2025 17th International Conference on COMmunication Systems and NETworks (COMSNETS) | 979-8-3315-3119-5/25/$31.00 ©2025 IEEE | DOI: 10.1109/COMSNETS63942.2025.10885663

PResCon: Physiological Response Augmented Continuous Emotion Annotation Tool

Swarnali Banik*, Akhilesh Adithya†§, Bivas Mitra¶, Sougata Sen*‡, Snehanshu Saha*‡, Surjya Ghosh*‡
*Department of Computer Science and Information Systems, BITS Pilani K K Birla Goa Campus, Goa, India
‡APPCAIR, BITS Pilani K K Birla Goa Campus, Goa, India
¶Department of Computer Science and Engineering, Indian Institute of Technology Kharagpur, India
†Amazon, India
Email: {p20210016, sougatas, snehanshus, surjyag}@goa.bits-pilani.ac.in, bivas@cse.iitkgp.ac.in, akhileshadithya311@gmail.com

***Abstract*—We present an annotation tool designed to collect fine-grain continuous emotion ground truth labels and physiological responses as users watch different videos. In the existing literature, continuous emotion annotations are typically collected as self-reports based on the 2D Circumplex model through auxiliary devices (a joystick or similar tools) during video consumption. However, the users face challenges as they need to use an auxiliary device and rely on the Circumplex model's valence and arousal axis without clear indication of distinct emotion labels. Therefore, to overcome these challenges, we introduce PResCon, a tool that allows to collect emotion annotations using keyboard arrow keys only (without using any auxiliary device) and guiding users towards the discrete emotions (e.g., amused, stressed, bored, relaxed) by embedding them on the Circumplex plane. The tool encompasses two major components - (a) a web interface that displays various stimuli videos and provision of collecting emotion labels (valence and arousal) by moving the keyboard arrow keys and (b) a sensor setup (based on Arduino) that allows to collect the physiological responses (heart rate and galvanic skin response) passively during video consumption. In summary, the PResCon tool facilitates research in the domain of affective computing by allowing to collect fine-grain continuous emotion annotations and the physiological responses in real-time during video consumption.**

I. Introduction

At the core of many web-based video platforms (e.g., online meeting [1], online tutoring [2], online gaming [3]) lies a machine learning model that infers emotion continuously to moderate the content delivery for an engaging user experience. Typically, to train such emotion inference models, the emotion labels are collected as self-reports using an auxiliary device (e.g., joystick [4], mouse [5], or similar device) during video consumption.

The existing annotation tools such as Feeltrace [5], GTrace [6], DARMA [7], ANVIL [8] require participants to continuously input the emotions as they watch the videos, while physiological signals (e.g., EDA, HR) get recorded passively. Similarly, Sharma et al. demonstrated that emotion annotations can be collected by continuously moving a joystick on the Circumplex plane [9]. The major challenge with the existing continuous emotion annotations tools are two-fold. First, participants often need to use an auxiliary device (e.g., joystick or slider), thus making the wide-scale applicability of the tool restricted. Second, as the Circumplex model represents emotions on a 2D plane (where the x and y axis represent valence and arousal respectively), the participants may not always relate with the discrete emotion (e.g., happy, sad, stressed, relaxed) that they perceive at a given moment, thus making the annotation challenging.

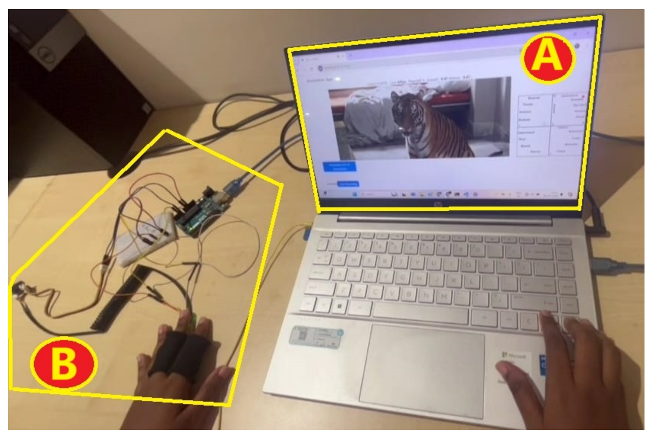


Fig. 1: The overview of the PResCon, which includes an annotation interface (marked as Ⓐ) for video playback and continuous emotion annotation based on the Circumplex model [9] and the physiological sensors setup (marked as Ⓑ).

