The EmojiGrid as an Immersive Self-report Tool for the Affective Assessment of 360 VR Videos

Alexander Toet[1](✉), Fabienne Heijn[1,2], Anne-Marie Brouwer[1], Tina Mioch[1], and Jan B. F. van Erp[1,3]

[1] TNO Human Factors, Soesterberg, The Netherlands
{lex.toet,anne-marie.brouwer,tina.mioch,jan.vanerp}@tno.nl
[2] University of Utrecht, Utrecht, The Netherlands
[3] University of Twente, Enschede, The Netherlands

Abstract. Immersive 360° VR systems are increasingly used in entertainment, marketing and design and development processes. Many of these applications involve emotional experiences. Since momentary emotions significantly determine a user's response and decisions, it is essential to understand the influence of media content and technological factors on the user's emotional response. To measure the emotional responses to immersive VR experiences, efficient and validated instruments are required. Most tools currently used to measure the emotional response of users compromise the immersive experience since they are cognitively demanding, time consuming, and their application requires the user to leave the VR. We investigated the validity of an immersive, efficient and intuitive EmojiGrid graphical self-report tool for the assessment of emotions evoked by 360º VR videos. Using the EmojiGrid, 40 participants rated 62 360º VR videos from a validated public database. The resulting mean valence and arousal values agree with the corresponding values provided with the database (obtained with an alternative validated rating tool). We conclude that the EmojiGrid is a valid self-report tool for the assessment of VR-evoked emotions.

Keywords: 360º VR · Valence · Arousal · Emotions · EmojiGrid

1 Introduction

1.1 Immersive VR as an Affective Medium

Immersive VR systems presenting the user a full 360º head rotation view can effectively induce a wide range of emotional responses in individuals [1, 2], comparable to those evoked by real-life scenarios [3]. The availability of low-cost head-mounted displays (HMDs) has increased the popularity of immersive 360º VR videos on video streaming platforms such as YouTube and Facebook and stimulated applications in many other areas where emotions play a role, such as online retail for studying and influencing consumer behavior and product evaluation [4, 5], the architectural design [6], and

P. Bourdot et al. (Eds.): EuroVR 2019, LNCS 11883, pp. 330–335, 2019.
https://doi.org/10.1007/978-3-030-31908-3_24

product design [7]. The news industry has adopted immersive 360º videos to establish an emotional connection between subject and viewer [8, 9]. The tourism industry has embraced immersive 360º VR systems to provide prospective travelers a compelling sneak-preview of their holiday experience [10]. Also, systems providing shared social experiences in VR between remote participants are becoming available, such as VR teleconferencing and VR social visits.

The increasing mass consumption of VR experiences in entertainment and many other fields, in combination with the lack of knowledge on the emotional effects of VR, calls for research in the area of emotions elicited by VR content [2]. Although it is generally assumed that immersive VR experiences elicit similar emotional responses as real-life experiences, a direct comparison of similar emotional experiences in both worlds is still lacking [11]. Also, it is still not clear how the emotional responses of users relate to the content of VR media [1]. To investigate the influence of technological factors on emotions and presence in VR, there is a need for further studies [12], specifically those comparing the emotional effects elicited by in vivo exposure to those evoked by VR exposure [1]. To enable these studies, immersive, efficient and validated instruments are needed to measure the emotional responses to VR experiences [2].

1.2 Measuring Emotions Evoked by Immersive VR

Most tools that are currently used for the subjective assessment of emotional responses to 360º VR systems are time consuming, demand cognitive effort (interpretation) and are typically applied outside the VR. They are therefore typically applied after a VR experience has ended [1–3]. As a result, it is not clear to what extent they reflect the momentarily experienced emotions during the VR presentation. Hence, there is a need for efficient affective self-report tools that can efficiently and reliably measure VR evoked emotions, even during the VR experience itself.

The EmojiGrid [13] is a graphical self-report tool that enables users to rate their subjectively experienced valence and arousal with a single response. The EmojiGrid is a Cartesian grid that is labeled with emoji showing different emotional facial expressions (Fig. 1). Users can report their subjective ratings of valence and arousal by marking the location on the grid that best represents their current emotional state. The tool is easy to administer, intuitive (the facial expressions speak for themselves and don't need additional labels) and efficient (the two principal affective dimensions are measured with a single response). This suggests that embedding the EmojiGrid in a VR could afford the immersive assessment of a user's emotional response to a VR experience in a minimally intrusive way.

