

Probing Dynamic Human Facial Action Recognition From The Other Side Of The Mean

Cristóbal Curio* Martin A. Giese† Martin Breidt* Mario Kleiner* Heinrich H. Bühlhoff*

*Max Planck Institute for Biological Cybernetics, Tübingen, Germany

†University of Wales Bangor, UK & ARL, Hertie Institute for Clinical Brain Research
University Clinic Tübingen, Germany



Animated peak expressions 'happy' (left) and 'disgusted' (right) next to corresponding 'anti'-expressions.

Abstract

Insights from human perception of moving faces have the potential to provide interesting insights for technical animation systems as well as in the neural encoding of facial expressions in the brain. We present a psychophysical experiment that explores high-level after-effects for dynamic facial expressions. We address specifically in how far such after-effects represent adaptation in neural representation for static vs. dynamic features of faces. High-level after-effects have been reported for the recognition of static faces [Webster and Maclin 1999; Leopold et al. 2001], and also for the perception of point-light walkers [Jordan et al. 2006; Troje et al. 2006]. After-effects were reflected by shifts in category boundaries between different facial expressions and between male and female walks. We report on a new after-effect in humans observing dynamic facial expressions that have been generated by a highly controllable dynamic morphable face model. As key element of our experiment, we created *dynamic 'anti-expressions'* in analogy to *static 'anti-faces'* [Leopold et al. 2001]. We tested the influence of dynamics and identity on expression-specific recognition performance after adaptation to *'anti-expressions'*. In addition, by a quantitative analysis of the optic flow patterns corresponding to the adaptation and test expressions we rule out that the observed changes reflect a simple low-level motion after-effect. Since we found no evidence for a critical role of temporal order of the stimulus frames we conclude that after-effects in dynamic faces might be dominated by adaptation to the form information in individual stimulus frames.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation; H.1.2 [Models And Principles]: User/Machine Systems—Human information processing

Keywords: human psychophysics, facial motion, facial animation

1 INTRODUCTION

Research on the neural encoding of faces in visual cortex, so far, has predominantly focused on the recognition of static faces [O'Toole et al. 2002]. In spite of the growing interest and high relevance of understanding facial expressions in emotional and social interaction only few results on the neural mechanisms of the processing of temporally changing dynamic faces exist. Specifically, almost no physiologically plausible neural theory of the recognition of dynamically changing faces has yet been developed. At the same time, present research in different research areas has started to investigate the statistical basis for the recognition of dynamic faces [Chang et al. 2004; Vlasic et al. 2005]. This makes the development of neural models in this area a timely topic in theoretical neuroscience. A deeper understanding of the mechanisms of dynamic face processing in biological systems might help to develop and improve technical systems for the analysis and synthesis of facial movements. It has been proven useful to encode faces in a multidimensional face-space [Valentine 1991]. Several recent studies have provided evidence about the neural representation of static faces and the properties that are reflected by continuous perceptual space models [Webster and Maclin 1999; Leopold et al. 2001; Loeffler et al. 2005; Leopold et al. 2006; Jiang et al. 2006]. A fundamental mechanism in perception research are after-effects that have been already reported for the recognition of neutral static faces [Webster and Maclin 1999; Leopold et al. 2001] and static facial expressions [Fox and Barton 2006]. Among other studies, in particular [Leopold et al. 2001] provides evidence for the neural representations of faces by using adaptation techniques and anti-faces, faces that are morphed to the opposite side of the mean in face space. The presentation of such static 'anti-faces' biases the

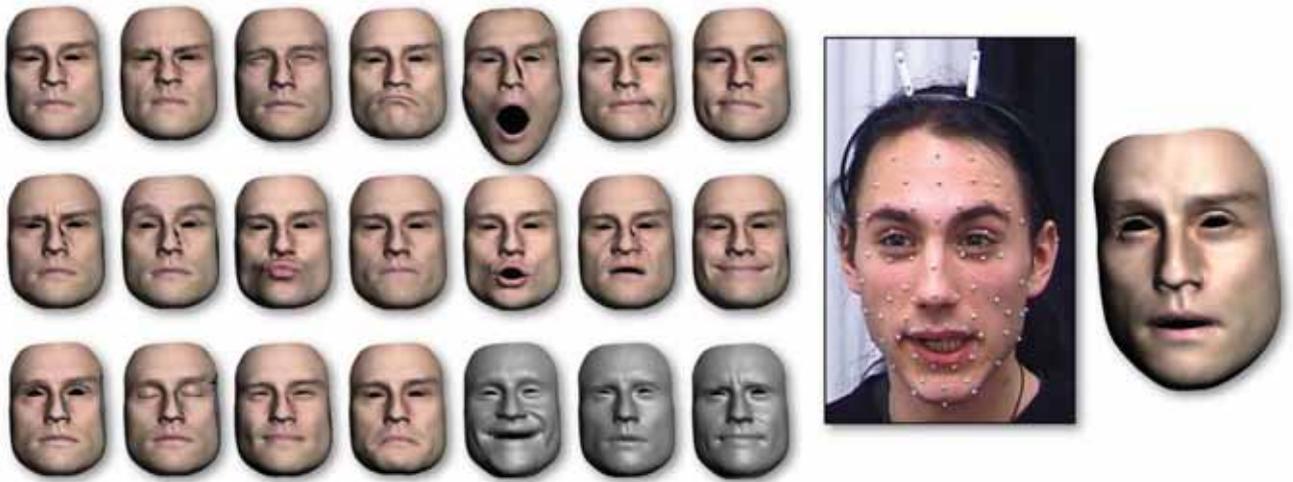


Figure 1: An Action Unit basis obtained from 3D scans and texture (left) and a Motion Capture system. Expressions can be retargetted frame-by-frame from Motion Capture to morph shapes by automatic decoding and application of Action Unit activation levels (right).

perception of neutral test faces temporarily towards the perception of specific identities or facial expressions. This indicates that the mean face may play an important role as a prototype in a perceptual space that reveals neural coding properties. Given the strong after-effects for static pictures of faces, it seems interesting to study whether after-effects exist also for dynamic facial expressions. In addition, it seems important to study how after-effects for dynamic and static faces are related. We try to answer some of these questions applying a novel technique for the synthesis of realistic and completely parameterized dynamic facial expressions. Section 3 gives an overview of the developed facial animation pipeline. Section 4 presents design of our adaption study based on our novel dynamic face space, and the main results are presented in Section 4.3. Finally, Section 6 discusses an analysis of the statistics of the optic flow patterns generated by the stimuli that rules out low-level motion after-effects as explanations of the results. Possible implications for neural models of the encoding of dynamic faces are discussed in Section 7.

