Real-time, Continuous Emotion Annotation Prototype for 360° VR Videos

Tong Xue
Beijing Institute of Technology
Centrum Wiskunde
& Informatica (CWI)
Amsterdam, The Netherlands
xuetong@bit.edu.cn

Abdallah El Ali
Centrum Wiskunde
& Informatica (CWI)
Amsterdam, The Netherlands
abdallah.el.ali@cwi.nl

Surjya Ghosh
Centrum Wiskunde
& Informatica (CWI)
Amsterdam, The Netherlands
surjya.ghosh@cwi.nl

Pablo Cesar
Centrum Wiskunde
& Informatica (CWI)
Delft University of Technology
Amsterdam, The Netherlands
p.s.cesar@cwi.nl

Gangyi Ding
Beijing Institute of Technology
Beijing, China
dgy@bit.edu.cn

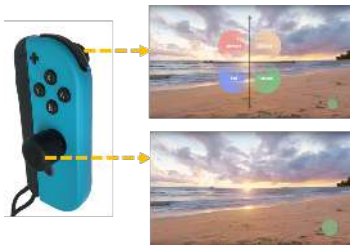
CHI'20,, April 25–30, 2020, Honolulu, HI, USA
ACM 978-1-4503-6819-3/20/04.
https://doi.org/10.1145/3334480.XXXXXXX

Abstract
With the increasing availability of head-mounted displays (HMDs) that show immersive 360° VR content, it is important to understand to what extent these immersive experiences can evoke emotions. Typically to collect emotion ground truth labels, users rate videos through post-experience self-reports that are discrete in nature. However, post-stimuli self-reports are temporally imprecise, especially after watching 360° videos. In this work, we developed a novel joystick-based emotion annotation system, in which users could report their valence and arousal continuously while watching 360° videos by HTC Vive Pro Eye HMD. Two suitable peripheral visualization techniques (HaloLight and DotSize) were designed aiming at minimizing workload and distraction. Our prototype provides a starting point for researchers to design momentary self-reports in virtual environments for precise emotion recognition.

Author Keywords
360° video; emotion annotation; continuous; visualization

CCS Concepts
•Human-centered computing → Human computer interaction (HCI); *Virtual Reality;*

(a) HTC VIVE Pro Eye HMD for displaying stimuli



(b) Joy-con Right Controller (Joystick for rating emotion and Trigger Button for activating helper function).

Figure 1: 360° VR video emotion annotation components.

Introduction

Immersive Virtual Reality (VR) experiences, such as watching 360° videos using head-mounted displays (HMDs), allow users to interact with content and feel immersed in the experience. An important aspect of this immersion is the capacity of content to evoke a wide range of emotions in individuals as they interact through head movements in this virtual space [10]. Whether the end goal is to evoke emotional responses and positive associations for educational purposes [1], for news engagement [14, 15], or improve tourism experiences through immersive previews [3], it is important to collect accurate and precise ground truth labels throughout the user's immersive experience.

Based on considerations for designing VR HMD-based interactions [6], we scope our work to HTC VIVE Pro Eye[1], and develop an input technique based on the Joy-Con controller[2]. For output, we draw on peripheral information visualization research, and follow a user-centric approach to design several alternatives for how to display user state feedback continuously while users are watching a 360° video.

Designing 360° Video Continuous Emotion Annotation Prototype

Our 360° video continuous annotation prototype is shown in Figure 1. It consists of two major components: (1) the HTC VIVE Pro Eye HMD (Figure 1(a)) with a resolution of 2880 x 1600 pixels, a 110° field of view and a refresh rate of 90Hz (2) the input device (a joystick on the Joy-con right controller) for emotion annotation (Figure 1(b)). A custom scene was constructed in the Unity Engine[3] to display 360° videos online at 30 fps and show the annotation feedback based on users' continuous ratings.

[1] https://enterprise.vive.com/us/product/vive-pro-eye/
[2] https://www.nintendo.com/switch/choose-your-joy-con-color/
[3] https://unity3d.com/

360° Video Stimuli

In this work, we select eight short 360° video clips with arousal and valence ratings from the database provided by Li et al [8], which contains arousal and valence ratings from 95 subjects. We selected two sample videos to represent each quadrant, ensuring the videos are matched as much as possible by duration. These videos are used as means to design our annotation tool.

Annotating Videos Continuously

Since users will be equipped with an HMD and cannot see the annotation tool, we need to ensure that annotating is comfortable,ergonomic, as well as precise. With a return spring, a physical joystick that provides additional proprioceptive feedback could aid realigning to center position under no force [12], which makes it suitable for continuous annotation while wearing an HMD. Based on Russell's Circumplex model[11], users can move the Joy-con joystick head into one of the four quadrants of the A-V model, as shown in Figure 3. To increase the emotion intensity, the user can move the joystick head further.

Visual Annotation Feedback

Since users will be watching 360° videos while annotating emotions continuously, this will lead to divided attention [7]. To this end, it is necessary to reduce the burden of annotating and conveying state feedback without interfering too much with the viewing experience. We draw on research on peripheral feedback [2], and consider GUI elements to lower interruptions which helps users keep awareness of the primary task [9, 4].

