Neural models for the cross-species recognition of dynamic facial expressions

Michael Stettler, Nick Taubert, Ramona Siebert, Silvia Spadacenta, Peter Dicke, Peter Thier, Martin A. Giese

Dynamic facial expression recognition is an essential skill of primate communication. While the neural mechanisms to recognize static facial expressions has been extensively investigated, they remain largely unclear for dynamic facial expressions. We studied plausible neural encoding mechanisms, exploiting highly controlled and realistic stimulus sets generated by computer graphics, which are also used in electrophysiological experiments. METHODS: Combining mechanisms from physiologically plausible neural models for the recognition of dynamic bodies (Giese & Poggio, 2003), static faces (Giese & Leopold, 2005) and architectures from computer vision (Simonyan et al., 2014), we devised models for the recognition of dynamic facial expressions. The first model is *example-based,* and encodes dynamic expressions as temporal sequences of snapshots, exploiting a sequence-selective recurrent neural network. The second model exploits *norm-referenced* encoding, where face-space neurons are tuned to the differences between the actual stimulus frame and a reference face, the neutral facial expression. The dynamic expressions are recognized by differentiating the responses of these face-space neurons. We tested the models with high-quality human and monkey avatars, animated with motion capture data from both species, controlling expression by motion morphing. RESULTS: Both models recognize reliably dynamic facial expressions of humans and monkeys. However, the predicted behaviour of face-tuned neurons is very different for both models. The norm-referenced model shows a highly gradual, almost linear dependence of the neuron activity with the expressivity of the stimuli, while the neurons in the example-based model show very limited capability of generalization to expressions with varying strength. We also explored if the models explain the human capability of humans to recognize human expressions on monkey faces (Taubert et al. 2020). CONCLUSIONS: The two physiologically plausible models accomplish the recognition of dynamic faces and make distinguishable predictions for physiological experiments, where norm-referenced encoding might support transfer of expression recognition across different head shapes.