2 RELATED WORK

Recent studies have tried to address the perceptual role of timing information in moving faces [Hill et al. 2005; Giese et al. 2002] based on animation technology. The neural representation of static face perception was investigated with an adaptation paradigm and high-level after-effects by [Leopold et al. 2001] based on a 3D face morph space [Blanz and Vetter 1999]. Recent studies have demonstrated high-level after-effects also for point-light walkers, resulting in shifts of perceived gender [Jordan et al. 2006; Troje et al. 2006]. Viewing of a gender-specific gait pattern for an extended period biased judgements for subsequent intermediate gaits toward the opposite gender. It seems unlikely that this adaptation effect on dynamic patterns can be explained by adaptation to local features of the stimuli. Instead, it likely results from adaptation of representations of the global motion of the figures. This suggests that high-level after-effects can be induced also in representations of highly complex dynamic patterns. Recent developments in animation technology provide a sound way to parameterize and explore a dynamic face space [Curio et al. 2006].

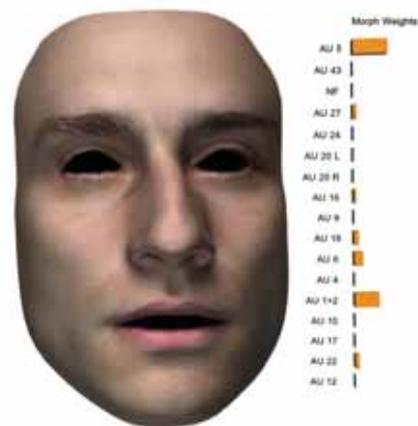


Figure 2: Sample of animation frame along with estimated weight vector \mathbf{w}^* denoting the activation level of employed Action Units.

3 MODELING OF DYNAMIC FACIAL EXPRESSION SPACE

For our perception experiment dynamic face stimuli were created using a three-dimensional dynamic morphable face model that is based on 3D scans of human heads [Curio et al. 2006]. For a better understanding of our new expression space in Section 3.3 we briefly give an overview of the two major components. The first one approximates face movements recorded with Motion Capture (MoCap) by a linear superposition of 3D Facial Action Units (AUs). Complex dynamic facial movements can be expressed by this low-dimensional basis (Figure 2) based on the Facial Action Unit Coding System (FACS) after [Ekman and Friesen 1978]. For each instance in time, the face can be characterized by a weight vector $\mathbf{w}(t)$ that specifies how much the individual Action Units contribute to the approximation of the present keyframe. The second step of the algorithm transfers this weight vector, or a paramet-

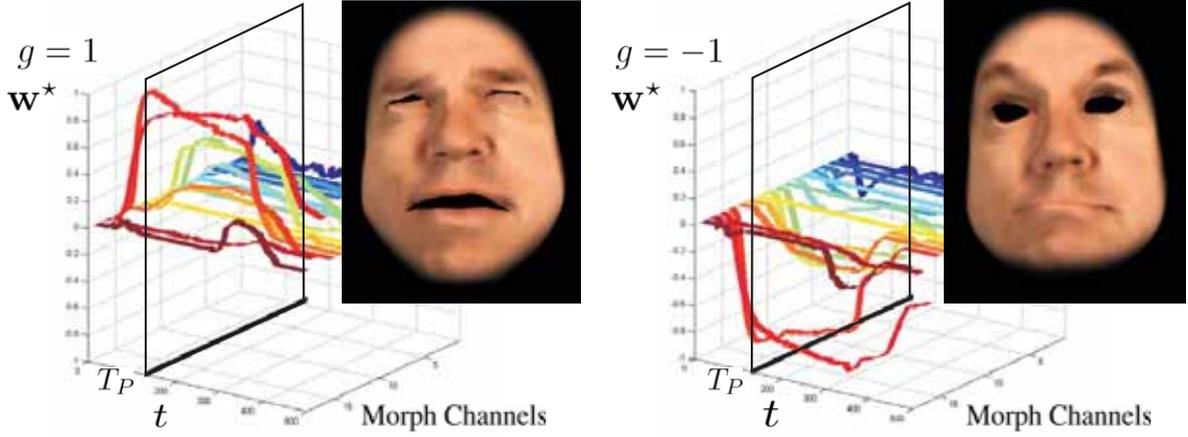


Figure 3: Estimated morph weight time courses $\mathbf{w}^*(t)$ of expression ‘disgust’ for original-expression (left) and anti-expression (right) and animated peak at $t = T_P$). Reduced and anti-expressions are obtained by variation of gain factor g in Eqn. 3.

rically modified version of it, $\mathbf{w}^*(t)$, to a dynamic morphable face model that creates photo-realistic animations. A more detailed description of the different analysis and rendering components of the animation algorithm follows.