We, in this paper, design and implement PResCon, a continuous emotion annotation tool (Fig. 1) to address the aforementioned challenges. The tool facilitates the continuous emotion annotation by allowing to record the annotation using keyboard arrow keys and projecting a number of discrete emotions on different quadrants of the Circumplex plane based on their valence and arousal representation. The provision of seeing the discrete emotions on the Cirumplex plane of the annotation tool assists participants to accurately identify the perceived emotions and record them in real-time. Moreover, as the users use keyboard arrow keys to indicate the emotion annotations, any auxiliary device is not required. The tool encompasses a web interface, which displays the stimuli videos and allows the participants to annotate them continuously. At the same time, physiological signals, such

§Work done by the author during the period of staying at BITS Goa.

as Galvanic Skin Response (GSR) and Heart Rate (HR), are recorded through sensors (connected via an Arduino board), capturing the user's physiological responses to complement the self-reported emotion labels. This combination of continuous emotion self-report and physiological response enriches the emotional profile recorded during the video stimuli.

To demonstrate the utility of PResCon, we performed a user study (N=12). The participants watched eight stimuli videos as displayed, passively recorded the physiological signals (GSR and HR), and continuously provided emotion annotations for the stimuli videos. The dataset gathered from the study reveals a fine-grain continuously annotated emotions coupled with physiological responses for the stimuli videos.

II. Implementation of PResCon Annotation Tool

In this section, we provide an overview of the PResCon, which consists of two key components as shown in Fig. 1 - (i) annotation interface and (ii) physiological sensor configuration. This annotation tool is designed to collect participants' physiological responses and continuous emotion annotations during video consumption. The collected data using this tool will help to validate the relationship between emotion variations and physiological responses.

A. Annotation Interface

The annotation interface of PResCon (see Fig. 2a) enables users to provide continuous emotion annotations using keyboard arrow keys while they watch the videos. These annotations are based on the Circumplex model [9], which maps human emotions onto a two-dimensional plane with two axes: valence (pleasure) and arousal (activeness), segmented into four quadrants (Fig. 2b). Notably, the interface displays a number of discrete emotions in each quadrant of the Circumplex model to facilitate the annotation process. When the application launches, a red dot cursor appears at the origin of this Circumplex model (Fig. 2b), indicating a neutral starting position. As the video plays, users navigate the arrow keys to reflect their experienced emotions in terms of valence and arousal. At each timestamp, the cursor's position reflects these ratings for the video content. Additionally, upon activating the annotation interface, the Arduino begins passively recording physiological signals as users annotate their emotions during the video playback.

B. Sensor Configuration

The sensor setup of PResCon consists of a pulse rate sensor (HW-827, World Famous Electronics LLC) and a galvanic skin response (GSR V1.2, Seeed Studio Grove) sensor, both connected to the GPIO pins on an Arduino Uno board. These sensors capture heart rate (HR) and skin conductance, respectively, each at a sampling rate of 10 Hz. The data collected by the sensors is then transmitted to a connected laptop via the serial port (Fig. 2c). The choice of these sensors balances affordability with real-time, accurate, and precise measurements, making the setup practical for a variety of applications [10].

III. Description of PResCon Annotation Demo

This section outlines the demo procedure of PResCon. We performed a user study in a lab-based environment. Prior to the study, IRB approval (HEC/BPGC/2023/005) was obtained from our institute's Human Ethics Committee (HEC).

Video ID	Emotion	Valence	Arousal	Duration (in sec)
1	Amusing	Med/High	Med/High	185
2				173
3	Boring	Low	Low	119
4				160
5	Relaxing	Med/High	Low	145
6				147
7	Scary	Low	High	197
8				144

TABLE I: The stimuli videos used in the PResCon annotation demo. These videos were previously utilized in studies focused on physiological responses and continuous emotion annotations [4].

