

A Social Interaction Interface Supporting Affective Augmentation Based on Neuronal Data

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

Larissa Brübach
University of Würzburg, HCI Group
larissa.bruebach@stud-mail.uni-wuerzburg.de

Franziska Westermeier
University of Würzburg, HCI Group
franziska.westermeier@stud-mail.uni-wuerzburg.de

Tobias Feigl
FAU Erlangen-Nürnberg
Fraunhofer IIS, Nürnberg
feiglts@iis.fraunhofer.de

Christian Schell
University of Würzburg, HCI Group
christian.schell@uni-wuerzburg.de

Marc Erich Latoschik
University of Würzburg, HCI Group
marc.latoschik@uni-wuerzburg.de

ABSTRACT

In this demonstration we present a prototype for an avatar-mediated social interaction interface that supports the replication of head- and eye movement in distributed virtual environments. In addition to the retargeting of these natural behaviors, the system is capable of augmenting the interaction based on the visual presentation of affective states. We derive those states using neuronal data captured by electroencephalographic (EEG) sensing in combination with a machine learning driven classification of emotional states.

CCS CONCEPTS

• **Human-centered computing** → *Mixed / augmented reality*.

KEYWORDS

Communication interfaces; embodiment; affective computing; avatars; brain-computer interfaces

ACM Reference Format:

Daniel Roth, Larissa Brübach, Franziska Westermeier, Tobias Feigl, Christian Schell, and Marc Erich Latoschik. 2019. A Social Interaction Interface Supporting Affective Augmentation Based on Neuronal Data. In *Symposium on Spatial User Interaction (SUI '19)*, October 19–20, 2019, New Orleans, LA, USA. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3357251.3360018>

1 INTRODUCTION

Today's platforms and research prototypes for communicative interfaces in Virtual-, Mixed-, and Augmented Reality (VR, MR, AR) often allow to replicate spatial communicative behaviors such as body motion, gaze, and facial expression [4, 7]. Recent research has begun to augment these natural behaviors with additional social information, such as visual transformations [5], affective displays [3], or modifications of natural behaviors [1, 6]. To explore such augmentations on the basis of neuronal data, we presented a Brain2Communicate, a prototype to perform such augmentations on the basis of affective states that are derived from EEG data [8]. In this demonstration, we

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

SUI '19, October 19–20, 2019, New Orleans, LA, USA

© 2019 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-6975-6/19/10.

<https://doi.org/10.1145/3357251.3360018>

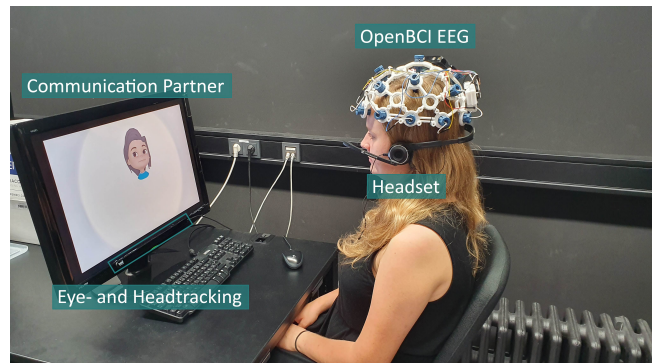


Figure 1: Hardware description. A Tobii eyetracker tracks the user's gaze and head movements. Movements of the communication partner are replicated to an avatar. Audio data is streamed and used for audio transmission and voice-to-mouth animation. An OpenBCI EEG headset senses neuronal activity and transmits EEG data. The data is preprocessed and classified using a recurrent neuronal network. Tracking data, audio data and classified affective states are streamed to a Unity 3D Network application to synchronize the environment.

demonstrate the prototype with additional features and an iterated processing pipeline.

2 APPROACH

2.1 Classification of Affective States

In previous work [8], we trained an SVM using the default implementation of the scikit-learn project, leveraging a polynomial kernel. While the accuracy was high during offline classification, we did notice that it was inaccurate during real-time classification. We therefore trained a Long short-term memory (LSTM) recurrent neural network using the PyTorch framework with one layer and a hidden layer size of 100 and the output feeding directly fed the softmax output layer. We chose happy and sad as affective states to identify, as those states were found to be evoked more often in natural interactions in pretesting. Video stimuli were used for data acquisition to trigger those affective states. The data was annotated

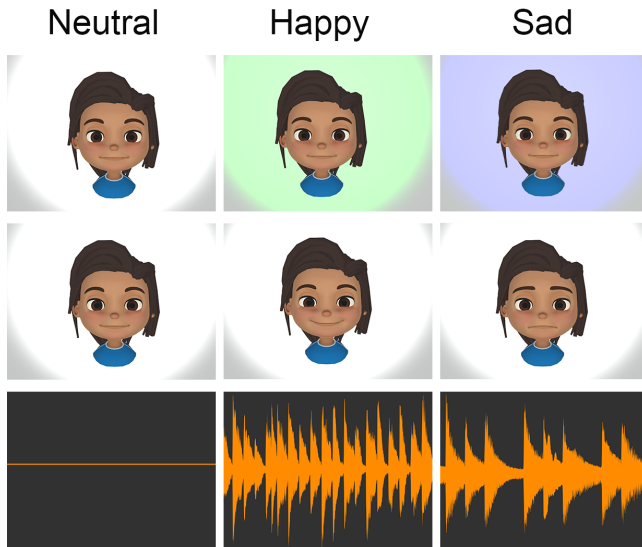


Figure 2: Illustration of the designed augmentation types. Top: Mood-based color transformation. Center: Facial expression changes. Bottom: Mood-based audio playback.

by a visual examination of the participant's facial expressions that were recorded during the acquisition. We evaluated the LSTM based on a training set that was split 6/7 (training) to 1/7 (testing). Out of the three possible states, the LSTM could correctly detect an affective state with 44% accuracy.

