Development of a Prototype AI System for Real-time Emotion Prediction and Mental State Adjustment

Akihiro Sasaki[1], Eriko Sugisaki[1], Roberto Legaspi[2], Yasushi Naruse[3], Nao Kobayashi[1]

[1] Healthcare Medical Group, Life Science Laboratories, and [2]AI Division, KDDI Research, Inc., Fujimino, Japan
[3]Center for Information and Neural Networks, National Institute of Information and Communications Technology, Kobe, Japan

1. Introduction

Our emotional state influences our daily performance. Recent reports have highlighted a hormetic relationship between stress and cognitive performance [1, 2], suggesting that understanding the optimal balance of emotional states, rather than adopting a simple negative-positive emotion perspective, could potentially enhance the optimization of behavioral performance. Our ultimate goal is to construct a system that discerns individual emotional states from physiological information, and generates music, visuals, or conversations as means to facilitate the individual's transition toward their desired emotional state. To achieve this, it is essential to evaluate complex emotional states on different emotional axes and assess possibly continuously fluctuating emotional states in real-time as possible.

2. Related works

There are several studies on emotion estimation using physiological data (ECG, GSR, EEG), such as predicting depressive mood when listening to news audio [3], and predicting the dynamics of mood during video game play [4, 5]. Particularly, the method of Ishikawa et al. [6] accomplished simultaneous estimation of emotional states along six axes, namely, Sad-Happy, Nervous-Relaxed, Fear-Relieved, Lethargic-Excited, Depressed-Delighted, and Angry-Serene, by incorporating cross-modal factors across multiple physiological modalities into the prediction model. Evaluating its accuracy by the mean absolute percentage error (MAPE), the method achieved less than 25% and 36% error rates for the Angry-Serene and the Fear-Relieved axes, respectively.

3. Emotion Estimation System During Music Listening

Here, we present an emotion estimation system that provides predicted emotional values on six emotional axes, similar to Ishikawa et al. [6]. In our system, notably, we have incorporated the real-time prediction capability, allowing us to update the predicted values every 0.5 seconds.
Model construction: We trained our model on 2,322 instances from 54 participants, each giving physiological (EEG and ECG) and emotional rating data. Participants listened to music for a minute while recording EEG and ECG, then rated their emotion on a 15-point scale for six emotional axes. Explanatory variables were taken from the 10 seconds of EEG and ECG data immediately preceding the emotion rating, and the emotion ratings served as the objective variables. The model was trained using XGboost due to its computational efficiency, smaller resource requirement, and faster training speed, enabling real-time estimation.
Implementation of emotion estimation: Users wear EEG and ECG devices and transmit the measured data to a computer via a smartphone using Bluetooth. Once the computer accumulates 10 seconds of data, it begins estimating the emotional state. The emotional state estimate is then

In: Kiemute Oyibo, Wenzhen Xu, Elena Vlahu-Gjorgievska (eds.): The Adjunct Proceedings of the 19th International Conference on Persuasive Technology, April 10, 2024, Wollongong, Australia
EMAIL: xakh-sasaki@kddi.com (A. Sasaki); no-kobayashi@kddi.com (N. Kobayashi);
ORCID: 0000-0002-2249-4975 (A. Sasaki); 0000-0002-7634-0031 (E. Sugisaki); 0000-0001-8789-0429 (Y. Naruse); 0009-0003-5533-2918 (N. Kobayashi)

updated every 0.5 seconds based on the preceding 10 seconds of data. This system, as shown in Figure 1, enables the real-time estimation of emotional state transitions during music or visual content consumption.

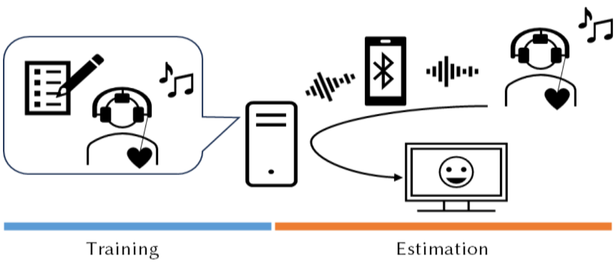

Figure 1: Schematic diagram of our system design: The model uses EEG, ECG, and emotion ratings for training. Implementation involves continuously transferring these data via Bluetooth to a computer, updating emotion predictions every 0.5 seconds based on the previous 10 seconds of data.

4. Prospects and potential applications

In the future, we plan a comprehensive evaluation of our models' accuracy using metrics, such as MAPE and other relevant measures, aiming for less than 20% error in both training and real-world application. Future developments will merge AI to generate music or visuals, guiding users to their desired emotional states, potentially enhancing presentations, work efficiency, and mental health.

Acknowledgements

This work was partially supported by Innovative Science and Technology Initiative for Security (JPJ004596), Acquisition, Technology, and Logistics Agency, Japan.

References

[1] A. Oshri, Z. Cui, M. M. Owens, C. A. Carvalho, and L. Sweet, "Low-to-moderate level of perceived stress strengthens working memory: Testing the hormesis hypothesis through neural activation," *Neuropsychologia*, vol. 176, p. 108354, Nov. 2022, doi: 10.1016/j.neuropsychologia.2022.108354.

[2] K. A. James, J. I. Stromin, N. Steenkamp, and M. I. Combrinck, "Understanding the relationships between physiological and psychosocial stress, cortisol and cognition," *Front. Endocrinol.*, vol. 14, 2023, doi: 10.3389/fendo.2023.1085950.

[3] K. Fuseda, H. Watanabe, A. Matsumoto, J. Saito, Y. Naruse, and A. S. Ihara, "Impact of depressed state on attention and language processing during news broadcasts: EEG analysis and machine learning approach," *Sci. Rep.*, vol. 12, no. 1, p. 20492, Nov. 2022, doi: 10.1038/s41598-022-24319-x.

[4] Y. Yokota, T. Soshi, and Y. Naruse, "Error-related negativity predicts failure in competitive dual-player video games," *PLOS ONE*, vol. 14, no. 2, p. e0212483, Feb. 2019, doi: 10.1371/journal.pone.0212483.

[5] Y. Yokota and Y. Naruse, "Temporal Fluctuation of Mood in Gaming Task Modulates Feedback Negativity: EEG Study With Virtual Reality," *Front. Hum. Neurosci.*, vol. 15, p. 536288, Jun. 2021, doi: 10.3389/fnhum.2021.536288.

[6] Y. Ishikawa *et al.*, "Learning Cross-Modal Factors from Multimodal Physiological Signals for Emotion Recognition," in *PRICAI 2023: Trends in Artificial Intelligence*, vol. 14325, F. Liu, A. A. Sadanandan, D. N. Pham, P. Mursanto, and D. Lukose, Eds., in Lecture Notes in Computer Science, vol. 14325, Singapore: Springer Nature Singapore, 2024, pp. 438–450. doi: 10.1007/978-981-99-7019-3_40.

[7] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco California USA: ACM, Aug. 2016, pp. 785–794. doi: 10.1145/2939672.2939785.