

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/224088194>

Pleasure-arousal-dominance driven facial expression simulation

Conference Paper · October 2009
DOI: 10.1109/ACII.2009.5349579 · Source: IEEE Xplore

CITATIONS
18

READS
510

4 authors, including:



Karl Grammer
University of Vienna
169 PUBLICATIONS 7,672 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Caddy - cognitive autonomous diving buddy [View project](#)



2D:4D research [View project](#)

Pleasure-Arousal-Dominance Driven Facial Expression Simulation

Hana Boukricha, Ipke Wachsmuth
Faculty of Technology
Bielefeld University
33594 Bielefeld - Germany
{hboukric,ipke}@techfak.uni-bielefeld.de

Andrea Hofstätter, Karl Grammer
Ludwig-Boltzmann-Institute for Urban Ethology
Althanstrasse 14
A-1090 Vienna - Austria
andrea.hofstaetter@gmx.at, karl.grammer@univie.ac.at

Abstract

Expressing and recognizing affective states with respect to facial expressions is an important aspect in perceiving virtual humans as more natural and believable. Based on the results of an empirical study a system for simulating emotional facial expressions for a virtual human has been evolved. This system consists of two parts: (1) a control architecture for simulating emotional facial expressions with respect to Pleasure, Arousal, and Dominance (PAD) values, (2) an expressive output component for animating the virtual human's facial muscle actions called Action Units (AUs), modeled following the Facial Action Coding System (FACS). A large face repertoire of about 6000 faces arranged in PAD-space with respect to two dominance values (dominant vs. submissive) is obtained as a result of the empirical study. Using the face repertoire an approach towards realizing facial mimicry for a virtual human based on backward mapping AUs displaying an emotional facial expression on PAD-values is outlined. A preliminary evaluation of this first approach is realized with AUs corresponding to the basic emotions Happy and Angry.

1. Introduction

In his book *The Expression of the Emotions in Man and Animals* Darwin focuses on a more detailed and first scientific description of the meaning of different facial expressions as well as the facial muscles accompanying them. He also underlines the specific and functional role of facial expressions in expressing and communicating emotions [13]. Thus facial expressions play an important role in social interactions since detecting and understanding the facial expressions displayed by others allow an access to their intentional and affective states.

Nowadays the eventuality of being confronted with virtual characters embedded in computer related applications such as teaching or therapy applications [8], interactive

museum guide applications [23], and movie-video applications, is increasing. Therefore features of human face-to-face interactions should be applied when designing human-computer interfaces, *e.g.*, features underlying kinds of facial displays which play an essential role as a nonverbal communication channel [7]. Facial expressions are crucial not only in expressing and communicating emotions but also in mimicking the facial expressions of others. In social behavior mimicry has a necessary role in contributing to build bondings between humans. Mimicry acts as a 'social glue that binds humans together' since it contributes empathy, liking, rapport, and affiliation [9]. Bavelas *et al.* [2] argue for the role of mimicry as a communicative function in social interaction:

By immediately displaying a reaction appropriate to the other's situation (*e.g.*, a wince for the other's pain), the observer conveys, precisely and eloquently, both awareness of and involvement with the other's situation. (p. 278)

In human-computer interaction Brave *et al.* [4] and Prendinger *et al.* [32] found that agents showing involvement with their partner's situation through behaving empathically are judged by humans as more likeable, trustworthy and caring. In our work the definition of mimicry as empathy arousing mode introduced by Hoffman [21] is followed. He defines mimicry as the process involving the imitation of another's facial expression, voice, and posture, which triggers an afferent feedback eliciting the same feelings in oneself as those of the others.

In this paper the methods and results of an empirical study consisting of rating randomly generated facial expressions of the virtual human Emma, which is the female 'counterpart' to the virtual human Max [24] as well as an approach towards realizing facial mimicry using the face repertoire resulting from the empirical study are presented. In the next section previous works on generating a control architecture for simulating a virtual human's facial expressions as well as extending a virtual human's or robot's be-

havior to mimicking human's facial or multimodal expressions are outlined. Description of the used methods and results of the empirical study, the modeling of the virtual human Emma's face inline with Ekman *et al.* FACS [16], and the linkage between the obtained faces and the virtual human Emma's emotion dynamics are topics of Section 3. In Section 4 a first approach towards realizing facial mimicry following the definition of mimicry introduced by Hoffman is presented. That is, based on the face repertoire resulting from the empirical study, we are mainly interested in exploring how the changes in the facial musculature of the virtual human Emma while imitating a facial expression can induce changes in her emotional state. Finally a summary of the main conclusions and an outlook to future work are given.