1.3 Current Study

This study was performed to evaluate the EmojiGrid as a self-report tool for the assessment of immersive VR evoked emotions. Thereto, we used the EmojiGrid to measure valence and arousal for immersive 360 VR video clips from the public database provided by Li et al. [14] and we compared the results with the corresponding normative ratings provided with this database, that were obtained with the validated Self-Assessment Mannikin (SAM: [15]).

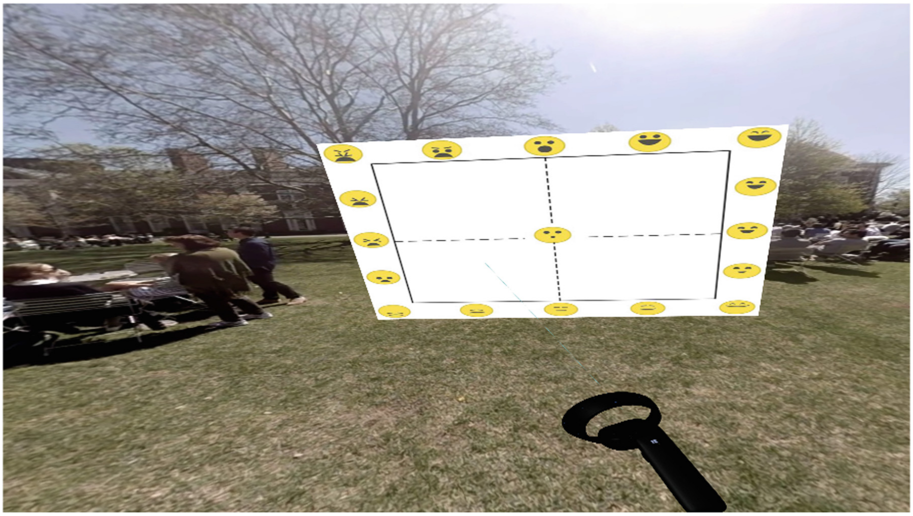

Fig. 1. Screenshots showing the virtual representation of the Samsung Odyssey Controller (bottom right) pointing a laser beam at the EmojiGrid that is projected over the VR scene.

2 Methods and Procedure

The stimuli used in this study were 62 immersive 360º VR videos from a publicly available validated database. All videos are of short length, require no explanation, and were found to induce different levels of valence and arousal [14]. For each video, mean valence and arousal ratings are provided as measured with the Self-Assessment Mannikin (SAM: [15]). The 62 immersive VR videos were divided into clusters with an approximate duration of 12 min per cluster, to prevent nausea or fatigue. This resulted in a total of 16 clusters, each consisting of two to six videos. The order of positive and negative valenced videos in the subsets was randomized. The presentation order of the clusters themselves was randomized over the participants. Three additional 360º VR videos (with respectively low, moderate and high valence) were downloaded from the internet and served to introduce the participants to the viewing and rating procedure. All videos were presented on a Samsung Odyssey Windows Mixed Reality headset (www.samsung.com), equipped with a Dual 3.5” AMOLED 1440 × 1600 resolution display, a 110º field of view and a refresh rate of 90 Hz. The video soundtracks were presented through a Sony MDR-1000X noise-canceling headset to prevent any distractions by ambient sounds.

The A-Frame (https://aframe.io) open source Javascript framework for creating (web-based) VR experiences was used to embed and display the EmojiGrid in the virtual environment at the end of each video. Node.js (https://nodejs.org) was used to set up a local server on an Alienware 13 R3 Notebook (Intel Core i7 7700HQ) which ran on Windows 10. To rate the videos, participants used a Samsung Odyssey remote control to point a virtual laser beam and place a check mark at the appropriate location on the EmojiGrid. The coordinates of the check marks (observer responses) on the EmojiGrid were also logged using A-Frame. These coordinates were rounded to two decimal places.