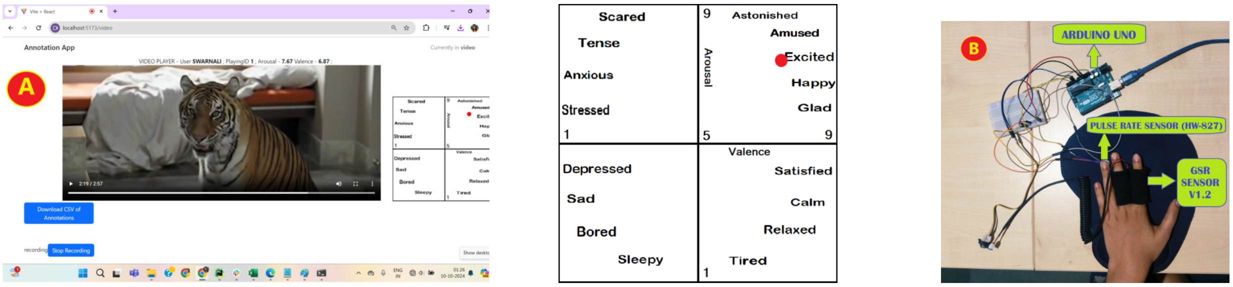
We recruited 12 participants, including 6 females and 6 males, aged between 20 and 40 years, from our university. During the study, each participant watched a set of 8 videos and continuously provided their emotion annotations using the annotation interface of PResCon. The sensors passively recorded physiological signals (i.e., GSR and HR) throughout each session. The selected videos (Table I) represent distinct emotions from each quadrant of the Circumplex model, covering emotions such as amused, relaxed, bored, and scared. These videos have been previously used in similar research [4]. Details of each video and its associated emotions are presented in Table I.

Every participant watched the videos in a predetermined random order, continuously recording valence and arousal scores on a 1 to 9 scale using the keyboard arrow keys, with scores automatically assigned based on the cursor position. During annotations, participants only needed to navigate the cursor using the arrow keys, and scores were generated automatically based on the cursor position. The annotation interface displayed a circumplex plane with discrete emotion labels (e.g., amused, stressed, bored, relaxed), assisting participants in positioning the cursor in line with their perceived emotions. As the cursor moved across the circumplex plane, valence and arousal values were recorded on a 9-point scale, effectively capturing participants' felt emotions without requiring specific numerical input. This process reduced cognitive load and guided participants toward accurate annotations. A two-minute blue screen was shown between videos to prevent any carry-over effect. Before the study, participants were briefed on the concepts of valence and arousal and instructed on using the arrow keys to record emotions during video playback.

IV. Preliminary Evaluation and Data Analysis

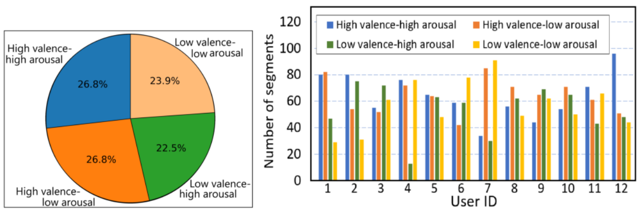
We performed preliminary analysis of the collected dataset to make sure that the tool could collect the continuous annotations and the physiological responses. We carried out the following pre-processing tasks.

First, we segmented the physiological responses into fixed-size (i.e., 5-second) windows (this window length was also



(a) The annotation interface of PResCon **(b)** Circumplex model [9] **(c)** The sensor setup of PResCon

Fig. 2: The PResCon setup: (a) the UI of the annotation interface for video playback and emotion annotation; (b) the Circumplex model of emotion, which guides the emotion annotation process on a 2D plane; and (c) the sensor configuration for collecting physiological responses.



