Design of Time-Continuous Emotion Rating Interfaces

Jason W. Woodworth
Jason.Woodworth1@louisiana.edu
University of Louisiana at Lafayette
Louisiana, USA

Christoph W. Borst
cwborst@gmail.com
University of Louisiana at Lafayette
Louisiana, USA

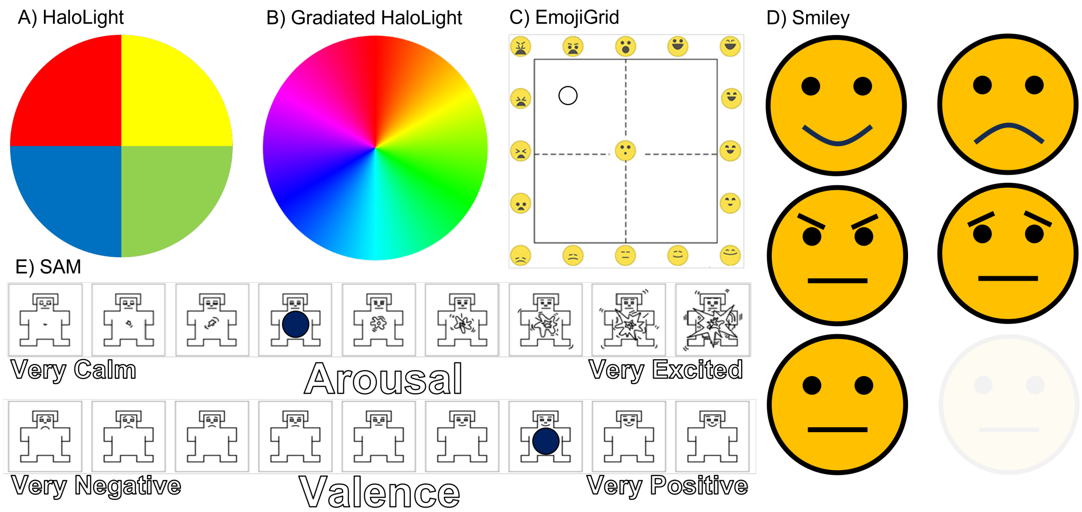

Figure 1: The five continuous rating interfaces compared in our study. Each visualizes user input on an arousal-valence scale.

ABSTRACT

We present a preliminary study on the design of visual interfaces for users to continuously rate their emotion while viewing VR content. Interfaces consist of two from the literature, a continuous adaptation of the popular SAM interface, and two novel interfaces. Designs were tested to discern what elements are intuitive or distracting. Study phases included initial impressions of interface visuals, tuning the interface control scheme, training by rating a list of emotion labels, and continuous rating of 360° video content. Results suggest that an interactive face icon (Smiley) is a promising design choice and suggest further evaluation of possible benefits.

KEYWORDS

Affective Computing, Emotion Recognition, Continuous Rating

ACM Reference Format:
Jason W. Woodworth and Christoph W. Borst. 2023. Design of Time-Continuous Emotion Rating Interfaces. In *29th ACM Symposium on Virtual Reality Software and Technology (VRST 2023), October 9–11, 2023, Christchurch, New Zealand*. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3611659.3617199

1 INTRODUCTION

We are exploring interface designs for users to give time-continuous, fine-grained reportings of emotional responses to VR content. Such reportings are used, for example, as data labels for developing emotion recognition models [3]. While these are often captured as arousal and valence measures at the end of an emotional stimulus via something like the self-assessment manikin (SAM) [1, 5], such discrete reporting does not capture the full emotional range a single stimulus elicits over time [6].

Time-continuous rating interfaces in VR have begun to be explored, with HaloLight and DotSize being recent examples [6]. We consider the need to explore a larger design space, including different control schemes and visuals. We thus introduce novel designs and continuous adaptations of existing methods with an exploratory preliminary study meant to guide future study direction.

2 INTERFACE DESIGN

All designs (Figure 1) visualize emotion based on a 2D input for an arousal-valence model of emotion. The input used Vive controller trackpads. Primary design considerations were that an interface should be intuitive, to minimize required training or reminders, and be minimally invasive, to support focus on the content.

We considered four control schemes. Two one-handed schemes used horizontal and vertical movement on one trackpad, with horizontal component being arousal (AXVY) or valence (AYVX). This allowed a user to easily specify quadrants by diagonal moves from trackpad center. Two two-handed schemes assigned arousal to left

Table 1: Summary of pilot results, indicating the number of times an interface was rated most intuitive (Int), most effective for watching videos (Eff), most invasive / distracting (MI), least invasive / distracting (LI), best overall, which control scheme was most preferred, and the most commonly given pros and cons.

	Int	Eff	MI	LI	Best	CS	Pro	Con
HL	0	0	1	1	0	AYVX	Simple	Colors not intuitive
GHL	0	2	2	1	2	AYVX	Color mixing helped precision	More colors is overwhelming
EG	1	0	1	0	0	AYVX	Labeling is intuitive	Invasive, too much to see
SAM	1	1	3	0	1	THX	Easy and precise after training	Requires training and visually large
Smiley	5	4	0	5	4	THY	Intuitive and non-invasive	Face doesn't capture all emotions

and valence to right hands, with values changed with either vertical (THY) or horizontal (THX) movement. This supported mental separation of the two dimensions. Interface visuals were placed in the bottom of the user's vision to be as unobtrusive as possible.