3.1 PARAMETERIZATION OF FACIAL MOTION DATA

The facial animation is based on facial MoCap data recorded with a VICON 612 motion capture system, using $K = 69$ reflective markers with a temporal resolution of 120 Hz (cf. Figure 1, Right). The recorded facial expressions \mathcal{E} start in neutral state at $t = T_N$ passing through maximal peak expression $t = T_P$ and going back to neutral state at $t = T$. The data can be characterized by a matrix $\mathbf{M}_{\mathcal{E}} = [\mathbf{p}_{T_N}, \dots, \mathbf{p}_t, \dots, \mathbf{p}_{T_P}, \dots, \mathbf{p}_T]$ with 3D marker coordinates concatenated in the $(K \times 3)$ vectors \mathbf{p} for each time point. The previous matrix is thus a function of time $\mathbf{M}_{\mathcal{E}}(t) = \mathbf{p}_t$. For all further processing we selected one neutral expression vector at $\mathbf{p}_{\text{Ref}} = \mathbf{p}_{T_N}$. To obtain a low-dimensional parametrization of the dynamic facial movements the motion capture data was approximated by a linear combination of dynamically recorded Facial Action Units (AUs). By recording actors that have been asked to execute individual AUs separately, we obtained 3D motion capture data for $N = 17$ AUs that can be characterized by the matrices $\mathbf{M}_{AU,i}(t)$, $i \in 1, \dots, N$. After manually selecting the extreme keyframes, $\mathbf{M}_{AU,i}(T_P)$, and aligning them to the reference vectors \mathbf{p}_{Ref} the vector residuals define our AU basis system, indicated by $\mathcal{M}_{AU} = \{\mathcal{M}_{AU,i} \in \mathcal{R}^{K \times 3} | i \in 0, \dots, N\}$. A low dimensional approximation of any dynamic facial expression is obtained by, first, aligning $\mathbf{M}_{\mathcal{E}}(t)$ with \mathbf{p}_{Ref} for each time step, and then expressing the resulting matrix $\mathcal{M}_{\mathcal{E}}(t)$ by a linear superposition of the Action Units $\mathcal{M}_{AU,i}$ according to the equation

$$\mathcal{M}_{\mathcal{E}}(t) = \mathbf{p}_{\text{Ref}} + \sum_{i=1}^N w_{\mathcal{E},i}(t) \mathcal{M}_{AU,i}. \quad (1)$$

This approximation problem was solved by least square optimization with positivity constraint $w_i \geq 0$ applying Quadratic Programming. This procedure is similar to [Choe and Ko 2001], the difference being that we approximate expressions by real-world recorded MoCap AUs, and also use 3D scans of AUs with photographed textures for animation (s.b.).

3.2 MORPHABLE MODEL OF FACIAL EXPRESSIONS

A method for the synthesis of realistic 3D faces based on high-resolution 3D scans by morphing between shapes (\mathcal{S}_i) and textures (\mathcal{T}_i), that are in dense correspondence, has been proposed in [Blaiz and Vetter 1999]. Convex linear combination of shapes $\sum_i^m a_i \mathcal{S}_i$ and textures $\sum_i^m b_i \mathcal{T}_i$ were suitable for the photo-realistic simulation of static faces with new identities \mathcal{S}_{ID} , modeled by dense triangle surface meshes of L 3D vertex points. We extended this method by combining it with a method for the simulation of facial movements based on linear combinations of Action Units. A basis of Action Units was obtained from AU scans of actors (Figure 1). This AU based shape actuation basis is denoted by $\mathcal{S}_{AU} = \{\mathcal{S}_{AU,i} \in \mathcal{R}^{L \times 3} | i \in 1, \dots, N\}$, with dimension $L \gg K$ and $N = 17$, where each of the recorded AUs semantically corresponds to one element of the basis \mathcal{M}_{AU} . Individual surface scans were manually set into correspondence with 3D software *headus CySlice*, establishing correspondence between the actors’ neutral expression scan with the available 3D head mesh. With the solution $\mathbf{w}_{\mathcal{E}}(t)$ of (1) we are able to animate heads with different facial identities by frame-by-frame morphing between the elements of the shape actuation basis:

$$\mathcal{S}_{\mathcal{E}}(t) = \mathcal{S}_{ID,0} + \sum_{i=1}^N w_{\mathcal{E},i}(t) \mathcal{S}_{AU,i}. \quad (2)$$

3.3 GENERATION OF ‘REDUCED EXPRESSION’ AND ‘ANTI-EXPRESSION’

With this animation pipeline we were able to define a low-dimensional metric space of dynamic expressions. This space allows in particular the generation of reduced expressions and anti-expressions by rescaling of the weight vector:

$$\mathbf{w}_{\mathcal{E}}^*(t) = g \mathbf{w}_{\mathcal{E}}(t), \quad g \in \mathcal{R} \quad (3)$$

Reduced expressions correspond to morphing gains $0 < g < 1$, while negative morphing gains $g < 0$ define dynamic anti-expressions (cf. Figure 3). For our experiment we used two different shape identities (ID A & B) and one common texture \mathcal{T} . For comparison with static adaptation effects, we also derived ‘static anti-expressions’. These stimuli were given by one keyframe that

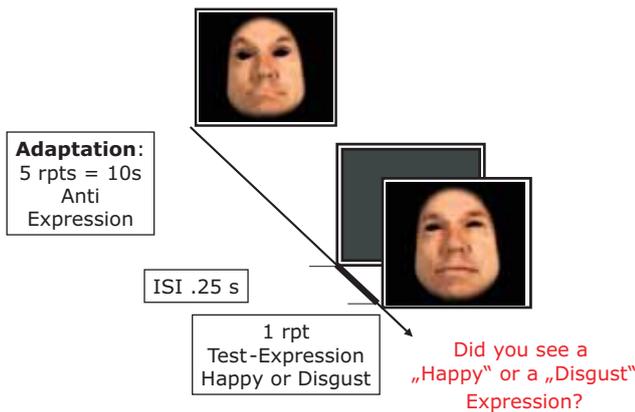


Figure 4: Adaptation study with dynamic expressions. In each trial subjects adapt to one anti-expression and are tested with randomly chosen test expression.

corresponded to the peak expression. Expression lengths were time aligned by subsampling, guaranteeing the same time intervals T_P between neutral and peak expressions.