A. Visual Annotation Feedback

We designed HaloLight and DotSize two peripheral feedback prototypes (Figure 2). Whereas HaloLight uses color opacity to indicate intensity, DotSize uses the size of the filled circle to indicate intensity. Both are presented in the

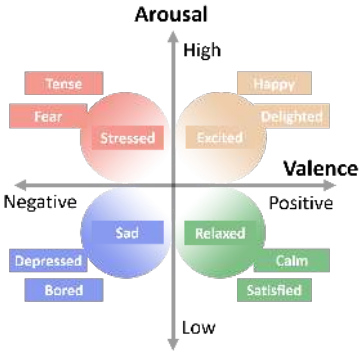


Figure 3: Arousal-Valence model space based on Russell's Circumplex model [11]. In our annotation prototype, four distinct colors are selected across quadrants (HEX values = #eecdac, #7fc087, #879af0, #f4978e for quadrants one to four clock-wise, respectively) [5].

(a) HaloLight: shaded halo arc in bottom-right viewport, which varies in transparency with emotion intensity.

(b) DotSize: circle dot in bottom-right viewport, which varies in size with emotion intensity.

Figure 2: Two peripheral feedback designs.

periphery of users' visual attention, where their position is fixed to the bottom right corner of the HMD viewport. The color of the component indicates the emotion users were annotating currently [5]. Four colors (HEX values = #eecdac, #7fc087, #879af0, #f4978e for quadrants one to four respectively) provided feedback to users on which emotion they were currently annotating. We selected four colors based on a simplified version of Itten's color system [13], which has been shown to be intuitive and easy for users to understand [5].

B. On-demand reference

We also developed an on-demand helper function, so that users who forget what color corresponds to a quadrant with corresponding emotions can use it for easy lookup. We enabled this on-demand reference functionality, activated through a joystick button press event. We show the helper function in Figure 1(b), where here we only include the most representative emotion keyword (by contrast to several keywords in Figure 3), that should serve as a reminder trigger for what a quadrant corresponds to.

In conclusion, our prototype provides a continuous emotion annotation method that is suitable for collecting more precise emotion ground truth labels for 360° VR video content. This demo submission can help both elicit and capture momentary emotions in virtual environments, which can be used for building more temporally precise 360° video-based emotion recognition and prediction models.

REFERENCES

[1] Fathima Assilmia, Yun Suen Pai, Keiko Okawa, and Kai Kunze. 2017. IN360: A 360-Degree-Video Platform to Change Students Preconceived Notions on Their Career. In *Proc. CHI EA '17*. ACM, New York, NY, USA, 2359–2365.

[2] Saskia Bakker, Doris Hausen, and Ted Selker. 2016. *Peripheral Interaction: Challenges and Opportunities for HCI in the Periphery of Attention*. Springer.

[3] Julia Beck and Roman Egger. 2018. Emotionalise Me: Self-reporting and Arousal Measurements in Virtual Tourism Environments. In *Information and Communication Technologies in Tourism 2018*, Brigitte Stangl and Juho Pesonen (Eds.). Springer International Publishing, Cham, 3–15.

[4] Carl Gutwin, Andy Cockburn, and Ashley Coveney. 2017. Peripheral popout: The influence of visual angle and stimulus intensity on popout effects. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 208–219.

[5] Dini Handayani, Abdul Wahab, and Hamwira Yaacob. 2015. Recognition of emotions in video clips: the self-assessment manikin validation. *Telkomnika* 13, 4 (2015), 1343.

[6] Jason Jerald. 2015. *The VR book: Human-centered design for virtual reality*. Morgan & Claypool.

[7] Daniel Kahneman. 1973. *Attention and effort*. Vol. 1063. Citeseer.

[8] Benjamin J Li, Jeremy N Bailenson, Adam Pines, Walter J Greenleaf, and Leanne M Williams. 2017. A public database of immersive VR videos with corresponding ratings of arousal, valence, and correlations between head movements and self report measures. *Frontiers in psychology* 8 (2017), 2116.

[9] Aristides Mairena, Carl Gutwin, and Andy Cockburn. 2019. Peripheral notifications in large displays: Effects of feature combination and task interference. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

[10] Tiago Oliveira, Paulo Noriega, Francisco Rebelo, and Regina Heidrich. 2018. Evaluation of the Relationship Between Virtual Environments and Emotions. In *Advances in Ergonomics in Design*, Francisco Rebelo and Marcelo Soares (Eds.). Springer, Cham, 71–82.

[11] James A Russell. 1980. A circumplex model of affect. *Journal of personality and social psychology* 39, 6 (1980), 1161.

[12] Karan Sharma, Claudio Castellini, Freek Stulp, and Egon L Van den Broek. 2017. Continuous, real-time emotion annotation: A novel joystick-based analysis framework. *IEEE Trans. Affective Computing* (2017).

[13] Anna Ståhl, Petra Sundström, and Kristina Höök. 2005. A foundation for emotional expressivity. In *Proc. Designing for User eXperience '05*. AIGA: American Institute of Graphic Arts, 33.

[14] Paul Hendriks Vettehen, Daan Wiltink, Maite Huiskamp, Gabi Schaap, and Paul Ketelaar. 2019. Taking the full view: How viewers respond to 360-degree video news. *Computers in Human Behavior* 91 (2019), 24 – 32.

[15] Guan Wang, Wenying Gu, and Ayoung Suh. 2018. The Effects of 360-Degree VR Videos on Audience Engagement: Evidence from the New York Times. In *HCI in Business, Government, and Organizations*, Fiona Fui-Hoon Nah and Bo Sophia Xiao (Eds.). Springer International Publishing, Cham, 217–235.