A total of 40 participants (22 females, 18 males) aged between 18 and 29 (M = 22.16; SD = 2.70) participated in this study. Before starting the actual experiment, they first inspected the EmojiGrid (Fig. 1) and were informed they could use this tool to report their response to a video by clicking on a point in the grid that best matched their emotions. No reference was made to the concepts of valence and arousal (the constructs underlying the axes of the EmojiGrid) since we wanted the participants to use the tool intuitively.

In the actual experiment, each participant was shown three clusters of immersive VR videos. At the end of each video, participants reported their emotions by pointing the virtual laser beam at the EmojiGrid. After clicking, the response (i.e., the coordinates of the check mark on the EmojiGrid) was stored, and the next video started playing. On average, each video was viewed and rated by a minimum of 7 participants. Between two clusters, the participants were given a five-minute break. During the breaks, participants could take off the HMD. After rating the third cluster, the HMD was removed, and the participants were debriefed about the purpose of the study. The entire experiment lasted about 80 min, including the debriefing.

3 Results

Figure 2 shows the relation between the mean (across all participants) valence and arousal ratings obtained with the EmojiGrid in this study and those obtained with the SAM in the earlier study by Li et al. [14]. This figure shows that the immersive VR videos successfully elicited a wide range of different emotions. To quantify the agreement between both results, we used IBM SPSS Statistics 25 (www.ibm.com) to compute the Intraclass Correlation Coefficient (ICC) estimates and their 95% confidence intervals, based on a mean-rating, consistency, two-way mixed-effects model. The ICC for valence was 0.91 [0.85–0.95] and the ICC for arousal was 0.73 [0.55–0.84]. Thus, the valence ratings obtained by both studies are in excellent agreement, whereas the arousal ratings moderately agree.

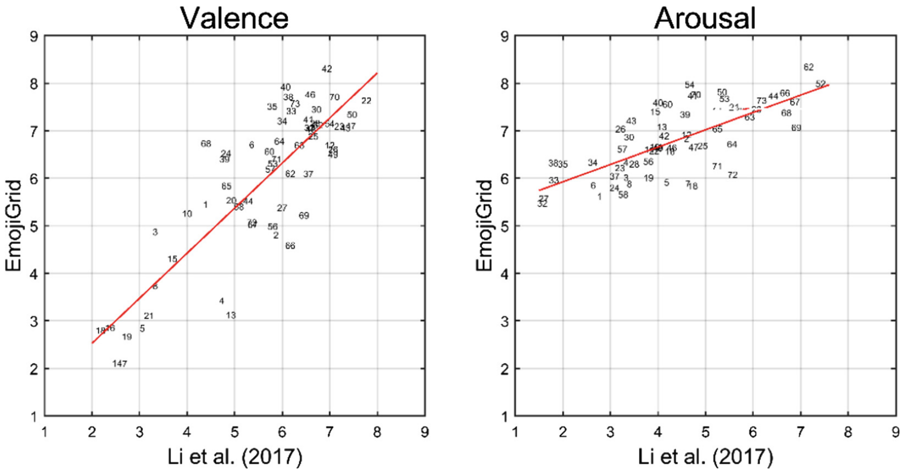

Fig. 2. Correlation plots illustrating the relationship between the valence (left) and arousal (right) ratings provided by Li et al. [14] and those obtained with the EmojiGrid in the current study. The numbers correspond to the original video identifiers in the database from Li et al. [14].

4 Conclusion and Discussion

We investigated the validity and effectiveness of the EmojiGrid for rating the emotional response elicited by immersive VR stimuli. We found that participants intuitively reported their experienced emotions (valence and arousal) after watching an immersive VR video by simply pointing at the EmojiGrid in the VR space. The EmojiGrid itself required no further explanation. The validity of the EmojiGrid for measuring the valence and arousal of immersive VR videos was assessed by comparing the mean (over all participants) subjective valence and arousal ratings for all videos to the corresponding normative ratings provided by [14] and obtained with the validated SAM rating tool. The mean valence ratings obtained in both studies are in excellent agreement, while the arousal ratings moderately agree. We therefore conclude that the EmojiGrid can serve as an efficient, immersive and intuitive affective self-report tool for the assessment of VR-evoked emotions. The EmojiGrid can be implemented in the VR itself (eliminating the need to leave the VR and thereby give up immersion), is more efficient (requires only a single response to rate both valence and arousal) than most existing methods and requires no cognitive effort (is intuitive and language independent).