2. Related Work

There are various attempts to create a control architecture for animating a virtual human's facial features. Lee *et al.* [26] proposed a facial expression animation system based on nonlinear deformations provided by learning a decomposable generative model. Stylized facial expressions with expressiveness control are synthesized and used to extract MPEG-4 facial animation parameters (FAPs) to animate a synthetic agent's face. Therefore it is not clear how the generated faces are linked to a dynamic control architecture that can be embedded into a virtual human's behavior system. Stoiber *et al.* [34] generate a control interface for animating synthetic faces by analyzing physical deformations of a human face displaying emotions. This is based on detecting principal geometrical and textural variation patterns and their reorganisation in a low-dimensional space. The resulting control space is very similar to Plutchik's emotion wheel [31].

The above described models all use image processing and tracking algorithms for animating synthetic faces. In this paper an approach based on empirical statistical and reverse engineering methods to generate a control architecture for a virtual human's emotional facial expressions, as introduced by [20], is applied. Following the theoretical emotion model of [33] a 3 dimensional space of pleasure, arousal, and dominance values is chosen as the facial expressions' control space. Theoretical models of emotions, such as categorical emotion models, are based on classifying facial expressions onto different simple or mixed basic emotions, *e.g.*, the six basic emotions as a result of a cross-cultural study by Ekman [15]. This discrete approach suffers from the fact that each basic emotion can be linked to more than one facial expression, thus it is unclear on which specific facial expression the result of an appraisal should be mapped. Dimensional emotion models indicate that emotional facial expressions can be collocated in a 2 or 3 dimensional space, *e.g.*, the pleasure, arousal, and dominance

space of semantic affective dimensions [33]. Compared to the categorical emotion model, the dimensional model is more convenient when dealing with research on facial expressions since blended facial expressions as well as facial expressions' intensities can be better represented.

With regard to mimicry there are various attempts in extending an agent's or robot's behavior to mimicking human's facial or multimodal expressions. The works of Caridakis *et al.* [6] and Courgeon *et al.* [12] consist of perceiving, interpreting, planning, and then animating the multimodal expression of the human. In [6] video recorded human's facial expressions and gestures are processed and analyzed. From an expression recognition module, Facial Animation Parameters (FAPs) are derived and expressed by the agent's face (the gesture's symbolic name is not being derived from the expression recognition module thus the gesture is manually communicated to the agent). Five expressivity parameters related to the movement's spatial volume, speed, energy, fluidity, and repetitivity are extracted from analyzing the image data and used to affect the quality of the agent's expressive behavior. In [12] from user's action on a 3D device (Joystick), a modulated target in PAD-space is computed and integrated with the output of a facial expression recognition module. The facial expression mirrored by the agent correspond to a blend of emotions derived from a modulated target in PAD-space and from combining facial expression recognition rates of seven basic emotions. Breazeal *et al.* [5] primarily concentrate on the imitation task related to mimicry. They explore how imitation as a social learning and teaching process contributes to building socially intelligent robots. The robot identifies one of the basic emotions as emotion related to the imitated facial expression and uses this information to link new facial expressions with emotion labels.

A common characteristic among these works is that they mainly investigate the function of how the agent or robot is better enabled to learn reproducing or mirroring humans' facial expressions. In contrast, in our work we aim at principally exploring how the changes in the facial musculature of the agent while imitating a facial expression can induce changes in its emotional state. That is, based on the rich face repertoire provided by the empirical study, we aim at developing a system of backward mapping AUs displaying an emotional facial expression on PAD-values since the intensity of an emotion as well as comparing different emotions is better measured by real numbers.

3. Pleasure-Arousal-Dominance Driven Facial Expression Simulation

In this section the expressive output component for animating the virtual human's facial actions modeled following FACS [16] and the control architecture for simulating emo-

tional facial expressions with PAD-values are presented.