3.4 NEUTRAL RIGID HEAD MOVEMENT

Rigid head movements can contribute essential information to the recognition of dynamic facial expressions. Since we wanted to study specifically the influence of non-rigid intrinsic facial movements, we eliminated the expression-specific rigid head movements from our stimuli. However, since we wanted to minimize the influence of possible low-level adaptation effects, we added a *neutral* rigid head movement that was given by one sinusoidal trajectory.

4 PSYCHOPHYSICAL ADAPTATION STUDY

We conducted a pre-study to perceptually calibrate our morphing space and two main experiments investigating the influence of dynamic vs. static adaptation stimuli, and the influence of temporal order on the adaptation effect. In addition, the transfer of adaptation to faces with different identity was tested. The animated facial expressions were rendered in real-time OpenGL Matlab Psychtoolbox-3 [Brainard 1997]. Subjects viewed the stimuli at 90cm distance and stimuli were perceived at a vertical visual angle of 12 degrees.

4.1 PRE-STUDY: CALIBRATION OF MORPHING LEVELS

The morphing parameter g in Equation 3 varies monotonically with the recognizability of the dynamic expressions. Yet, the relationship of this parameter and recognition rates might vary across subjects and expressions. We thus calibrated this space separately for each subject and expression by measuring a psychometric function that models the relationship between the morphing gain g and the recognition probability of that expression. In this calibration experiment, stimuli with a morphing gain $g \in \{.1 .15 .2 .25 .5\}$ were presented in random order and subjects had to respond in a 3AFC (Alternative Forced Choice) task whether they perceived one of the two expressions or a neutral expression ($g = 0$). The measured recognition probabilities $P(g)$ were fitted by the normalized cumulative probability function $P = \Phi(z) = .5(1 + \text{erf}(-z/\sqrt{2}))$ with $z = (g - \mu)/\sigma^2$. Expressions with reduced strength were de-

fined by the morphing weights that, according to this fitted function, corresponded to a recognition probability of $P_C = 0.2$.

4.2 EXPERIMENT 1: DYNAMIC VS. STATIC ADAPTORS AND INFLUENCE OF IDENTITY

This experiment tested the influence of dynamic vs. static adaptors (anti-expressions vs. extreme keyframes from the anti-expression) on the perception of expressions with reduced strength. Also, the identity of the adapting face was identical or different from the one of the test face. In addition, subjects were adapted with neutral dynamic expressions of one identity, in order to provide a baseline for expression-specific adaptation. Adaptation stimuli were presented for 10s, corresponding to 5 repetitions of the individual anti-expressions followed by a short ISI (400ms). For the static condition, at each repetition static anti-expressions were rendered at peak frame for the same time interval as the dynamic stimulus runs. Subsequently, the test stimulus cycle was presented, and naïve subjects ($N = 8$) had to decide in a 2AFC task whether they perceived a ‘happy’ or a ‘disgust’ expression (Figure 4). The experiment was organized as three factorial design (identity A/B, stimulus type Dynamic vs. Static and two expressions: ‘happy’ and ‘disgust’). We used two reduced test-expressions ‘happy’ and ‘disgust’ of identity A. This results in 16 different combinations of adaptation and test stimuli. Stimuli were presented in 12 blocks. Within each block the order of the 16 stimuli was randomized.

4.3 EXPERIMENT 2: FORWARD VS. REVERSE TEMPORAL ORDER

This experiment served mainly to test how the temporal order of the adapting stimulus affects the adaptation effects. Specifically, if the observed adaptation process would be based purely on the adaptation of individual keyframes no influence of the temporal order on the adaptation effects would be expected. The design of this experiment was identical with experiment 1. However, for half of the trials we replaced the static adaptation condition with dynamic anti-faces played in reverse temporal order. Also, we used different naïve subjects ($N = 9$).

Prediction: In general, based on the results on adaptation with static faces [Leopold et al. 2001] one might predict that adaptation with anti-expressions should increase the sensitivity of perception for compatible expressions with reduced strength. Conversely, one might expect that such adaptation with incompatible anti-expressions might decrease the sensitivity for faces with reduced expression strength (Figure 5).

5 RESULTS

Experiment 1: Results of Experiment 1 are shown in Figure 6 (left). All tested adaptation conditions resulted in an increase of the recognition of the matching expression compared to the baseline (adaptation with a neutral face). This increase was significant for the conditions with matching identity of adaptation and test face ($t > 3.4$; $p < 0.002$). For dynamic and static adapting faces with different identities this increase failed to reach significance. For non-matching expressions recognition rate was significantly reduced for all adaptation conditions ($t < -2.33$; $p < 0.05$). A three-way dependent measures ANOVA shows a significant influence of the Test Expression (‘happy’ vs. ‘disgust’) ($F(1, 91) = 48.3$; $p < 0.001$) and of the factor matching vs. non-matching expression (adapt-test) ($F(1, 91) = 94$; $p < 0.001$). The third factor, adaptation stimulus (dynamic vs. static face with same or different identity), did not have a significant influence, and all interactions were non-significant. This result shows that the observed

