these modalities can be used (separately) to detect affective states, our system collects all these information, and it does it in a synchronous manner.

2 PROPOSED SYSTEM

2.1 Game flow

Our VR game is a single user, based on the manipulation and assembly of virtual objects to be performed within a limited amount of time (Fig. 1). We intentionally avoided cognitively demanding tasks and elaborate narratives [3]. The HTC Vive headset is used for visualization, while interaction is made simple and intuitive through the use of HTC Vive controllers.

After a demo scene used to familiarize with the game interaction modality and the task, the actual game starts. There are two playing conditions, set by the experimenter: normal and manipulated. In both conditions a timer is shown on the top left part of the VR display, at the beginning and during the last 10 seconds of the game. In the normal condition, participants have up to 2, 3 or 4 minutes to complete the suit, according to the experimenter choice. In the manipulated condition, whenever the player is about to complete the suit (i.e., when 11 of the 13 suit parts have been positioned), the countdown drops down to 10 seconds, making it nearly impossible to successfully complete the game.

According to Roseman's theory, the *situational state* appraisal variable permits to distinguish joy and frustration. More precisely, *circumstance-caused* events *inconsistent* with personal motives are

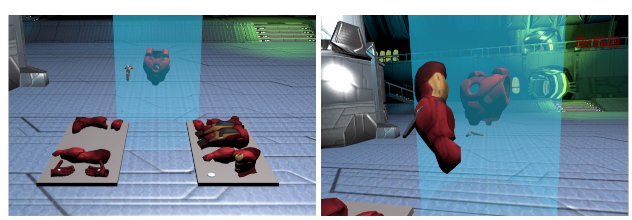

Figure 1: Two screenshots of the VR game.

*From 2nd to the 6th author are in the alphabetical order.

Woodstock '18, June 03–05, 2018, Woodstock, NY

ACM ISBN 978-1-4503-9999-9/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

expected to elicit frustration, whereas *circumstance-caused* events *consistent* with ones' motives supposedly elicit joy. In our case, frustration is expected when there is not enough time to complete the task, while joy - when the player is able to assemble the exoskeleton.

2.2 Multimodal recording system

We implemented the multimodal recording system by exploiting the EyesWeb XMI open research platform[1]. Several machines are connected through a wired network, on which a SMPTE timecode signal is constantly transmitted via UDP packets and acts as a synchronization clock between the machines (Fig. 2). Each machine has an internal clock that is used to generate timestamps for the recorded data. The wired network is a local gigabit Ethernet connection which ensures a high speed transmission of the UDP packets containing the SMPTE timecodes.

To our purpose we integrated 5 devices in our setup (Fig. 2): a head mounted wireless microphone to register the audio of respiration; the HTC Vive, providing its own tracking system; the MYO device[2], capturing forearms EMG signal; Empatica E4[3] bracelet, used to collect physiological signals and a webcam camera recording player movements. Also, the second video stream is recorded in EyesWeb XMI by capturing a portion of the screen of the machine running SteamVR[4]. It is worth noting, however, that the innate modularity of our system allows for adding more devices without impairing the recording system capabilities.

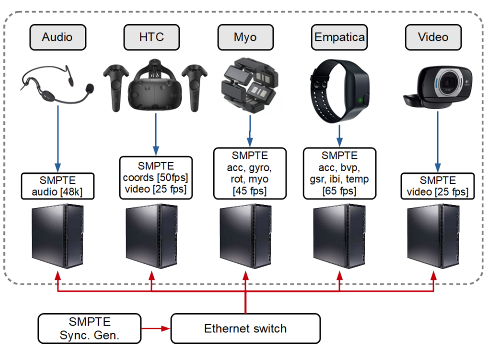

Figure 2: Sensors and data streams of the recording system.