3.1. Implementing the Facial Action Coding System in a Virtual Human's Face

Since every facial movement emerges from contracting different facial muscles, Ekman and Friesen [14] argue that a comprehensive expression coding system could be obtained by discovering the impact of muscle contraction on changing the visible appearance of the face. Thus based on an anatomical analysis of the facial behavior they introduced a method called Facial Action Coding System (FACS) to describe any facial movement with anatomically based Action Units (AUs). AUs are the core elements of FACS describing the contraction of single or multiple facial muscles (46 AUs) as well as eye and head movements (12AUs).

FACS is the most commonly used expression coding system in the behavioral sciences since it allows human observers measuring facial behavior in terms of visually observable facial muscle actions. Furthermore, new approaches to automated facial expression analysis attempt at recognizing AUs from a human face since detection of AUs allows a more flexible and versatile interpretation of facial expressions. That is, the interpretation is not restricted to recognizing the emotional states related to a facial expression, but also the related mental cognitive states can be recognized [1] [17] [28] [35].

Hence the virtual human Emma's face is modeled following FACS [16]. The implementation of AUs in Emma's face was performed at the *Department of Anthropology at the University of Vienna* with the help of experienced FACS coders [22]. Emma's face replicates 44 AUs including 34 AUs representing 9 upper face units and 25 lower face units. The remaining AUs represent head and eye units. The AUs are implemented directly on the head mesh at their maximum contraction. The appearance of Emma's face (see Figure 1) was designed by using front and side photographs of a real woman's face.

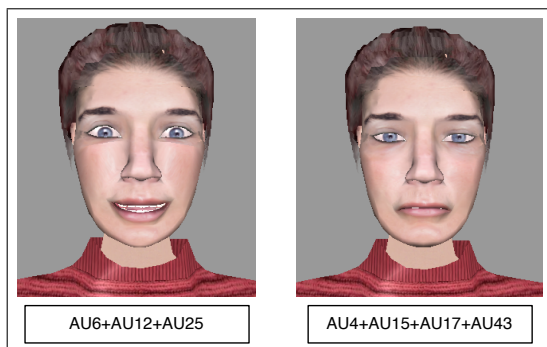


Figure 1. Emma's face with two example facial expressions: Happy and Sad.

3.2. Reconstructing Facial Expressions in Pleasure-Arousal-Dominance Space

The following empirical study [22] consists of human participants rating randomly generated facial expressions of the virtual human Emma with the bipolar adjectives from the "Semantic Differential Measures of Emotional State or Characteristic (Trait) Emotions" [29] (translated into German). As result a face repertoire of about 6000 faces arranged in PAD-space with respect to two dominance values (dominant vs. submissive) was obtained. In the following sections the methods for generating the face repertoire are reported.

3.2.1 Empirical Study

The empirical study took place at the *Biozentrum of the University of Vienna* in Austria. A total of 3517 randomly generated facial expressions were rated by 353 adults ranging between 18 and 65 years of age. 1418 (40.3%) of the faces were rated by 142 male subjects and 2099 (59.7%) of the faces were rated by 211 female subjects. Each of the subjects rated respectively 10 facial expressions with a German translation of bipolar adjectives from the "Semantic Differential Measures of Emotional State or Characteristic (Trait) Emotions" [29]. Each facial expression was rated with 18 bipolar adjectives on a 1 to 7 Likert-Scale. Following [30] each group of 6 bipolar adjectives is used to represent one of the dimensions of pleasure, arousal, or dominance.

A face randomizer program allowed randomly altering the intensities of the activations of single AUs. For generating random facial expressions the following rule was applied: Among 44 AUs chose randomly 10 AUs with random intensities between 0% and 50% from overall intensity. By this rule the generation of unnatural exaggerated faces is avoided. Furthermore constraint-based blending rules following Wojdel *et al.* [36] were applied in order to avoid generating anatomically impossible faces. These rules include dominance, exclusion, and opposition rules first introduced by Ekman and Friesen in [14] and implemented later by Wojdel *et al.* [36]. The faces were visualized and presented to subjects in POSER 6 (Curious Labs, Santa Cruz, CA). As reference face the neutral face was blended as transition between two presented faces.