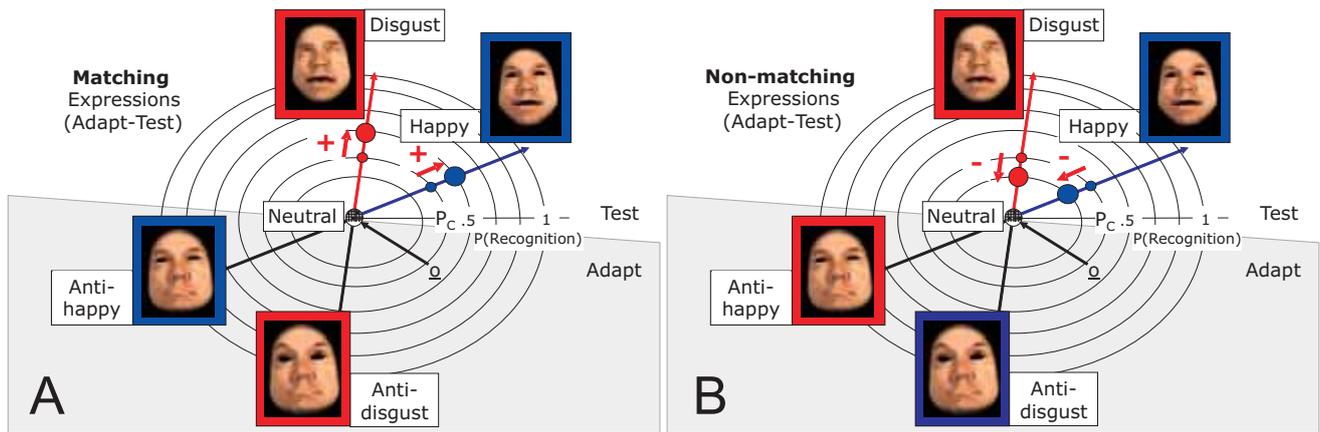


Figure 5: Predictions for recognition performance after adaptation. Increase for matching expressions (Adapt-Test) that are on the same expression morph axis (A) and a decrease for non-matching expressions (Adapt-Test) that are on different morph axes (B). The elliptic lines indicate contours of equal recognition probability $P(\text{Recognition})$.

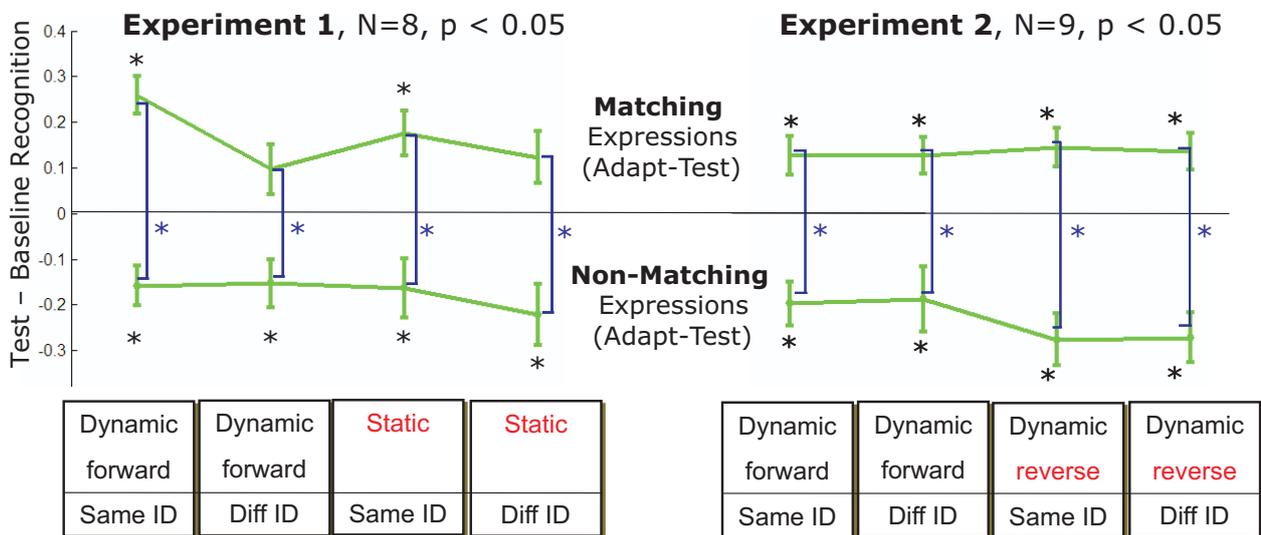


Figure 6: Results of recognition rates after testing expressions relative to baseline recognition performance for both experiments.

after-effect shows a clear selectivity for the tested expression. A more detailed analysis of the responses for the matching expressions only reveals a significant influence of the Test Expression $F(1, 45) = 26; p < 0.001$ and a marginally significant influence of the Adaptation stimulus $F(1, 45) = 2.8; p < 0.062$, where a post-hoc comparison shows that adaptation with a dynamic anti-expression and same identity induced significantly higher increases of the recognition rates than adaptation with static or dynamic anti-expressions with different identity ($p < 0.04$). This indicates that adaptation effects for matching conditions were particularly strong for matching facial identity.

Experiment 2: Results of Experiment 2 are shown in Figure 6 (right). Again, for all tested conditions we observed significant adaptation effects with a significant increase of the recognition of the matching expressions compared to the baseline and a significant decrease for non-matching expression conditions ($|t_{17}| > 2.5; p < 0.02$). A more detailed analysis using a 3-way dependent

measures ANOVA shows a significant influence of Matching vs. Non-matching adapting expression ($F(1, 98) = 21.3; p < 0.001$), and of the Test Expression ('happy' or 'disgust') ($F(1, 98) = 118; p < 0.001$). There was no significant difference between different types of adapting stimuli (forward vs. reverse temporal order, and matching vs. non-matching identity), i.e. no influence of the factor type of adaptor. In addition, all interactions between the three factors were non-significant. These results show again expression-specific adaptation. The comparison between the different conditions, and specifically between adaptors with forward and reverse sequential order did not reveal any significant differences. This implies that the observed adaptation effect is equally strong for stimuli presented in forward and reverse temporal order. This suggests that the representation of facial expressions might be dominated by the perception of the form of faces in individual keyframes. However, further experiments with stimuli suitable for separating form and motion cues are required to clarify this point.

