Towards Incorporating Appraisal into Emotion Recognition: A Dynamic Architecture for Intensity Estimation from Physiological Signals

Robert Jenke[1], Angelika Peer[2]

***Abstract*— Current approaches to emotion recognition do not address the fact that emotions are dynamic processes. This work concerns itself with the development of a gray-box framework for dynamic emotion intensity estimation that can incorporate findings from appraisal models, specifically Scherer's Component Process Model. It is based on Dynamic Field Theory which allows the combination of theoretical knowledge with data-driven experimental approaches. Further, we conducted an exemplary user study applying the proposed model to estimate intensity of negative emotions from physiological signals. Results show significant improvements of the proposed model to common methodology and baselines. The flexible cognitive architecture opens a wide field of experiments and directions to deepen the understanding of emotion processes as a whole.**

I. Introduction

Current efforts in Human-Robot-Interaction (HRI) aim at finding ways to make interaction more natural. In this, knowledge of the user's emotional state is considered an important factor. Methods of automatic estimation of affective states from various modalities, including physiological signals, have therefore received much attention lately.

Recent work in emotion theory, e.g. Scherer's Component Process Model (CPM), points out the dynamic nature of emotion processes which therefore, "require a dynamic computational architecture" [1]. To date, however, most work on emotion recognition concerns itself with static prediction of emotion labels from a window of time series data using machine learning methods (i.e. black-box approach).

Our main research objective is to design a gray-box model for emotion recognition from physiological signals, which is capable of combining theoretical knowledge incorporated in the CPM with experimental data to train parameters the model. In this paper, we address the hitherto neglected dynamic evolvement of the affective state by proposing an architecture for emotion intensity estimation based on the Dynamic Field Theory (DFT) [2].

II. Model

In the CPM, the subjective feeling (i.e. affective state) is characterized by the emotion intensity I of an emotion quality ϑ at time t and can generally be written as $\mathrm{I}(\vartheta, t)$. The CPM provides detailed relations between the so-called stimulus evaluation checks (SECs) that happen in the appraisal process and their effects on physiology. For example, a novelty check can lead to an increase in skin conductance or the obstructiveness of an event changes the heart rate of a person.

[1]Chair of Automatic Control Engineering, Technische Universität München, Munich, Germany, www.lsr.ei.tum.de, E-mail: rj@tum.de

[2]Bristol Robotics Laboratory, University of the West of England, Bristol, UK, www.brl.ac.uk, E-mail: angelika.peer@brl.ac.uk

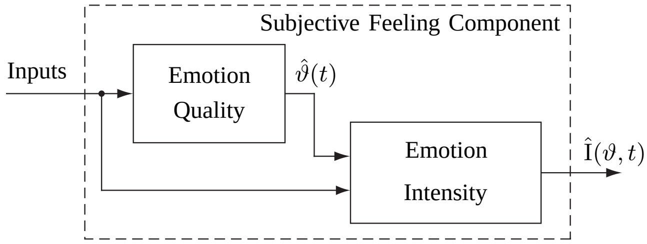

Fig. 1. The subjective feeling component is divided into consecutive estimation of emotion quality $\hat{\vartheta}(t)$ and emotion intensity $\hat{\mathrm{I}}(\vartheta, t)$.

Similar to Bailenson et al. [3], we separate estimation of emotion quality and intensity (see Fig. 1). We control the former by experimental design, i.e. we assume $\hat{\vartheta}(t)$ to be a known input to our model. The architecture of our dynamic model is based on DFT. These fields usually span over physical dimensions such as space or angle and model dynamic changes along this dimension. Fields are governed by differential equations and can represent functionalities like memory (for details, see [4]).

For our model, we define the field over the emotion quality ϑ as shown in Fig. 2. The core part of the model is the intensity layer $i(\vartheta, t)$ together with a memory layer $m(\vartheta, t)$, which model the changes in the subjective feeling, i.e. the output $\hat{\mathrm{I}}(\hat{\vartheta}, t)$. The second part are the input layers, where we use one layer for each prediction from the SECs provided by the CPM, e.g. $u(\vartheta, t)$ in Fig. 2. For example, a change in skin response would be an input layer.

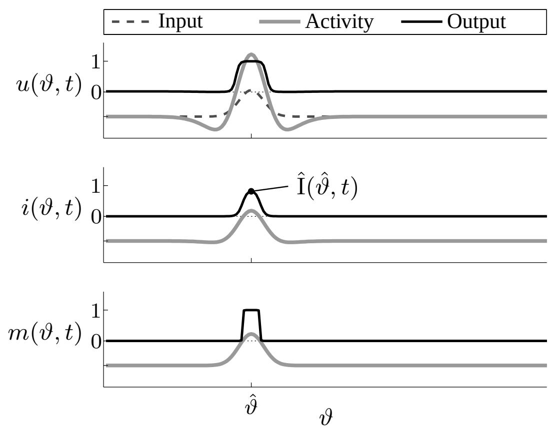

Fig. 2. Architecture of the proposed dynamic model: three-layer field spanned over the dimension of emotion quality ϑ.