3.2.2 Methods and Results

A principal component analysis with varimax-rotation was conducted on the rating adjectives ($N = 3482$) and yielded as expected a three factor solution. The first factor is the dominance factor, which explains 25.0% of variance, the second factor is the pleasure factor, which explains 22.5% of variance, and the third factor is arousal with 19.2% of variance. Further analysis of the data was restricted to con-

sidering these three factors and their correlation with AUs' activation. A multiple multivariate linear regression was performed on all AUs with each of the three factors. The multiple linear regression shows that most of the AUs (26 AUs) correlate ($r = 0.505$, $N = 3482$, $p < 0.01$) with either negative or positive pleasure. AU25, AU12, AU13, AU14, AU34, AU55, AU24, AU26, AU22, AU30 left, AU27, AU51 and AU18 correlate with positive pleasure. AU4, AU64, AU15, AU43, AU1, AU10, AU63, AU17, AU54, AU9 and AU33 correlate with negative pleasure. 18 AUs correlate ($r = 0.387$, $N = 3482$, $p < 0.01$) with either negative or positive dominance. AU4, AU9, AU43, AU2, AU27, AU31, AU53, AU24, AU54, AU38, AU55, AU23 and AU29 correlate with positive dominance. AU1, AU61 and AU62 correlate with negative dominance. 10 AUs correlate ($r = 0.557$, $N = 3482$, $p < 0.01$) with either high or low arousal. AU64, AU27, AU25, AU2, AU61, AU62, AU5, AU16 and AU31 correlate with high arousal. AU43, AU63, AU56, AU55, AU18, AU34 and AU54 correlate with low arousal. Furthermore a two dimensional regression analysis of AU activation with two of the three factors was performed with a LOESS-regression (cf. [10], [11]). As result 3 dimensional non linear regression planes for each AU in the pleasure-arousal, pleasure-dominance, and arousal-dominance space were obtained. Furthermore regression of pleasure and arousal was calculated for those faces that had been rated positive concerning dominance and those that had been rated as showing negative dominance (*e.g.*, see Figure 2). Since the PA-space is a continuous space, the AU intensity values of new faces can be interpolated in the regression planes, *e.g.*, using splines. By combining all planes for all AUs' activations the facial expression corresponding to each point in the space can be recomposed and a facial expressions' control space is reconstructed (see Figure 3).

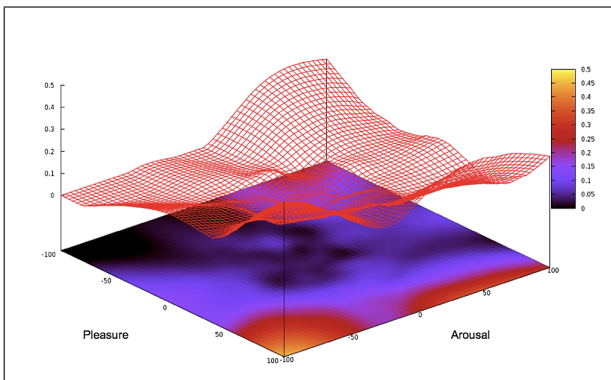


Figure 2. The regression plane of AU12 in PA-space for high dominance. The maximum intensity of AU12 correlates with high pleasure.

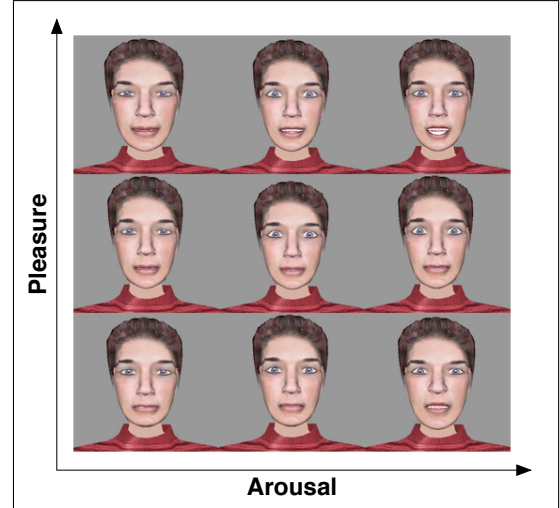


Figure 3. The resulting face repertoire in PA-control space for high dominance. In the middle is the neutral face. At the corners the faces corresponding resp. to high-low and negative-positive arousal and pleasure are shown (head and eye units are not included here).

