Probabilistic model for the online synthesis of stylized reactive movements in Virtual Reality

Nick Taubert[1,2], Junru Li[1,2], Dominik Endres[3], Andrea Christensen[2], Beatrice de Gelder[4], Martin A. Giese[2]

[1]Equal Contribution, [2]Section for Computational Sensomotorics, Dept. for Cognitive Neurology, Hertie Institute for Clinical Brain Research & Centre for Integrative Neurocience, University Clinic Tuebingen, [3]Theoretical Neuroscience, Department of Psychology, Philipps-University Marburg, Brain and Emotion Laboratory, [4]Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University

Sektion Theoretische Sensomotorik
Section Computational Sensomotorics

Motivation

Embodiment theories hypothesize that the perception of emotions from body movements involves an activation of brain structures that are involved in motor execution during social interaction [4, 5]. This predicts that, for identical visual stimulation, bodily emotions should be perceived as more expressive when the observers are involved in social motor behavior. We tested this hypothesis, exploiting advanced VR technology, requiring participants to judge the emotions of an avatar that reacted to their own motor behavior.

Basic Concept

Based on motion capture data from four human actors, we learned generative models for the body motion during emotional pair interactions, exploiting a framework based on Gaussian Process Latent Variable Models (GP-LVM) [1] and have been proposed as a powerful approach for high dimensional data modeling through dimensionality reduction. It has been shown that GP-LVMs are able to capture subtle emotional style changes and convey the information to human observers during reconstruction [3].

GP-LVMs are a probabilistic representation of dual PCA that map nonlinear a low-dimensional latent variable $\mathbf{x}$ on the data $\mathbf{y}$:

$$\mathbf{y} = f(\mathbf{x}) + \varepsilon, \quad f(\mathbf{x}) \sim GP(m_Y(\mathbf{x}), k_Y(\mathbf{x}, \mathbf{x}')),$$

where $f(\mathbf{x})$ is drawn from a Gaussian process with mean function $m_Y(\mathbf{x})$ and kernel function $k_Y(\mathbf{x}, \mathbf{x}')$. We assume a zero mean function $m_Y(\mathbf{x}) = 0$ and use a non-linear radial basis function (RBF) kernel [2] for a high dimensionality reduction and smooth trajectories in latent space. Furthermore, the variance term for ε can be absorbed into this kernel via the noise precision γ_3:

$$k_Y(\mathbf{x}, \mathbf{x}') = \gamma_1 \exp\left(-\frac{\gamma_2}{2}|\mathbf{x} - \mathbf{x}'|^2\right) + \gamma_3^{-1}\delta_{\mathbf{x},\mathbf{x}'},$$

where γ_1 is the output scale and γ_2 the inverse width of the RBF term. Let $\mathbf{K}_Y$ denote the $N \times N$ kernel covariance matrix, obtained by applying the kernel function to each pair of data points.

Motion Generation

Generation by nonlinear 2nd order AR model,

- style per emotion and actor encoded by latent variables $\mathbf{e}$ and $\mathbf{c}$,
- constraining manifold in latent space by parametrization using periodic functions in kernel based regression,
- spase approximation of kernal to make method real-time capable.

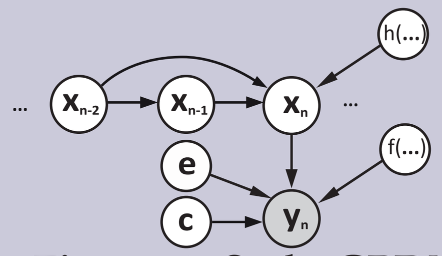


Figure 1: Style-GPDM

Preliminary Experiments

The first experiment investigated whether the Style-GPDM can synthesize motions with recognizable emotional styles.

perceived emotion	intended emotion: Anger	Neutral	Fear
Anger	**0.70769231**	0.06153846	0.12051282
Neutral	0.09230769	**0.89230769**	0.13076923
Fear	0.20000000	0.04615385	**0.74871795**

Table 1: Classification Results (N = 26)

In a second experiment participants had to rate the emotional style per actor, morphed in five steps. The resulting psychometric functions were used to normalize the emotional expressiveness levels between different actors.