3 AUGMENTATION TYPES

To address multiple options and use cases, We designed three augmentation types for the current prototype that include a visual, a behavioral, and an auditive augmentation.

- **Color transformations.** On the basis of a detected affective state, the background color changes according to a color representing the affect to underline the mood.
- **Audio background.** If sad or happy states are detected, the background audio changes to a sad (minor) or happy (major) tone.
- **Facial expression augmentation.** To directly transmit the affective state, we conceptualized a mapping between affective state and resulting avatar facial expression.

The goal of the design was to provide support for mood detection in various communicative scenarios in multi-user virtual interactions. For example, with the support of audio augmentations users could infer affective states without the visual presence of avatars, for example in VR applications where the orientation is not facing the communication partner. Whereas an augmentation of facial expressions may be beneficial in a general context, a supportive visual transformation may have a benefit for users with difficulties to infer emotional states from nonverbal behavior.

4 DISCUSSION

The current prototype can detect states based on a training set that is captured from a single participant. While further improvements

to the detection method and preprocessing should be performed, the LSTM performs better in real-time classification compared to the support vector machine used in previous work [8]. A current limitation is the data set used for the training, which is relatively small. We therefore aim to collect a database that includes multiple users with the hope to generalize the detection. However, based on the current testing, an individual training template is beneficial. Furthermore, ethical considerations need to be addressed regarding the use of social augmentations based on neuronal data, which is subject to future work. In future iterations, we will also explore avatar plasticity and thus form and appearance changes explored in related works [2, 5].

5 CONCLUSION

In this demonstration, we present a prototype to augment social interactions in avatar-mediated communication scenarios based on neuronal data acquired through a brain-computer interface. We designed three distinct augmentation types, a visual transformation, a behavior augmentation, as well as an auditive augmentation. In contrast to a previous approach [8], we chose an alternative classification method to improve the accuracy of real-time classification in natural interaction scenarios. A formal evaluation is subject to present work.

ACKNOWLEDGMENTS

We thank Roman Eyck and Peter Kullmann for their assistance in developing prototypes of the avatar platform.

REFERENCES

- [1] Jeremy N Bailenson, Andrew C Beall, Jack Loomis, Jim Blascovich, and Matthew Turk. 2004. Transformed social interaction: Decoupling representation from behavior and form in collaborative virtual environments. *Presence: Teleoperators & Virtual Environments* 13, 4 (2004), 428–441.
- [2] Jeremy N Bailenson, Nick Yee, Dan Merget, and Ralph Schroeder. 2006. The effect of behavioral realism and form realism of real-time avatar faces on verbal disclosure, nonverbal disclosure, emotion recognition, and copresence in dyadic interaction. *Presence: Teleoperators and Virtual Environments* 15, 4 (2006), 359–372.
- [3] Jonathon D Hart, Thammathip Piumsombon, Louise Lawrence, Gun A Lee, Ross T Smith, and Mark Billinghurst. 2018. Demonstrating Emotion Sharing and Augmentation in Cooperative Virtual Reality Games. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*. ACM, 117–120.
- [4] David Roberts, Robin Wolff, John Rae, Anthony Steed, Rob Aspin, Moira McIntyre, Adriana Pena, Oyewole Oyekoya, and Will Steptoe. 2009. Communicating eye-gaze across a distance: Comparing an eye-gaze enabled immersive collaborative virtual environment, aligned video conferencing, and being together. In *2009 IEEE Virtual Reality Conference*. IEEE, 135–142.
- [5] D. Roth, C. Kleinbeck, T. Feigl, C. Mutschler, and M. E. Latoschik. 2018. Beyond Replication: Augmenting Social Behaviors in Multi-User Virtual Realities. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 215–222. <https://doi.org/10.1109/VR.2018.8447550>
- [6] 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>
- [7] Daniel Roth, Kristoffer Waldow, Felix Stetter, Gary Bente, Marc Erich Latoschik, and Arnulph Fuhrmann. 2016. SIAMC: A Socially Immersive Avatar Mediated Communication Platform. In *Proceedings of the 22Nd ACM Conference on Virtual Reality Software and Technology (VRST '16)*. ACM, New York, NY, USA, 357–358. <https://doi.org/10.1145/2993369.2996302>
- [8] Daniel Roth, Franziska Westermeier, Larissa Brübach, Tobias Feigl, Christian Schell, and Marc Erich Latoschik. 2019, to appear. Brain 2 Communicate: EEG-based Affect Recognition to Augment Virtual Social Interactions. In *Proceedings of the Mensch und Computer 2019 Workshop on User-Embodied Interaction in Virtual Reality*.