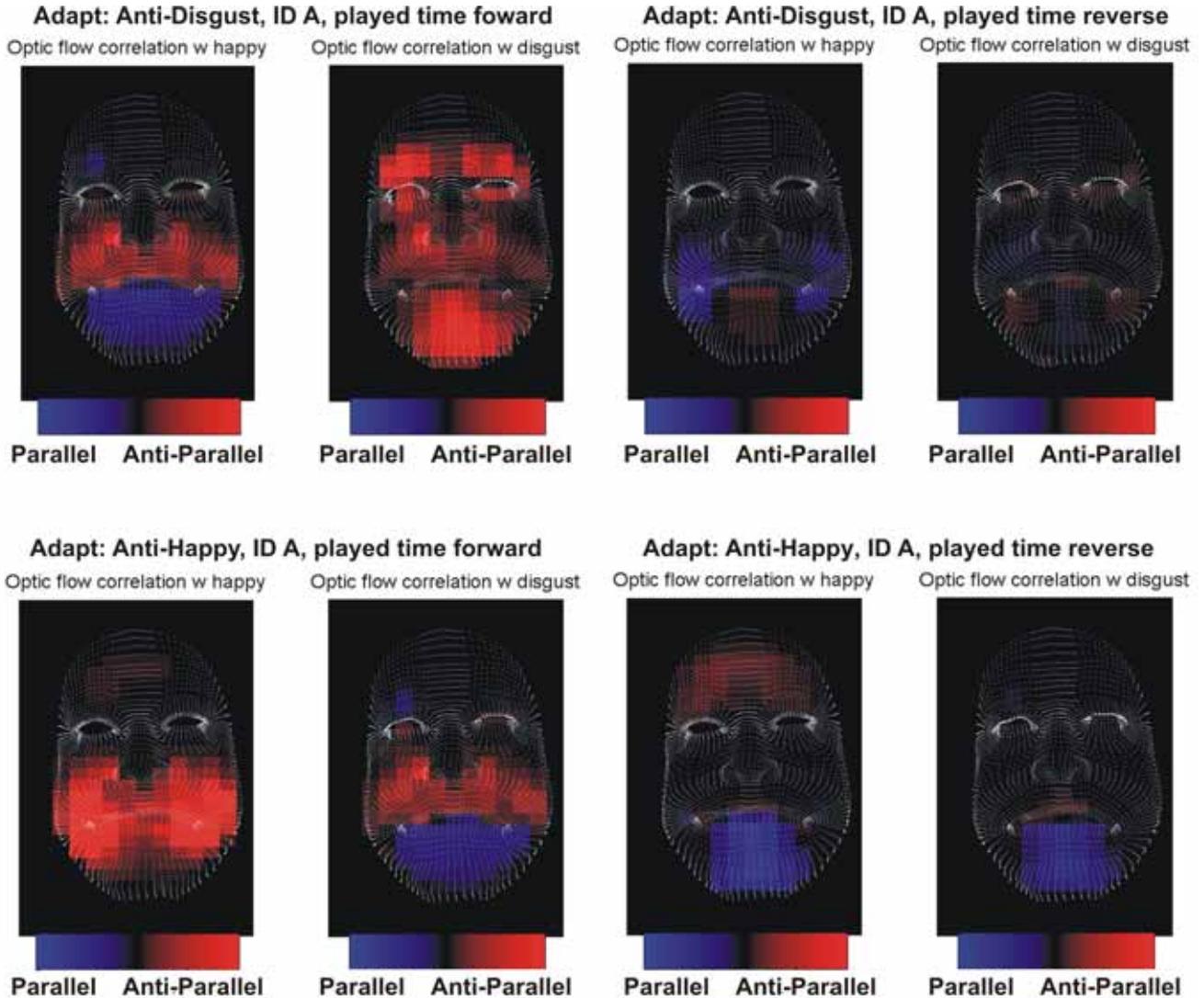


Figure 7: Comparison of the optic flow patterns between anti- and the two test expressions of Experiment 2. Color indicates the strength of correlation measure of the two optic flow fields between two consecutive frames (Eqn. 4). Results are rendered just for adaptation identity A.

6 A COMPUTATIONAL STUDY TO RULE OUT LOW-LEVEL MOTION ADAPTATION

Dynamic facial expressions specify dense optic flow patterns which produce low-level motion after-effects. One might thus raise the concern that the observed effects might mainly reflect adaptation of low-level motion processing mechanisms. The applied technique of stimulus generation (Section 3.2) allowed us to control for this possible confound. By computing the optic flow patterns generated by normal expressions and the anti-expressions played in forward and reversed order we were able to measure the size of possible low-level after-effects. For that, we determined the similarity of the local motion patterns generated by the adaptation and the test stimuli. Highly correlated local motion patterns (with motion in the same direction) should lead to reduced recognition of the test pattern, by adaptation of local motion detectors activated by the adaptor. However, anti-parallel motion patterns predict an increased recognition of the test pattern, since in this case the adapting motion pattern induces a kind of ‘waterfall illusion’.

6.1 Optic flow computation

To obtain a measure for the size of possible low-level motion after-effects we computed the correlations of the optic flow patterns generated by adaptation and test stimuli, aligning for the positions of adaptation and test faces. From the 3D face model all corresponding vertex positions in the image plane $\mathbf{x} = \{\mathbf{x}_i \in \mathcal{R}^2 | \mathbf{x}_1, \dots, \mathbf{x}_Q\}$ can be computed. From subsequent frames for each point in time the corresponding 2D optic flow vectors $\mathbf{v}_i = \{\mathbf{v}_1, \dots, \mathbf{v}_Q\}$ can be computed. As coarse estimate of the overlap of the local motion information, we sampled these optic flow fields with regular reference grid, \mathcal{G}_g (20×20), where the valid motion vector for each grid point was determined by nearest neighbor interpolation. Each grid point $\mathbf{x}_g \in \mathcal{G}_g$ was assigned the flow vector $\mathbf{v}(\mathbf{x}_g) = \mathbf{v}_i$ with $i = \operatorname{argmin} \|\mathbf{x} - \mathbf{x}_i\|$, inside the face regions of two consecutive snapshots of the animated expression. Outside the face regions the flow vector was set to zero. The correlation measure between the flow vector fields was defined as the sum over all scalar vector products of flow vectors. The summation was car-