(a) Distribution in total dataset **(b)** User-wise distribution

Fig. 3: Distribution of valence-arousal across video segments (a) in the total dataset, (b) in each user.

adopted in earlier works [11]) for every participant and video combination. Hence, we obtained a total of 2871 segments and, on average 239 segments (SD: 1.16) from each user. Next, we labelled each of these segments based on the emotion annotation provided by the user during that segment. In this analysis, we denoted the valence (or arousal) as high when the ground truth is greater than 5 and as low when the ground truth is 5 or below. Since scores are on a 1 to 9 scale, we used 5 as the midpoint to label valence (or arousal) as high (or low) for every user. Therefore, we obtained four distinct levels of valence and arousal, i.e., high valence–high arousal. high valence–low arousal, low valence–high arousal, and low valence–low arousal. We have 770 segments for high valence–high arousal, another 770 segments correspond to high valence–low arousal, 646 to low valence–high arousal, and 685 to low valence–low arousal (as shown in Fig. 3a) out of 2871 total segments. The four distinct levels of valence and arousal for each user are shown in Fig. 3b. These details reveal that the PResCon can collect the physiological responses and the continuous emotion annotations.

V. Conclusion and Future Work

In this paper, we designed and implemented PResCon, a tool to capture continuous emotion annotations and passively sense physiological responses (GSR, HR) during video consumption. Unlike traditional continuous annotation tools, this tool does not rely on any auxiliary device for annotation collection and facilitate emotion self-reporting by embedding a number of discrete emotions on the circumplex model. We demonstrated the utility of PResCon in a lab-based controlled study (N=12), which reveals that it can collect physiological responses tagged with continuous emotion annotations. We aim to evaluate the performance of PResCon in a more realistic setting with a large and diverse population in our future work.

VI. Acknowledgement

This research has been supported by the AI4ICPS research grant (TRP3RDED02178, dated: 20/02/2024), Chanakya Ph.D. Fellowship of AI4ICPS (IIT Kharagpur), and the SURE grant (SUR/2022/001965) of SERB, DST, Government of India.

References

[1] P. Murali, J. Hernandez, D. McDuff, K. Rowan, J. Suh, and M. Czerwinski, "Affectivespotlight: Facilitating the communication of affective responses from audience members during online presentations," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–13.

[2] M. A. Hasan, N. F. M. Noor, S. S. B. A. Rahman, and M. M. Rahman, "The transition from intelligent to affective tutoring system: a review and open issues," *IEEE Access*, vol. 8, pp. 204 612–204 638, 2020.

[3] Y. Gao, Y. Jin, S. Choi, J. Li, J. Pan, L. Shu, C. Zhou, and Z. Jin, "Sonicface: Tracking facial expressions using a commodity microphone array," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 4, pp. 1–33, 2021.

[4] K. Sharma, C. Castellini, F. Stulp, and E. L. Van den Broek, "Continuous, real-time emotion annotation: A novel joystick-based analysis framework," *IEEE Transactions on Affective Computing*, vol. 11, no. 1, pp. 78–84, 2017.

[5] R. Cowie, E. Douglas-Cowie, S. Savvidou, E. McMahon, M. Sawey, and M. Schröder, "'FEELTRACE': An instrument for recording perceived emotion in real time," in *ITRW Speech-Emotion*, 2000.

[6] R. Cowie, M. Sawey, C. Doherty, J. Jaimovich, C. Fyans, and P. Stapleton, "Gtrace: General trace program compatible with emotionml," in *2013 humaine association conference on affective computing and intelligent interaction*. IEEE, 2013, pp. 709–710.

[7] J. M. Girard and A. G. C. Wright, "Darma: Software for dual axis rating and media annotation," *Behavior research methods*, vol. 50, pp. 902–909, 2018.

[8] M. Kipp, "Anvil-a generic annotation tool for multimodal dialogue," in *Seventh European conference on speech communication and technology*. Citeseer, 2001.

[9] J. A. Russell, "A circumplex model of affect." *Journal of personality and social psychology*, vol. 39, no. 6, p. 1161, 1980.

[10] V. Mishra, G. Pope, S. Lord, S. Lewia, B. Lowens, K. Caine, S. Sen, R. Halter, and D. Kotz, "Continuous detection of physiological stress with commodity hardware," *ACM transactions on computing for healthcare*, vol. 1, no. 2, pp. 1–30, 2020.

[11] A. Adithya, S. Tiwari, S. Sen, S. Chakraborty, and S. Ghosh, "Ocean: Towards developing an opportunistic continuous emotion annotation framework," in *2022 IEEE International Conference On Pervasive Computing And Communications Workshops And Other Affiliated Events (PerCom Workshops)*. IEEE, 2022, pp. 9–12.