3.2.3 Discussion

The visualization of the resulting face repertoire shows quite accurate facial expressions (*e.g.*, see Figure 3) since AUs occurring in expressions of positive emotion characterize positive pleasure, *e.g.*, AU12 and AU25. Negative pleasure is expressed essentially through AU4, AU15, AU1, AU17, and AU10 which occur mostly in faces expressing negative emotions. High arousal is characterized by AU25, AU27, and AU5. Low arousal is primarily expressed by AU43. High dominance is expressed essentially by AU4 and AU9 and negative dominance is characterized by AU1. AU25 and AU27 characterize both high and low dominance. Lance and Marsella [25] found that during gaze shifts head pitch is significantly related to dominance. That is, head turns upwards indicate high dominance and head turns downwards indicate low dominance. Compared to this, the multiple linear regression shows that in our data head pitch (AU53/AU54) correlates with high dominance. This might be explained by lacking motion (dynamics) while displaying static facial expressions to the subjects. Another argument might be the occurrence of AU54 with AUs showing high dominance (*e.g.*, AUs showing anger). In their analyses Lance and Marsella [25] only concentrate on the relationship of dominance and arousal with head movement during gaze shifts, thus it would be interesting to find out if including pleasure would influence the rating of head pitch with respect to dominance.

The study outlined in this paper has a general caveat. In the case of randomly generated faces no rule as to which AUs can really occur together in spontaneous human facial

expressions is followed. Furthermore the facial expressions were displayed statically with lacking contextual information. This makes the rating process more difficult and sometimes conflicting. Moreover the neutral face was also rated by 10 subjects. Most of them perceived the neutral face as aroused and negatively pleased (this was taken into account in the statistical analysis of the data).

3.3. Facial Expression Simulation through a Virtual Human's Emotion Dynamics

As mentioned in the introduction (see Section 1) the virtual human Emma is the female 'counterpart' to the virtual human Max [24]. As Max, Emma also has a cognitive architecture composed of an emotion simulation module [3] and a Belief-Desire-Intention (BDI) module [27]. The emotion simulation module has as inputs values of emotional valences. These values are either positive or negative and originate from perceiving, through sensory input, positive or negative stimulus, or from the deliberative component of the BDI module, *e.g.*, achieving a desired goal is rewarded with positive values of valence. These values drive the emotion dynamics of the virtual human over time. At each point in time, the emotion simulation system outputs values of pleasure, arousal, and one of two possible values of dominance (dominant vs. submissive). Using the face repertoire resulting from the empirical study each PAD-value output by the emotion simulation system over time is expressed by its corresponding face in the face repertoire thus relating the facial expression dynamically to a virtual human's emotion dynamics.

4. Towards Facial Mimicry for a Virtual Human

As previously mentioned in Section 3.1 approaches in facial expression analysis [1] [17] [28] [35] attempt at recognizing AUs from a human face to obtain a more flexible and versatile interpretation function of facial expression. By these approaches laborious facial expression imitation learning methods for reproducing and mirroring humans' facial expressions can be avoided when the agent's or robot's face is modeled following FACS.

[1] and [19] are developing a system of facial expression analysis with AU recognition in spontaneous expressions since spontaneous expressions occur more frequently in everyday interaction. Following neuropsychological studies (cf. [18]) Bartlett *et al.* [1] state the importance of analyzing spontaneous facial expressions as they differ from posed facial expressions in their dynamics and in which muscles are moved. Spontaneous (involuntary) facial expressions are initiated subcortically and are characterized by fast and smooth onsets with different facial muscles (AUs) peaking simultaneously, while posed (voluntary) facial expressions

are initiated cortically and are characterized by slow and jerky onsets with different facial muscles more often not peaking simultaneously.

Because currently we do not have data at hand from a system of facial expression analysis as described above, the starting point of our conception to realize facial mimicry is a vector of AUs' intensities available from simulating AUs expressing emotion with the virtual human Emma's face. In a first investigation of the idea of developing a system of backward mapping AUs displaying an emotional facial expression on PAD-values, we start up with some assumptions in order to reduce the complexity of this task. First we assume that the simulated facial expression has the same characteristics as a spontaneous facial expression. That is, the facial expression has a fast and smooth onset with different AUs peaking simultaneously. And second, PA-courses related to facial expression onset are output by the system. In this paper only PA-courses for different patterns (AU combinations) of the basic emotions Happy and Angry are presented.