During the debriefings several participants remarked they found it difficult to give single overall affective rating to videos that contained both pleasant and unpleasant episodes. Future studies should therefore afford users to continuously report perceived affect (e.g., by moving a pointer-controlled beam over the support of the grid). This feature may also be useful for the affective annotation of multimedia, for personalized affective multimedia retrieval or multimedia recommender systems, for real-time affective appraisal of multimedia entertainment or as an affective input (feedback) tool for serious gaming applications and affective multimedia (e.g., music) generation.

References

1. Riva, G., Mantovani, F., Capideville, C.S., et al.: Affective interactions using virtual reality: the link between presence and emotions. CyberPsychol. Behav. **10**(1), 45–56 (2007)
2. Oliveira, T., Noriega, P., Rebelo, F., Heidrich, R.: Evaluation of the relationship between virtual environments and emotions. In: Rebelo, F., Soares, M. (eds.) AHFE 2017. AISC, vol. 588, pp. 71–82. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-60582-1_8
3. Chirico, A., Gaggioli, A.: When virtual feels real: comparing emotional responses and presence in virtual and natural environments. Cyberpsychol. Behav. Soc. Netw. **22**(3), 220–226 (2019)
4. Sinesio, F., Moneta, E., Porcherot, C., et al.: Do immersive techniques help to capture consumer reality? Food Qual. Prefer. **77**, 123–134 (2019)
5. Bonetti, F., Warnaby, G., Quinn, L.: Augmented reality and virtual reality in physical and online retailing: a review, synthesis and research agenda. In: Jung, T., tom Dieck, M.C. (eds.) Augmented Reality and Virtual Reality. PI, pp. 119–132. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-64027-3_9
6. Portman, M.E., Natapov, A., Fisher-Gewirtzman, D.: To go where no man has gone before: virtual reality in architecture, landscape architecture and environmental planning. Comput. Environ. Urban Syst. **54**, 376–384 (2015)
7. Hilfert, T., König, M.: Low-cost virtual reality environment for engineering and construction. Vis. Eng. **4**(1), 2 (2016)
8. Vettehen, P.H., Wiltink, D., Huiskamp, M., et al.: Taking the full view: how viewers respond to 360-degree video news. Comput. Hum. Behav. **91**, 24–32 (2019)
9. Wang, G., Gu, W., Suh, A.: The effects of 360-degree VR videos on audience engagement: evidence from the New York Times. In: Nah, F.F.-H., Xiao, B.S. (eds.) HCIBGO 2018. LNCS, vol. 10923, pp. 217–235. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-91716-0_17
10. Marasco, A., Buonincontri, P., van Niekerk, M., et al.: Exploring the role of next-generation virtual technologies in destination marketing. J. Destin. Mark. Manag. **9**, 138–148 (2018)
11. Diemer, J., Alpers, G.W., Peperkorn, H.M., et al.: The impact of perception and presence on emotional reactions: a review of research in virtual reality. Front. Psychol. **6**(26), 1–9 (2015)
12. Seth, A., Suzuki, K., Critchley, H.: An interoceptive predictive coding model of conscious presence. Front. Psychol. **2**, 395 (2012)
13. Toet, A., Kaneko, D., Ushiama, S., et al.: EmojiGrid: a 2D pictorial scale for the assessment of food elicited emotions. Front. Psychol. **9**, 2396 (2018)
14. Li, B.J., Bailenson, J.N., Pines, A., et al.: A public database of immersive VR videos with corresponding ratings of arousal, valence, and correlations between head movements and self report measures. Front. Psychol. **8**, 2116 (2017)
15. Lang, P.J.: Behavioral treatment and bio-behavioral assessment: computer applications. In: Sidowski, J.B., Johnson, J.H., Williams, T.A. (eds.) Technology in Mental Health Care Delivery Systems, pp. 119–137. Ablex Publishing Corporation, Norwood, USA (1980)