3 PILOT STUDY AND CONCLUSION

A preliminary data collection was carried out on five participants (3 females). Five tasks we leveraged for eliciting both positive and negative emotions: (T1) Kitty Rescue game; (T2) a set of videos associated to amusement, no emotions and awe [8, 9]; (T3) our suit assembling game; (T4) Shinrin-yoku: Forest Meditation and Relaxation; (T5) RideOp - VR Thrill Ride Experience. The Tasks T1-T2, and T4-T5 were chosen to elicit other than joy and frustration positive and negative emotions (e.g., fear and awe). All participants had to perform the five tasks in a fixed order. After each task, they were asked to self-report their affective state by selecting one or more emotions from a 16 labels list. According to the self-reports, awe was experienced 8 times (twice in T2, 3 times in T4 and T5), surprise 6 times (once in T1, T3, and T4, 3 times in T2), joy 11 times (once in T1, T3 and T5, 6 times in T2, twice in T4), relief 6 times (twice in T1 and T4, once in T2 and T5), fear 11 times (5 times in T1, twice in T2, 4 times in T5), frustration 5 times (twice in T1, and 3 times in T3), anger twice in T2, sadness 2 times (in T1 and T2). Three emotions were reported once: hope (in T2), distress (in T1), pride (in T3). No emotion was reported 7 times (6 times in T2, and once in T3), other emotion (in this case disgust) was reported once in T2. Two labels: guilt and regret were never chosen by participants.

To sum up, participants' self-reports indicate that different positive and negative emotions were elicited during Tasks T1-T5, suggesting the feasibility of using VR-based methods to collect affect-related data. In particular, frustration during T3 was successfully elicited in 3 out of 5 participants. It is worth noting, however, that, according to appraisal theories, the same event can result in different emotions being elicited, depending on the person's appraisal. Interestingly, one of the participants reported in T3 other than joy positive emotion, i.e., pride. In the Roseman's model joy is differentiated from pride by *agency* appraisal. The first is caused by *circumstance-caused* events, and the second by *self-caused* events. Therefore, this model gives the plausible explanation of pride emotion, as it is reasonable to assign the agency to oneself when wining this type of a game.

In the future, the assessment of the emotional states could be enriched by measuring how players evaluate their experience in terms of appraisal variables, instead of only asking them for emotional labels. Moreover, physiological, kinematic, visual and auditory data will be analysed in order to understand their correlations with emotional states changes. Finally, we aim to develop more custom-made VR-games to elicit a wider spectrum of emotions.

REFERENCES

[1] M. K. Uhrig, N. Trautmann, U. Baumgärtner, R.-D. Treede, F. Henrich, W. Hiller, and S. Marschall, "Emotion elicitation: A comparison of pictures and films," *Frontiers in psychology*, vol. 7, p. 180, 2016.
[2] C. van Reekum, T. Johnstone, R. Banse, A. Etter, T. Wehrle, and K. Scherer, "Psychophysiological responses to appraisal dimensions in a computer game," *Cognition and Emotion*, vol. 18, no. 5, pp. 663–688, 2004.
[3] B. Meuleman and D. Rudrauf, "Induction and profiling of strong multi-componential emotions in virtual reality," *IEEE Transactions on Affective Computing*, pp. 1–1, 2018.
[4] B. Cebeci, U. Celikcan, and T. K. Capin, "A comprehensive study of the affective and physiological responses induced by dynamic virtual reality environments," *Computer Animation and Virtual Worlds*, vol. 0, no. 0, p. e1893, e1893 cav.1893.
[5] I. Roseman and A. Evdokas, "Appraisals cause experienced emotions: Experimental evidence," *Cognition and Emotion*, vol. 18, no. 1, pp. 1–28, 2004.
[6] G. Castellano, S. D. Villalba, and A. Camurri, "Recognising human emotions from body movement and gesture dynamics," in *Affective computing and intelligent interaction*. Springer, 2007, pp. 71–82.
[7] V. Lussu, R. Niewiadomski, G. Volpe, and A. Camurri, "The role of respiration audio in multimodal analysis of movement qualities," *Journal on Multimodal User Interfaces*, Apr 2019.
[8] A. Chirico, D. B. Yaden, G. Riva, and A. Gaggioli, "The potential of virtual reality for the investigation of awe," *Frontiers in Psychology*, vol. 7, p. 1766, 2016.
[9] A. Chirico, P. Cipresso, G. Riva, and A. Gaggioli, "A process for selecting and validating awe-inducing audio-visual stimuli," in *Pervasive Computing Paradigms for Mental Health*, N. Oliver, S. Serino, A. Matic, P. Cipresso, N. Filipovic, and L. Gavrilovska, Eds. Cham: Springer International Publishing, 2018, pp. 19–27.

[1] http://www.infomus.org/eyesweb_eng.php
[2] https://support.getmyo.com/hc/en-us
[3] https://www.empatica.com/en-eu/research/e4/
[4] https://www.steamvr.com