III. Experimental Design

We control the emotion quality $\hat{\vartheta}$ in our experimental design by fixing it through choice of emotion induction. This results in a simplified dynamic model at one location of the fields, i.e. three neurons and their governing equations.

For the dataset, we recorded the galvanic skin response (GSR) of subjects. Additionally, we used a slider device interface to record the emotion intensity experienced by the subject. For emotion induction, we used standardized IAPS pictures of a fixed emotion quality, here, negative emotions [5]. After segmentation, we had collected 7 trials of $110\,\mathrm{s}$ recordings for each of three subjects.

The change in GSR is computed as a prediction of SECs and used as model input. The continuously recorded intensity measures of the slider served as ground truth. For training of the dynamic model, free parameters are determined by means of experimental data applying leave-one-out cross validation. In this, we minimize the error between the output of the dynamic model and the ground truth s.t. boundary conditions.

IV. Results

First, we compare the accuracy of our model with common static methods and baselines, i.e. linear regression and random regressors. We use the match of estimate with ground truth plus an acceptable error margin as accuracy measure. In summary, the dynamic model performs significantly better than common methodology and baselines. Limitations of the model become apparent for small error margins.

Secondly, capabilities and limitations of the model in its current version are examplified in Fig. 3. In the upper graph, we see the changes in GSR, which characterize the onset as well as the increase of intensity well. The memory layer (bottom graph) helps to stabilize the decay at an appropriate rate. However, limitations of the current model are apparent, as the third change in GSR should not have any impact on the intensity. This points towards the need to include additional input layers where appropriate interaction can avoid this behavior.

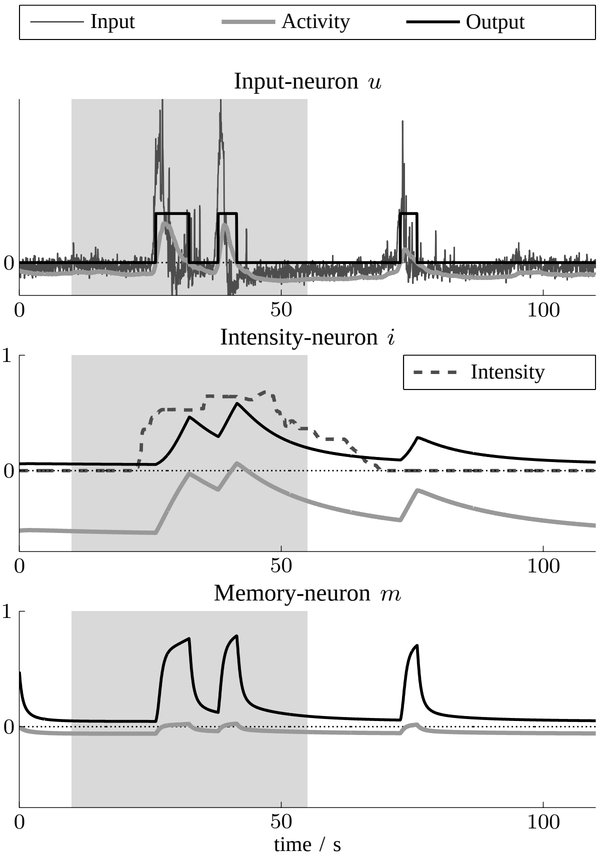

Fig. 3. Example of a single location of all layers over time.

V. Conclusion

For the first time, a dynamic gray-box model framework based on DFT has been proposed for emotion recognition, which allows to include theoretical knowledge into the model and learn free parameters from experimental results. We designed and carried out an exemplary study to estimate emotion intensity from physiological signals. In this, the dynamic model performed significantly better than baselines. We also identified current limitations and ways to improve the model. Future work includes several extension to the architecture as well as carrying out experiments to further evaluate the model.

References

[1] K. R. Scherer, "Emotions are emergent processes: they require a dynamic computational architecture." *Phil. Trans. of the Royal Society, Series B*, vol. 364, pp. 3459–74, Dec. 2009.

[2] G. Schöner, "Dynamical systems approaches to cognition," in *Cambridge Handbook of Computational Psychology*, R. Sun, Ed. Cambridge University Press, 2008, pp. 101–126.

[3] J. Bailenson, E. Pontikakis, I. Mauss, J. Gross, M. Jabon, C. Hutcherson, C. Nass, and O. John, "Real-time classification of evoked emotions using facial feature tracking and physiological responses," *Int.Journal of Human-Computer Studies*, vol. 66, no. 5, pp. 303–317, May 2008.

[4] Y. Sandamirskaya, "Dynamic neural fields as a step toward cognitive neuromorphic architectures." *Frontiers in Neuroscience*, vol. 7, pp. 1–13. art. 276, Jan. 2014.

[5] P. Lang, M. Bradley, and B. Cuthbert, "International Affective Picture System (IAPS): Affective ratings of pictures and instruction manual," University of Florida, Tech. Rep., 2005.