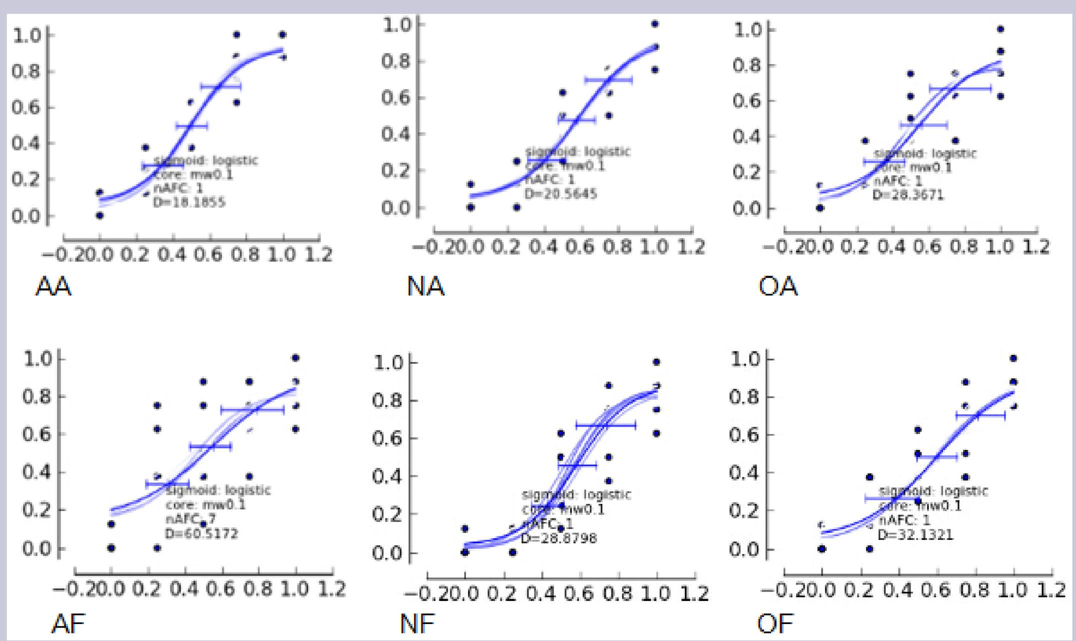
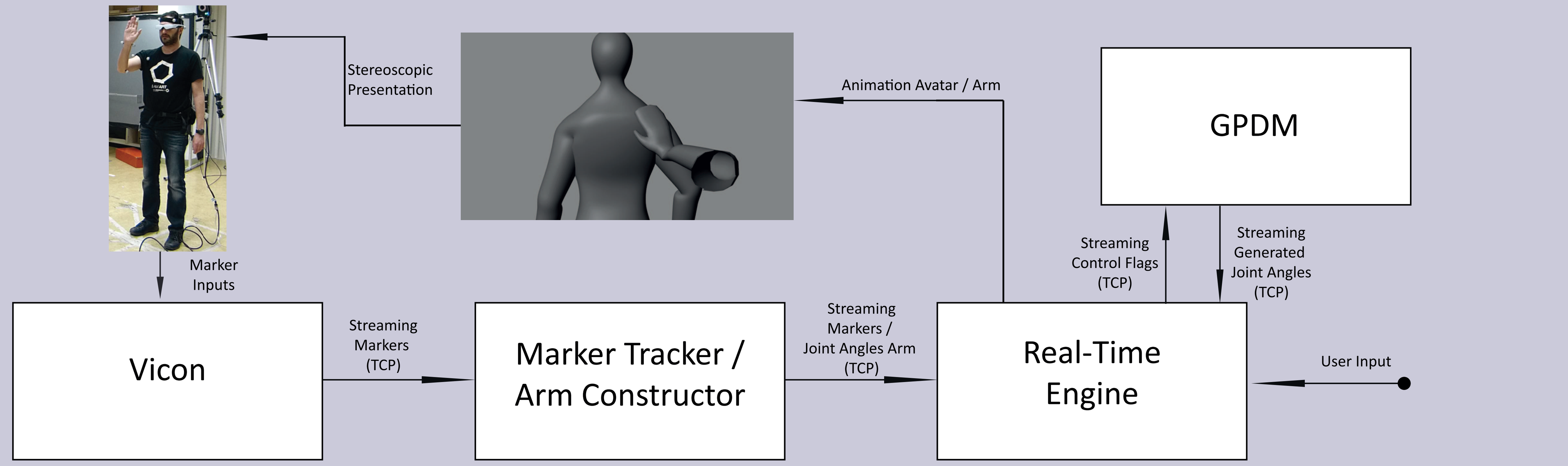


Figure 2: Perceived expressivity of morphs (top: angry, bottom: fearful) for different actors

Interactive System Layout



Task for Participants:

- Tip the avatar on its shoulder from behind.
- Rate the perceived emotion.
- Repeat the classification non-interactively (replaying the observed trajectories from previous trials).