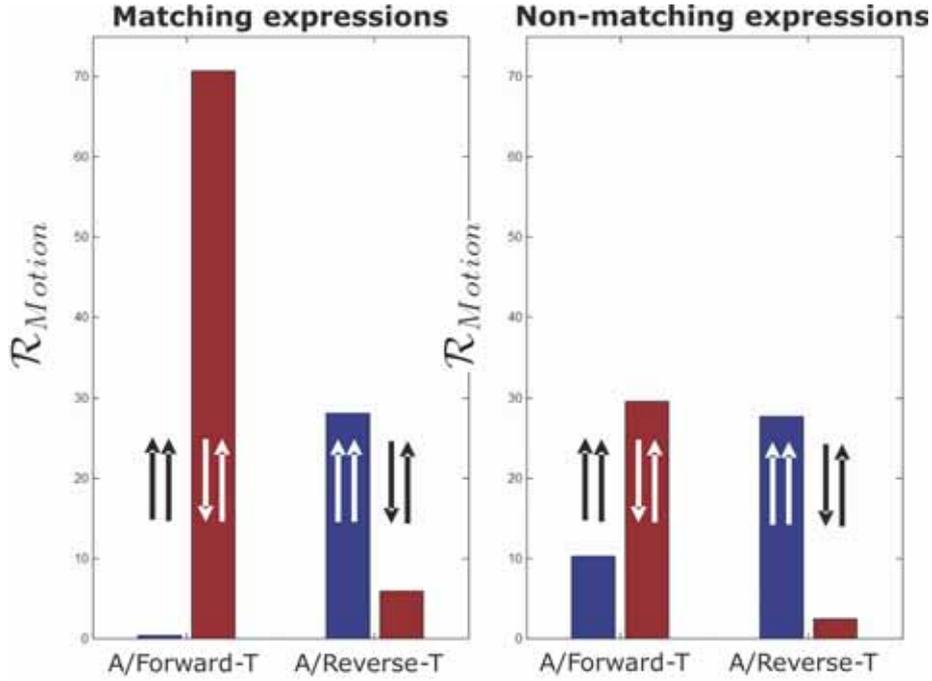


Figure 8: Correlation measures $\mathcal{R}_{Motion,\uparrow\uparrow}$ for parallel ($\uparrow\uparrow$) and $\mathcal{R}_{Motion,\downarrow\uparrow}$ for anti-parallel ($\downarrow\uparrow$) optic flow components between all dynamic adaptation conditions (forward / reverse, matching and non-matching expression) and the test expression.

ried out separately over products with positive and negative sign, defining a measure $\mathcal{R}_{Motion,\uparrow\uparrow}$ for the amount of ‘parallel’ and a measure $\mathcal{R}_{Motion,\downarrow\uparrow}$ for the amount of ‘anti-parallel’ optic flow of the (A)nti and (T)est stimuli. Signifying the corresponding flow vector fields by \mathbf{v}_A and \mathbf{v}_B , at each time t and grid point the scalar product was given by

$$\mathcal{C}(\mathbf{x}_g, t) = \langle \mathbf{v}_A(\mathbf{x}_g, t), \mathbf{v}_T(\mathbf{x}_g, t) \rangle. \quad (4)$$

The correlation measures were obtained by integrating Equation 4 over space and time

$$\mathcal{R}_{Motion,\uparrow\uparrow} = \frac{1}{Z} \int_{\mathcal{G}_g} \int_{t=T_N}^{T_P} [\mathcal{C}(\mathbf{x}_g, t)]_+ dt d\mathbf{x}_g \quad (5)$$

$$\mathcal{R}_{Motion,\downarrow\uparrow} = \frac{1}{Z} \int_{\mathcal{G}_g} \int_{t=T_N}^{T_P} [-\mathcal{C}(\mathbf{x}_g, t)]_+ dt d\mathbf{x}_g, \quad (6)$$

where Z denotes some normalization constant and $[\]_+$ the linear threshold function.

6.2 Computational results

The overall correlation measures $\mathcal{R}_{Motion,\uparrow\uparrow}$ and $\mathcal{R}_{Motion,\downarrow\uparrow}$ were computed for all conditions with dynamic adapting stimuli (cf. examples in Figure 7). The overall optic flow correlations can be depicted in Figure 8. Matching anti-expressions played in forward temporal order induce a strong anti-parallel optic flow field. Adaptation with the same expressions played in reverse order induces a strong parallel optic flow field between adaptor and test stimulus. If the observed after-effects were mainly based on low-level motion adaptation, this would predict that there should be opposite adaptation effects induced by the anti-face played in forward and backward temporal order. This clearly contradicts the experimental results that indicate no strong influence of the temporal order on the

observed after-effects for matching expressions. This rules out low level motion after-effects as major account for the observed after-effects in dynamic facial expressions. For non-matching adaptation and test expressions the optic flow analysis shows weaker correlations. However, again it shows dominant anti-parallel optic flow components for anti-expressions played in forward temporal order and dominant parallel flow components for expressions played in reverse order. Parallel optic flow components for adaptation and test predict a reduced recognition of the test stimulus. This effect would match the non-significant tendency in Experiment 2 that the recognition of non-matching expressions is somewhat reduced for adaptation stimuli played in reverse temporal compared to forward temporal order.

7 DISCUSSION

We presented the first study on high-level after-effects in the recognition of dynamic facial expressions. Our study was based on a novel computer graphics method for the synthesis of parametrically controlled photo-realistic facial expressions. The core of this method was the parametrization of dynamic expressions by Facial Action Units, defining an abstract low-dimensional morphing space of dynamic facial expressions. We found consistent after-effects for adaptation with dynamic anti-expressions resulting in an increased tendency to perceive matching test expressions, and a decreased tendency to perceive non-matching test expressions. This result is compatible with similar observation for adaptation effects with static faces and points to the existence of continuous perceptual spaces of dynamic facial expressions. In addition, these results show that the observed after-effects are highly expression-specific. Comparison between adapting stimuli played in forward and reverse temporal order did not reveal significant differences between the observed after-effects. This indicated that the involved after-effect might not rely on representations that are strongly sequence selective. A more detailed analysis shows also some hints

for a stronger adaptation for some expressions, if adaptation and test stimulus represent the same facial identity. This indicates a partial overlap of neural representations on a level of face cells selective for dynamic facial expressions and facial identity.