HaloLight (HL): This method from prior work [6] changed color of a circle to generally indicate the reported emotion. Circle opacity represented emotion "intensity", or distance from neutral.

Gradiated HaloLight (GHL): HaloLight with one color per quadrant does not provide variation within quadrants, e.g., tension and anger look the same. GHL addressed this by varying hue continuously according to an angle of the input coordinate, e.g., low arousal with neutral valence would color the circle cyan. This gave finer-grained feedback showing different emotions.

EmojiGrid (EG): A continuous adaptation of [4], EmojiGrid presented a 2D arousal-valence grid with emojis representing emotions around its edges. A moving dot showed the user's reported arousal (vertically) and valence (horizontally).

SAM: A continuous adaptation of a classic self-assessment manikin [1]. Two scales were shown with dots showing the current ratings.

Smiley: A novel approach with a dynamic face reminiscent of AffectButton [2]. Valence controlled a mouth curve, with the amount of smiling proportional to valence (neutral input gives a straight line, and lower values give a frown). Arousal was mapped to either *opacity*, with low arousal making the face closer to transparent, or *eyebrow* shape, with low arousal turning the eyebrows outward and high arousal turning them inward. We intend the face to give an intuitive visual of valence and arousal, making emotion easier to report.

3 PRELIMINARY STUDY

We gathered opinions of the five interface designs. Seven subjects participated, using a Vive Pro Eye headset and Vive controllers.

Subjects first used interfaces in an empty VR world while giving their initial impressions of what the visuals meant. Subjects then exited VR and were given brief training on the arousal-valence dimensions and how to give input. They then reentered VR and tried each interface with each of the four control schemes, picking the control scheme they preferred for each interface. They then practiced giving ratings by reading an emotion word (e.g. "happy" or "angry") and using each interface to report that emotion. Finally, they watched five 60-second 360° videos using each interface to continuously rate their emotions. Afterwards, they answered a questionnaire about their experience.

4 RESULTS

Table 1 summarizes subjective results. We see a trend of subjects finding the Smiley interface intuitive, minimally invasive, and effective for watching videos. Comments noted that the face's mouth was a clear indication of valence, and that subjects could easily find most expressions that matched their emotions. Subjects were split on if arousal should be represented by opacity (4) or eyebrows (3). The eyebrows made more sense as an emotional indicator, but could be misinterpreted in too many ways. The opacity was more abstract, but simpler to interpret once subjects knew what it meant. Most (5) subjects preferred two-handed input, saying it was easier to focus on each dimension individually and avoid accidental changes.

Results on HL and GHL suggest that the colors are not universal. During initial impressions, only 2 subjects correctly interpreted the emotions represented by all 4 colors. Those who liked them generally preferred GHL for its increased precision; the 4 colors made subjects feel that they were choosing from only 4 options. Feedback on SAM showed it was intuitive and precise, but only after training on the arousal-valence model, and it was large and distracting during video. Feedback on EG showed that the faces were intuitive and subjects liked seeing their input directly mapped to a grid, but showing all of the faces at once was too distracting during video. These results will be used to further improve designs and continue their comparison to other interfaces.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 1815976.

REFERENCES

[1] M Bradley and Peter J Lang. 1994. Measuring Emotion: The Self-Assessment Manikin and the Semantic Differential. *Journal of Behavior Therapy and Experimental Psychiatry* 25 (1994), 49–59. Issue I.

[2] Joost Broekens and Willem-Paul Brinkman. 2013. AffectButton: A method for reliable and valid affective self-report. *International Journal of Human-Computer Studies* 71, 6 (2013), 641–667.

[3] Javier Marín-Morales, Carmen Llinares, Jaime Guixeres, and Mariano Alcañiz. 2020. Emotion recognition in immersive virtual reality: From statistics to affective computing. *Sensors (Switzerland)* 20 (2020), 1–26. Issue 18.

[4] Alexander Toet, Daisuke Kaneko, Shota Ushiama, Sofie Hoving, Inge de Kruijf, Anne-Marie Brouwer, Victor Kallen, and Jan B. F. van Erp. 2018. EmojiGrid: A 2D Pictorial Scale for the Assessment of Food Elicited Emotions. *Frontiers in Psychology* 9 (2018).

[5] Jason W Woodworth and Christoph W Borst. 2021. Designing immersive affective tasks for emotion elicitation in virtual reality. In *2nd momentary emotion elicitation and capture workshop*.

[6] Tong Xue, Abdallah El Ali, Tianyi Zhang, Gangyi Ding, and Pablo Cesar. 2021. CEAP-360VR: A Continuous Physiological and Behavioral Emotion Annotation Dataset for 360 VR Videos. *IEEE Transactions on Multimedia* (2021).