In this first investigation, the task of deriving the PA-courses related to onset of simulated AUs expressing emotion is modeled as a coarse nearest neighbor search problem in multiple dimensions. Since each face in the face repertoire is a combination of different AUs with different intensities, each face can be represented as a multidimensional vector of AUs' intensities. Using a Euclidean metrical distance function the face vector including the most similar AUs' intensities to given AUs' intensities is extracted and the PA-values related to this face vector are returned as the predicted PA-values. That is, given a vector of intensities of simulated AUs, $f_{sim} = \langle i_{sim}(AU_{i_1}), i_{sim}(AU_{i_2}), \dots, i_{sim}(AU_{i_k}) \rangle \in R^k$, $\{i_1, i_2, \dots, i_k\} = ID \subseteq AI$, AI is the set of overall AU Identifiers and a repertoire of faces arranged in PA-space, $FR = \{f_1 f_2 \dots f_m\}$ with $f_{fr} = \langle i_{fr}(AU_{j_1}), i_{fr}(AU_{j_2}), \dots, i_{fr}(AU_{j_l}) \rangle \in R^l$, $1 \leq fr \leq m$, $\{j_1, j_2, \dots, j_l\} = AI$, then the function of returning the face from face repertoire (FR) including the most similar AUs' intensities to the given AUs' intensities can be described as

$$\operatorname{argmin}_{f_{fr} \in FR} \{dist(f_{sim}, f_{fr})\} = f_{min} \quad (1)$$

with $f_{min} \in FR$ and f_{min} including the most similar AUs' intensities to the given AUs' intensities. The function $dist$ is defined as follows

$$dist(f_{sim}, f_{fr}) = \sum_{e \in ID} \sqrt{(i_{sim}(AU_e) - i_{fr}(AU_e))^2} \quad (2)$$

During activation of the AUs simulated with the virtual human Emma's face the values of increasing AUs' intensities are sequentially processed with the function argmin

(1) thus getting the PA-courses related to facial expression onset.

In order to reduce the dimension of the search space, only faces from face repertoire arranged in PA-space of highest dominance with values of positive pleasure and high arousal, and negative pleasure and high arousal are considered to respectively calculate the PA-courses related to the onsets of the facial expressions Happy and Angry. This is since the emotions Happy and Angry correlate with respectively positive and negative pleasure values, high arousal values, and positive dominance values (cf. [33]).

The PA-regression planes of each AU are freeforms showing a nonlinear course with different local maxima and minima of AUs' intensities within the PA-space, *e.g.*, see Figure 2. This makes the task of backward mapping AUs displaying an emotional facial expression on PA-values ambiguous in that given AUs' intensities can be mapped on multiple PA-values. Nevertheless, in this first investigation with the coarse nearest neighbor search function, an overall increase in the values of pleasure and arousal is recorded from onsets of different patterns (cf. [16]) of the facial expression Happy, (AU6, AU12), (AU6, AU12, AU25), and (AU12, AU25). An overall decrease in pleasure and increase in arousal is recorded from onsets of different patterns (cf. [16]) of the facial expression Angry, (AU4, AU5, AU7, AU10) and (AU4, AU5, AU7, AU10, AU27), (*e.g.*, see Figure 4). The PA-courses show more jerky patterns in the interval [0, 0.3] of increasing intensities. This is due to the coarse nearest neighbor classification that returns exactly one nearest neighbor and thus, currently, can not handle the ambiguity described above. A smoother course of PA-values can be recorded by searching for the k-nearest neighbor with more adequate PA-values.

5. Conclusion and Future Work

Based on statistical and reverse engineering methods a control architecture for simulating a virtual human's facial expressions with PAD-values was devised. As result of an empirical study with 353 adult human participants rating randomly generated facial expressions modeled following FACS [16], a repertoire of faces arranged in PAD-space was obtained and integrated in the virtual human's cognitive architecture. Based on the face repertoire resulting from combining the regression planes of all AUs, a first investigation of backward mapping AUs displaying emotion to PAD-values using a nearest neighbor search function was introduced. Since the AUs in each considered facial expression of the emotions Happy and Angry are activated with the same intensity values thus having the same values of apex, as a next step we aim at altering these values of apex in order to better investigate the impact of each AU on the PA-courses. Furthermore the PA-courses of different AU combinations of additional basic emotions such as Sad and

