EmotiCon
Interactive emotion control for virtual characters

Bernhard Bittorf
CogVis/MMC, Faculty of Media,
Bauhaus-University Weimar
Bauhausstrasse 11
99423 Weimar, Germany
bernhard.bittorf@uni-weimar.de

Charles Wuethrich
CogVis/MMC, Faculty of Media,
Bauhaus-University Weimar
Bauhausstrasse 11
99423 Weimar, Germany
charles.wuethrich@uni-weimar.de

ABSTRACT

Emotional Expressions are an integral component of the reliability when animating virtual characters. But the sheer amount of possible emotions and the complexity in their selection makes it difficult to develop an intuitive way of controlling arbitrary facial expressions interactively in real-time. This work aims at finding a decent representation of emotions and their expressions, as well as their interactive control by suitable tools. It develops a valid cultural spreading system to classify emotions and thus control emotional expressions interactively. Different complexity stages are presented and evaluated for being able to satisfy different application scenarios.

Keywords

Facial Expressions, Emotion Control, Facial Animation, Virtual Humans, Knowledge-based Animation, Motion Synthesis, Emotion and Personality

1 INTRODUCTION

Realistic facial animation is one of the most difficult tasks for computer animators. The face is the main instrument for communication and defining a person's character. Animating human-like faces is typically done either by using keyframes or through Motion Capturing. Keyframing descends from 2D hand-drawn animation techniques. The usage of Motion Capturing and the evolution of modern graphics cards have led to the possibility of animating human-like faces in real-time. The sheer amount of possible motions requires new methods for controlling these animations in real-time. There are some approaches on defining scripting-languages or parameterized facial models but the field lacks an intuitive way of controlling arbitrary facial expressions interactively in real-time.

The main goal of this research is to find a reasonable model for controlling the state of a character rather than animating the face itself. We introduce a method for interactively controlling the emotions (and thereby the emotional expressions) of virtual characters in real-time which is modelled on the real-life interaction between director and actor. In order to achieve this, we have analyzed existing classifications of emotions and deduced methods and prototypes to control emotional expressions, interactively. Finally an experimental evaluation was performed.

2 RELATED WORK

Controlling human-like faces means basically controlling atomic movements. Such movements depend on the geometric model. The idea behind this work is to decrease the amount of possible movements or actions needed to be done by an animator by providing an abstraction layer which employs different interdependencies and metaphors.

The AMA-System [TMPT88] uses "Abstract Muscle Actions" to simulate muscle movements with single vertices that form the face. Those atomic expressions form so called tracks, which are a chronological sequence of keyframes. Tracks can be subsequently mixed like sound tracks in a recording studio.

Prem Kalra developed the SMILE-System at the University of Geneve. Different abstraction layers are defined to separate muscles, "Minimum Perceptible Actions" (MPAs), phonems and expressions as well as words and emotions from each other [KMMTT91, KMTM+98]. MPAs contain informations about the frame number where to start, the minimal action to be animated and the intensity to which the minimal action shall evolve. The System provides a High Level Script Scheduler (HLSS)

which allows the user to build emotions and words out of those MPAs and arrange the whole animation [MTM00]. The general syntax would be:
`while` <*duration*> `do` <*action*>.

Different approaches have been made to define such a scripting language which enables the animator to act as a director. But in this case a scripting language is not an interface for intuitive interactive work because the process of animating needs flexibility and responsiveness. However, the idea of creating an abstraction layer which enables the user to control a large number of parameters intuitively seems quite convenient though.

The MPEG-4 standard [Koe02] includes a set of Facial Definition Parameters (FDPs) describing the face and 68 associated Facial Animation Parameters (FAPs) which control rigid rotation of the head, eyeballs, eyelids and mandible. The standard also defines parameters that indicate the translation of corresponding feature points. FAPs are also used to describe the most common facial expressions and the visemes. In 2009 Rodrigues et al. developed a dynamic emotion model to facial expression generation using the MPEG-4 standard [RSV09] which was designed to work for the autmatic synthesis but not for interactive, intuitive control.

The Facial Action Coding System [EF78] is a widely used standard to systematically describe all observable facial actions that are independent from each other. Ekman and Friesen found 46 so called Actions Units (AUs).