Future experiments will have to clarify the exact nature of dynamic integration in such representations of facial expressions and their relationship to such high-level after-effects. A detailed analysis of the optic flow patterns generated by the dynamic face stimuli rules out that the obtained after-effect is simply induced by standard low-level motion after-effects. However, more detailed studies will be required to study the influences of local motion versus shape on the perceptual responses of dynamic face stimuli.

For neural models of the perception of dynamic faces our study suggests the existence of continuous perceptual spaces for dynamic faces that might be implemented in similar ways as perceptual spaces for static face stimuli [Giese and Leopold 2005; Jiang et al. 2006; Burton et al. 1999]. However, we did not find strong evidence for a central relevance of sequence selectivity for such dynamic stimuli as opposed to the perception of body motion [Giese and Poggio 2003]. This raises the question about the influence of local motion information on the perception of dynamic facial expressions. Future experiments that minimize form cues in individual frames, at the same time providing consistent local motion information, might help to clarify the relative influences of form and motion in the perception of dynamic facial expressions.

Acknowledgements

This work was supported by EU-Projects BACS FP6-IST-027140 and COBOL (FP6-NEST), DFG Perceptual Graphics, HFSP.

References

- BLANZ, V., AND VETTER, T. 1999. A morphable model for the synthesis of 3d faces. *ACM SIGGRAPH*, 187–194.
- BRAINARD, D. 1997. The psychophysics toolbox. *Spatial Vision* 4, 433–436.
- BURTON, A., BRUCE, V., AND HANCOCK, P. 1999. From pixels to people: a model of familiar face recognition. *Cognitive Science* 23, 1–31.
- CHANG, Y., HU, H., AND TURK, M. 2004. Probabilistic expression analysis on manifolds. In *Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, vol. 2, 520–527.
- CHOE, B., AND KO, H.-S. 2001. Analysis and synthesis of facial expressions with hand-generated muscle actuation basis. In *Proceedings of Computer Animation*, IEEE Computer Society, IEEE, 12–19.
- CURIO, C., BREIDT, M., KLEINER, M., VUONG, Q., GIESE, M. A., AND BÜLTHOFF, H. H. 2006. Semantic 3d motion re-targeting for facial animation. In *Proceedings of the 3rd Symposium on Applied Perception in Graphics and Visualization*, ACM Press, ACM, 77–84.
- EKMAN, P., AND FRIESEN, W. 1978. Facial action coding system: A technique for the measurement of facial movement. *Consulting Psychologists Press*.
- FOX, C., AND BARTON, J. 2006. What is adapted in face adaptation? the neural representations of expression in the human visual system. *Brain Research* 1127, 80–89.
- GIESE, M., AND LEOPOLD, D. A. 2005. Physiologically inspired neural model for the encoding of face spaces. *Neurocomputing* 65–66 (June), 93–101.
- GIESE, M., AND POGGIO, T. 2003. Neural mechanisms for the recognition of biological movements and action. *Nature Review Neuroscience* 4, 179–192.
- GIESE, M. A., KNAPPMAYER, B., AND BÜLTHOFF, H. H. 2002. Automatic synthesis of sequences of human movements by linear combination of learned example patterns. In *Proceedings of the Second International Workshop on Biologically Motivated Computer Vision*, Springer-Verlag, London, UK, 538–547.
- HILL, H., TROJE, N., AND JOHNSTON, A. 2005. Range- and domain-specific exaggeration of facial speech. *Journal of Vision* 5, 10 (12), 793–807.
- JIANG, X., ROSEN, E., ZEFFIRO, T., VANMETER, J., BLANZ, V., AND RIESENHUBER, M. 2006. Evaluation of a shape-based model of human face discrimination using fmri and behavioral techniques. *Neuron* 50, 1, 159–72.
- JORDAN, H., FALLAH, M., AND STONER, G. 2006. Adaptation of gender derived from biological motion. *Nature Neuroscience* 9, 738–739.
- LEOPOLD, D., O'TOOLE, A., VETTER, T., AND BLANZ, V. 2001. Prototype-referenced shape encoding revealed by high-level after effects. *Nature Neuroscience* 4, 89–94.
- LEOPOLD, D., BONDAR, I., AND GIESE, M. 2006. Norm-based face encoding by single neurons in the monkey inferotemporal cortex. *Nature* 442, 7102 (July), 572–575.
- LOEFFLER, G., YOURGANOV, G., WILKINSON, F., AND WILSON, H. R. R. 2005. fmri evidence for the neural representation of faces. *Nature Neuroscience* (September).
- O'TOOLE, A., ROARK, D., AND ABDI, H. 2002. Recognizing moving faces: a psychological and neural synthesis. *Trends in Cognitive Science* 6, 6, 261–266.
- TROJE, N., SADR, J., GEYER, H., AND NAKAYAMA, K. 2006. Adaptation aftereffects in the perception of gender from biological motion. *Journal of Vision* 6, 8, 850–857.
- VALENTINE, T. 1991. A unified account of the effects of distinctiveness, inversion and race in face recognition. *Quarterly J. Experimental Psychology* 43, 161204.
- VLASIC, D., BRAND, M., PFISTER, H., AND POPOVIĆ, J. 2005. Face transfer with multilinear models. *ACM Trans. Graph.* 24, 3, 426–433.
- WEBSTER, M., AND MACLIN, O. 1999. Figural after-effects in the perception of faces. *Psychonomic Bulletin and Review* 6, 647–653.