Brain 2 Communicate: EEG-based Affect Recognition to Augment Virtual Social Interactions

Daniel Roth
daniel.roth@uni-wuerzburg.de
University of Würzburg, HCI Group

Franziska Westermeier
University of Würzburg, HCI Group
Würzburg

Larissa Brübach
University of Würzburg, HCI Group
Würzburg

Tobias Feigl
FAU Erlangen-Nürnberg
Fraunhofer IIS, Nürnberg

Christian Schell
University of Würzburg, HCI Group
Würzburg

Marc Erich Latoschik
University of Würzburg, HCI Group
Würzburg

ABSTRACT

The perception and expression of emotion is a fundamental part of social interaction. This project aims to utilize neuronal signals to augment avatar-mediated communications. We recognize emotions with a brain-computer-interface (BCI) and supervised machine learning. Using an avatar-based communication interface that supports head tracking, gaze tracking, and speech to animation, we leverage the BCI-based affect detection to visualize emotional states.

KEYWORDS

affective computing, avatars, brain-computer interfaces

Figure 1: The current communication platform prototype. Upon detection of emotional sates, a visual effect appears.

1 INTRODUCTION

Social applications for Virtual-, Mixed, and Augmented Reality (VR, MR, AR) allow to represent users as avatars, virtual characters driven by human behavior [1]. Nonverbal behavior such as body movements, facial expressions, and gaze can be sensed and replicated to avatars to large extent [8]. Yet, it remains challenging to sense and display a user's feelings and intentions. Emotions, for example, are dependent from context and environment, culture, and personality. They are evoked through perception and physiological stimulation, and are relevant for human decision making [2]. Even sophisticated systems may not capture emotions to their full subtlety and diversity. In addition to merely replicating nonverbal behavior, the present work includes non-perceivable neuronal signals acquired through a BCI. BCIs measure brain activity, for example based on electroencephalography (EEG) [5], to allow for its interpretation and consecutively the interaction with computers. First, using such sensing mechanisms could be beneficial for people with communication disorders as interpreting and especially non-verbally expressing emotions can be challenging for them. Second, it could eliminate confusion as emotions could be visually transformed and displayed in a clear and understandable way which could reduce the overall amount necessary sensors.

2 APPROACH

For a first prototype, we use a desktop-based avatar interaction platform implemented with Unity3D (see Fig. 1). Similar to a video conference the networked simulation allows two users to exchange verbally via a headset [7]. However, instead of transmitting a video signal, we retarget head and gaze behavior to avatars and use voice-to-animation to visualize the verbal exchange. To derive affective states, we interface the OpenBCI EEG-based device with a Python application. Similar to other approaches for affect recognition [4], we used 8 electrodes of the frontal brain area (see Fig. 2) as data input. The data is preprocessed using bandpass and notch filtering and detrending to acquire beta and gamma waves respectively (16 features), and their ratio of asymmetry between the hemispheres was calculated (8 features). Together with the time dimension, this resulted in 25 features.

MuC'19 Workshops, Hamburg, Deutschland

https://doi.org/10.18420/muc2019-ws-571

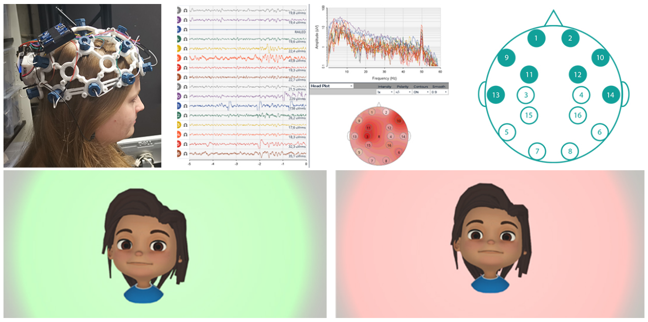

Figure 2: *Top:* The BCI, the EEG signals in the OpenBCI software, and the selected electrodes. *Bottom:* The augmentations for happy (green) and angry (red).

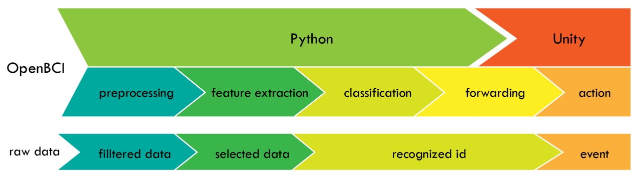

Figure 3: Processing Pipeline. Features (β, γ, and measures from opposing electrodes) are preprocessed and classified using a SVM. Interfacing with Unity, the Python application updates the current state and triggers the augmentations.