The Facial Expression Repertoire [BW10] is a set of atomic face movements which is based on the Facial Action Coding System but extends it with asymmetric movement. The main contribution is the mapping of Basic Emotion to a list of Action Units.

In 2004 Helzle et al. [HBSL04] developed a system for mapping captured Motion Curves onto the vertices of a humanlike head. For each existing Action Unit there are 100 captured timesteps that describe the movement from a neutral position to the Action Unit at its extreme value. The Adaptable Facial Setup (AFS) is a a standardized facial rig, driven by a set of nonlinear deformations. In 2007 Schmidt developed a system for animating all possible Action Units via the Adaptable Facial Setup in realtime. [Sch07]

Motivation

Finding a reasonable model for controlling the emotional state of a character and therewith his emotional expressions interactively requires a reasonable classification of emotions. The sheer amount of possible atomic movements (46 in FACS, 68 in MPEG-4, 120 in AFS) - not to mention their intensities - yields the need for an abstraction layer. By controlling the emotions of a virtual character we provide the user with an intuitive abstraction. To satisfy different needs we used different models and developed different interfaces for working with them respectively.

The Adaptable Facial Setup works with 120 Facial Actions multiplied by 68 Influence Points which results in 8160 datasets (Motion Curves) with 100 timesteps each and provides us with the opportunity to render humanlike characters in real-time, and thus explore and evaluate different interfaces and interaction techniques.

3 EMOTICON

The idea of manipulating emotions instead of directly manipulating the face itself provides the user with an abstraction such that an intuitive real-time interaction is possible. But there are two open questions: How can the user control the emotional state of the figure in a way that as many emotional expressions as possible can be achieved? And how are these emotions mapped to emotional expressions or Facial Actions?

Classification

Basically there are two approaches to classify emotions. The *Dimensional Classification* was developed in 1896 by Wundt [Rei00] and describes an emotion by a set of coordinates in a n-dimensional space. In 1986 Russell developed a circumplex model [Rus80, Rus86, Rus02, Rus03], in which all possible emotions can be defined in terms of a pair of xy-coordinates on a plane described by the axes *arousal* and *pleasure* (see Figure 1). The origin marks a neutral state and an emotion results from mixing different states of arousal and pleasure.

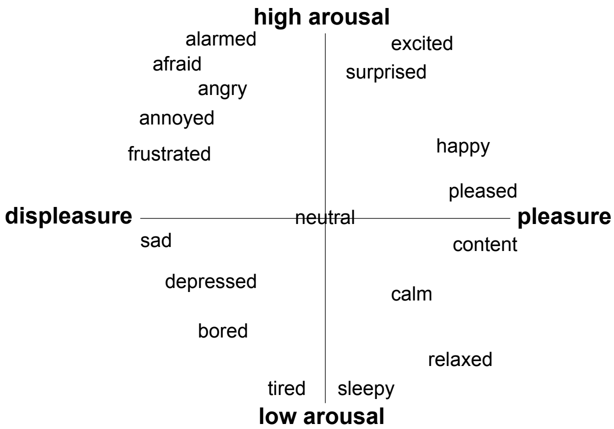

Figure 1: Classification according to James A. Russell

The *Categorical Classification* proposes some basic atomic emotions which can form more complex ones when they are mixed together. Different authors state different sets of Basic Emotions. But it can be shown that the field agrees in at least eight Basic Emotions. Table 1 shows the usage of the 14 most considered emotions in this scientific field whereas "‘+"’ indicates the usage and "‘-"’ indicates the rejection of the emotion respectively.

Emotion	McDOUGALL	PLUTCHIK	EKMAN
Fear	+	+	+
Anger	+	+	+
Disgust	+	+	+
Joy	+	+	+
Sadness	+	+	+
Interest	-	+	+
Surprise	+	+	+
Contentment	-	+	+
Compliance	+	-	-
Tenderness	+	-	-
Contempt	-	-	+
Amusement	-	-	+
Pride	-	-	+
Relief	-	-	+

Table 1: Different Basic Emotions according to different authors

We state that there is a more or less consensus in at least eight Basic Emotions: *Fear*, *Anger*, *Disgust*, *Joy*, *Sadness*, *Interest*, *Surprise* and *Trust*. Robert Plutchik used particularly these eight Basic Emotions for his classification of emotions [Plu62, Plu80]. Figure 2 shows a top view of Plutchiks model. The eight Basic Emotions are represented by eight petals which are subdivided into three different intensity levels respectively. *Surprise* for example increases to amazement and decreases to distraction. Related emotions are adjacent and polar ones are opposed. Mixing different Basic Emotions leads, depending on how far they are away from each other, to primary (*Trust* + *Joy* = *Love*), secondary (*Anger* + *Joy* = *Proudness*) or tertiary dyads (*Fear* + *Disgust* = *Shame*).