Experimental Setup:

- After the tipping the avatar turned around in one of three emotional styles: $fearful, angry$ in five interpolation steps.
- Balanced design with two groups, one with open loop and one with closed loop first).
- One block for each emotional style ($fearful, angry$).
- Stimuli presented in random order.
- Experiment consisted of four blocks with 90 trials (open loop with training blocks first).

Results

Emotional expressiveness of the stimuli was rated higher when the participants initiate the emotional reaction of the avatar in the VR setup by their own behavior, as compared to pure observation ($\mathbf{F(1,17) = 8.701}$ and $\mathbf{p < 0.01, N = 18}$). This effect was particularly prominent for anger expressions.

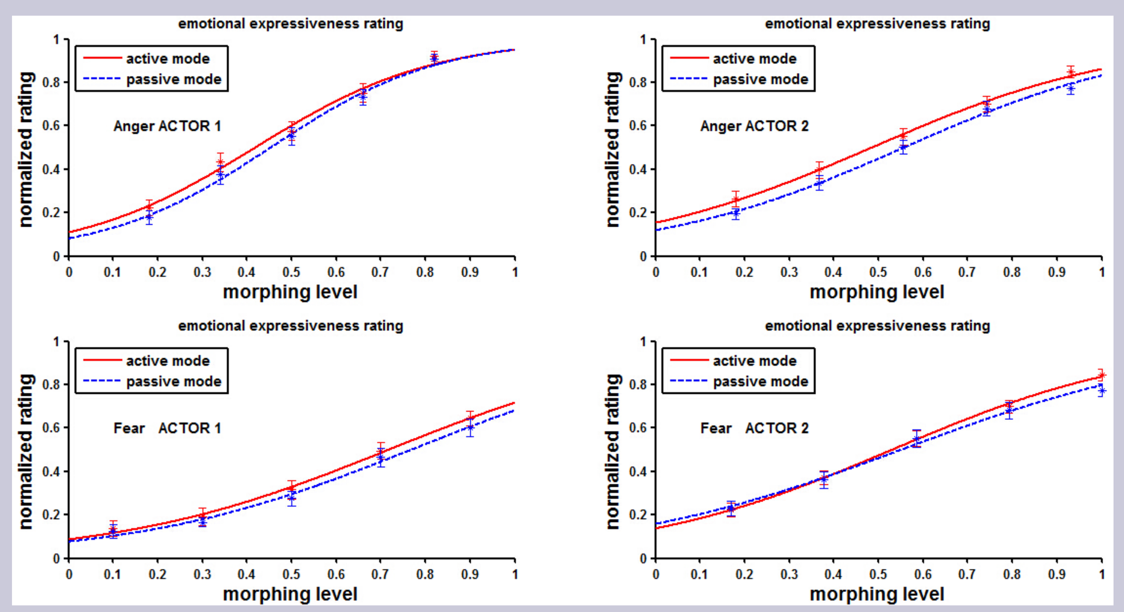


Figure 3: Psychometric functions per emotion and actor for active and passive mode.

Conclusion

Consistent with theories about embodied perception of emotion, the involvement in social motor tasks seems to increase perceived expressiveness of bodily emotions. For **future work** we will test the hypothesis with other emotions, e.g. $happy, sad$.

References

[1] N. D. Lawrence. Probabilistic non-linear principal component analysis with gaussian process latent variable models. *Journal of Machine Learning Research*, 6:1783–1816, 2005.

[2] Carl Edward Rasmussen and Christopher K. I. Williams. Gaussian processes for machine learning. *J. Am. Stat. Assoc.*, 103:429–429, 2008.

[3] Nick Taubert, Andrea Christensen, Dominik Endres, and Martin A. Giese. Online simulation of emotional interactive behaviors with hierarchical Gaussian process dynamical models. In *Proc. SAP'12*, pages 25–32, New York, New York, USA, 2012. ACM Press.

[4] Bruno Wicker, David I Perrett, Simon Baron-Cohen, and Jean Decety. Being the target of anotherâĂŹs emotion: a {PET} study. *Neuropsychologia*, 41(2):139 – 146, 2003. The cognitive neuroscience of social behavior.

[5] Daniel M. Wolpert, Kenji Doya, and Mitsuo Kawato. A unifying computational framework for motor control and social interaction. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 358(1431):593–602, March 2003.

Acknowledgements

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 604102 (HBP), Koroibot FP7-ICT-2013-10/ 611909, AMARSi- EC FP7-ICT-248311; DFG GI 305/4-1, DFG GZ: KA 1258/15-1; BMBF, FKZ: 01GQ1002A, FP7-PEOPLE-2011-ITN(Marie Curie): ABC PITN-GA-011-290011,CogIMon H2020 ICT-23-2014 /644727.