For a first proof of concept, we trained a Support Vector Machine (SVM; one-vs-one, polynomial kernel $(\gamma\langle x, x'\rangle + r)^d$, with $d$egree=3, and r=0) with labeled data from 208,498 samples of one user watching video material that evokes angry or happy emotional reactions (e.g., [9]). The data (2729 windows with 60 frames each with 25 features each, no overlap) was split into training (78%) and test (22% cross-fold). The trained SVM is then used to perform a real-time classification during a social interaction. In the case of detection of an affective state, the Python application informs the Unity application about the change in state (see Fig. 3) and the respective visual effect is displayed (see Fig. 2). Emotions were visualized by a background effect. We chose red for the negative state as it is a color of aggression and green for a positive state as it indicates relaxation and happiness [3].

3 PRELIMINARY EVALUATION

Assessing the current prototype, the detection rate of an offline testing is rather reliable (neutral: 0.96; happy: 0.89; angry: 0.92). Yet, for a live detection, the preprocessing and training is not sufficient to always reliably detect the emotional state. We found this to be mainly the cause of the training stimuli and the restricted dataset. While we used funny videos or videos that evoke angry and negative emotions to train the SVM, emotions expressed and perceived in social interactions may result in different neuronal processing than mere passive reactions. However, it is complex to gather such natural training data. Furthermore, a recurrent neural network may be an alternative approach to our application. Last, while we did not perform formal latency measurements yet, we noticed that the preprocessing is introducing some delay, which is why we aim at further improving the preprocessing pipeline to reduce the computational cost.

4 DISCUSSION

In this paper we presented an approach to augment user-embodied virtual interactions through affect recognition with an EEG-based BCI. Further work should also explore different visualization possibilities. For example, the avatar itself could change the skin color, or emotional states could trigger blendshape activities such as a smile. Right now our approach is trained for an individual subject and thus the system has to be trained on the individual. To generalize it we need data from a high number of subjects, possibly exposed to more natural stimuli. Future work will further make use of EEG hyperscanning to allow for multi-user augmentations. Furthermore, our approach could be contrasted to other typical approaches to derive affective states, such as facial expression classification or physiological sensing [6], for example, with regard to detection rates, system pricing, usability, and noise resistance. Overall, we think our approach could be beneficial to further explore the plasticity of virtual social interactions, which could be benficial for assistive technologies, training applications, and social virtual environments.

REFERENCES

[1] Jeremy N Bailenson and Jim Blascovich. 2004. Avatars. In *Encyclopedia of Human-Computer Interaction*. Berkshire Publishing Group, 64–68.

[2] Manfred Holodynski. 2006. *Emotionen - Entwicklung und Regulation*. Springer-Verlag.

[3] Naz Kaya and Helen H. Epps. 2004. Relationship between color and emotion: A study of college students. 38, 3 (2004), 396–405.

[4] Y. Lin, C. Wang, T. Jung, T. Wu, S. Jeng, J. Duann, and J. Chen. 2010. EEG-Based Emotion Recognition in Music Listening. *IEEE Transactions on Biomedical Engineering* 57, 7 (2010), 1798–1806. https://doi.org/10.1109/TBME.2010.2048568

[5] Luis Fernando Nicolas-Alonso and Jaime Gomez-Gil. 2012. Brain Computer Interfaces, a Review. *Sensors* 12, 2 (2012), 1211–1279. https://doi.org/10.3390/s120201211

[6] Rosalind W Picard. 2000. *Affective computing*. MIT press.

[7] Daniel Roth, Peter Kullmann, Gary Bente, Dominik Gall, and Marc Erich Latoschik. 2018. Effects of Hybrid and Synthetic Social Gaze in Avatar-Mediated Interactions. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 103–108. https://doi.org/10.1109/ISMAR-Adjunct.2018.00044

[8] Daniel Roth, Kristoffer Waldow, Marc Erich Latoschik, Arnulph Fuhrmann, and Gary Bente. 2017. Socially immersive avatar-based communication. In *2017 IEEE Virtual Reality (VR)*. IEEE, 259–260. https://doi.org/10.1109/VR.2017.7892275

[9] Ghost Tobi. 2019. 400 Flachwitze Teil 1 - YouTube. https://www.youtube.com/watch?v=Odtga14v9Ic