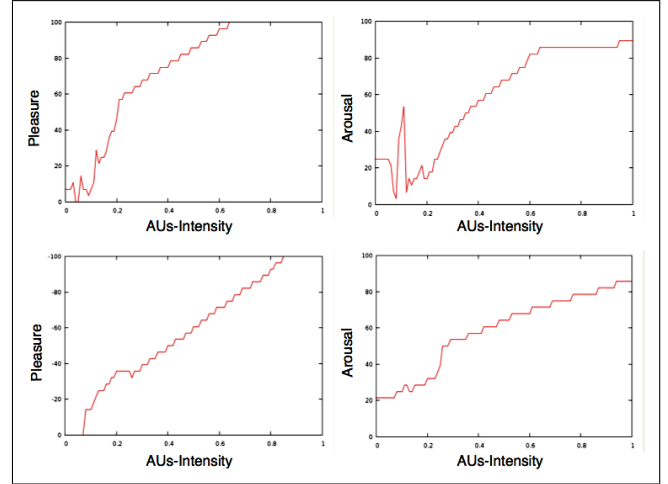


Figure 4. Plots of Pleasure and Arousal (PA) courses, showing pleasure over AU intensity (left) and Arousal over AU intensity (right). The upper plots show PA-courses corresponding to AU6+AU12. The lower plots show PA-courses corresponding to AU4+AU5+AU7+AU10. The AUs of each facial expression are activated with the same intensity values.

Fearful will be investigated. In future work we aim at investigating the impact of Emma's simulated facial expressions while interacting with a human partner. Our long term objective is to enable our virtual human Emma to predict PAD-values while mimicking perceived facial expressions and to use these values to adjust her subsequent behavior during interaction.

6. Acknowledgement

This research is partially supported by the Deutsche Forschungsgemeinschaft (DFG) in the Collaborative Research Center 673.

References

- [1] M. S. Bartlett, G. C. Littlewort, M. G. Frank, C. Lainscsek, I. R. Fasel, and J. Movellan. Automatic recognition of facial actions in spontaneous expressions. *Journal of Multimedia*, 1:22–35, 2006.
- [2] J. B. Bavelas, A. Black, N. Chovil, C. R. Lemery, and J. Mullett. Form and function in motor mimicry topographic evidence that the primary function is communicative. *Human Communication Research*, 14:275–299, 2006.
- [3] C. Becker-Asano and I. Wachsmuth. Affective computing with primary and secondary emotions in a virtual human. *Autonomous Agents and Multi-Agent Systems*, 2009. DOI: 10.1007/s10458-009-9094-9.
- [4] S. Brave, C. Nass, and K. Hutchinson. Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. *International Journal of Human-Computer Studies*, 62:162–178, 2005.

- [5] C. Breazeal, D. Buchsbaum, J. Gray, D. Gatenby, and B. Blumberg. Learning from and about others: Towards using imitation to bootstrap the social understanding of others by robots. *Artificial Life*, 11:31–62, 2005.
- [6] G. Caridakis, A. Raouzaoui, E. Bevacqua, M. Mancini, K. Karpouzis, L. Malatesta, and C. Pelachaud. Virtual agent multimodal mimicry of humans. *Language Resources and Evaluation*, 41:367–388, 2007.
- [7] J. Cassell. Nudge nudge wink wink: Elements of face-to-face conversation for embodied conversational agents. In J. Cassell, J. Sullivan, S. Prevost, and E. Churchill, editors, *Embodied Conversational Agents*, pages 1–27. The MIT Press, 2000.
- [8] J. Cassell. Towards a model of technology and literacy development: Story listening systems. *Applied Developmental Psychology*, 25:75–105, 2004.
- [9] T. L. Chartrand, W. W. Maddux, and J. L. Lakin. Beyond the perception-behavior link: The ubiquitous utility and motivational moderators of nonconscious mimicry. In R. Hassin, J. Uleman, and J. Bargh, editors, *The new unconscious*, pages 334–361. New York: Oxford University Press, 2005.
- [10] W. Cleveland. Robust locally weighted regression and smoothing. *Journal of the American Statistical Association*, 74:829–836, 1979.
- [11] W. S. Cleveland and S. J. Develin. Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association*, 83:596–610, 1988.
- [12] M. Courgeon, J.-C. Martin, and C. Jacquemin. User’s gestural exploration of different virtual agent’s expressive profiles. *Proceedings of the 7th International Conference on Autonomous Agents and Multiagent Systems*, 3:1237–1240, 2008.
- [13] C. Darwin. *The Expression of the Emotions in Man And Animals*. The University of Chicago Press, 1969.
- [14] P. Ekman and W. Friesen. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Consulting Psychologists Press, Palo Alto, CA, 1978.
- [15] P. Ekman and W. V. Friesen. Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17(2):124–129, 1971.
- [16] P. Ekman, W. V. Friesen, and J. C. Hager. *Facial Action Coding System: Investigator’s Guide*. Research Nexus, a subsidiary of Network Information Research Corporation, Salt Lake City UT, USA, 2002.
- [17] R. ElKaliouby and P. Robinson. Mind reading machines: Automated inference of cognitive mental states from video. In *IEEE International Conference on Systems, Man and Cybernetics*, 2004.
- [18] M. Frank and P. Ekman. Not all smiles are created equal: the differences between enjoyment and nonenjoyment smiles. In *Humor: The International Journal for Research in Humor*, volume 6, pages 9–26, 1993.
- [19] A. Goneid and R. ElKaliouby. Facial feature analysis of spontaneous facial expression. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.108.6373&rep=rep1&type=pdf>.
- [20] K. Grammer and E. Oberzaucher. The reconstruction of facial expressions in embodied systems: New approach to an old problem. In *ZiF:Mitteilungen*, volume 2/2006, pages 14–31, 2006.
- [21] M. L. Hoffman. *Empathy and Moral Development*. Cambridge University Press, 2000.
- [22] A. Hofstätter. Implementierung des facial action coding systems in einem Avatar. Master’s thesis, Universität Wien, 2008.
- [23] S. Kopp, L. Gesellensetter, N. Kraemer, and I. Wachsmuth. A conversational agent as museum guide – design and evaluation of a real-world application. In *Intelligent Virtual Agents*, pages 329–343. Springer, 2005.
- [24] S. Kopp, B. Jung, N. Lessmann, and I. Wachsmuth. Max – a multimodal assistant in virtual reality construction. *KI Zeitschrift (German Journal of Artificial Intelligence), Special Issue on Embodied Conversational Agents*, 4/03:11–17, 2003.
- [25] B. Lance and S. C. Marsella. Emotionally expressive head and body movement during gaze shifts. In *Intelligent Virtual Agents*, volume 4722 of *Lecture Notes in Computer Science*, pages 72–85. Springer Berlin / Heidelberg, 2007.
- [26] C.-S. Lee, A. Elgammal, and D. Metaxas. Synthesis and control of high resolution facial expressions for visual interactions. In *IEEE International Conference on Multimedia and Expo (ICME)*, pages 65–68, 2006.
- [27] N. Lessmann, S. Kopp, and I. Wachsmuth. Situated interaction with a virtual human - perception, action, and cognition. In G. Rickheit and I. Wachsmuth, editors, *Situated Communication*, pages 287–323. Mouton de Gruyter, Berlin, 2006.
- [28] Y. li Tian, T. Kanade, and J. F. Cohn. Recognizing action units for facial expression analysis. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 23, pages 97–115, 2001.
- [29] A. Mehrabian and J. A. Russel. *An Approach to Environmental Psychology*. Cambridge, MIT Press, 1974.
- [30] Osgood, Suci, and Tannenbaum. *The Measurement of Meaning*. University of Illinois Press, Champaign, 1957.
- [31] R. Plutchik. The nature of emotions. In *American Scientist*, volume 89, page 344, 2001.
- [32] H. Prendinger, C. Becker, and M. Ishizuka. A study in users’ physiological response to an empathic interface agent. *International Journal of Humanoid Robotics*, 3(3):371–391, September 2006.
- [33] J. A. Russel and A. Mehrabian. Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, 11:273–294, 1977.
- [34] N. Stoiber, R. Seguier, and G. Breton. Automatic design of a control interface for a synthetic face. In *International Conference on Intelligent User Interfaces*, pages 207–216, 2009.
- [35] M. Valstar and M. Pantic. Biologically vs. logic inspired encoding of facial actions and emotions in video. In *IEEE International Conference on Multimedia and Expo*, pages 325–328, 2006.
- [36] A. Wojdel, L. Rothkranz, and J. Wojdel. Fuzzy-logical implementation of co-occurrence rules for combining AUs. In *Computer Graphics and Imaging (CGIM)*, pages 1–6, 2003.