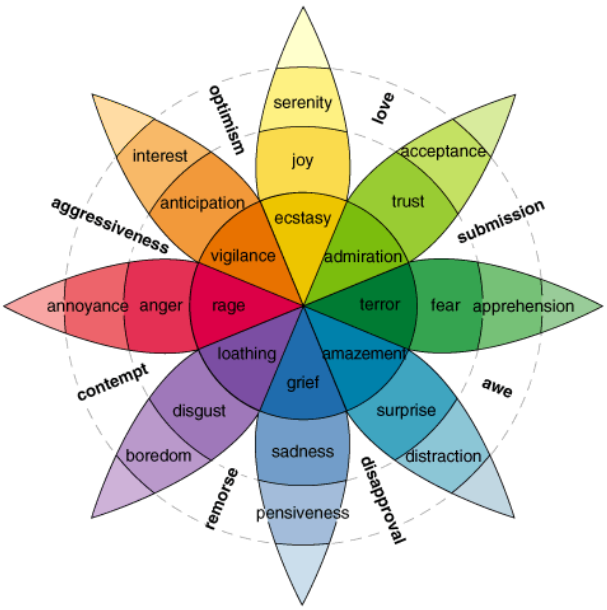

Figure 2: Classification according to Robert Plutchik

Interaction modes

On the basis of the *Categorical* and the *Dimensional Classification* of emotions and their inherent models we developed two interaction models. Because animating humanlike faces can be done in different ways and with different contexts we decided to implement different interaction modes based on different models.

Direct Control

Using Plutchiks classification of emotions as an interaction model allows the user to specify exactly which emotion the virtual character experiences at the moment. Mixing the eight Basic Emotions creates all other possible emotional states and therefore emotional expressions. We developed three interfaces for using this model.

There are basically two ways of controlling the eight Basic Emotions: either one chooses a graphical representation of the emotional space which provides the user with an overview and works as an interface or the input parameters are mapped onto a physical analogon (prop). Both methods were implemented.

The simplest way to interact with Plutchiks Classification is to use the flower-like graph as seen in Figure 2 as a GUI which can be controlled by the mouse. This leads to a very direct interaction but allows the user only to create primary dyads (see section Classification) and thus not all possible emotions. The reduction of possible emotions and the fact that related emotions are adjacent in the GUI leads to a reduction of the cognitive load for the user. The interface allows the animator to define vaguely the mood and its intensity by clicking a petal. The intensity increases to the outer sections of the petal. Thus the midpoint of the flower defines a neutral face which allows the user to fade between different (especially polar) emotions without any visual break.

Another approach is to map the Basic Emotions onto a physical analogon. We decided to implement a more artistic and a more technical prototype. We used one octave of a standard MIDI-keyboard as interface and mapped the eight Basic Emotions onto the eight white keys (as shown Figure 3). Keyboard sensitivity determines the intensity of the particular emotion. Extreme pressure on the keys thus means intenser emotions. This prototype allows the user to mix up to five different emotions per hand - each of them with one finger. But it is not possible to lock a certain emotion and its intensity for improving the other emotions at will because there is no way to lock the keys in a certain position. Furthermore refining the intensity values lacks accuracy.

By trying to combine the best properties of the two interfaces we used a sound mixer as a prop and mapped the eight Basic Emotions onto the eight sliders (see Figure 4). Each slider ranges from 0 to 255 whereas zero means a neutral state and 255 is the most emotional state. This allows the user to mix all eight Basic Emotions independently. At the same time the intensities can be locked and adjusted in a very accurate way.

Figure 3: Analogon 1 - MIDI-keyboard

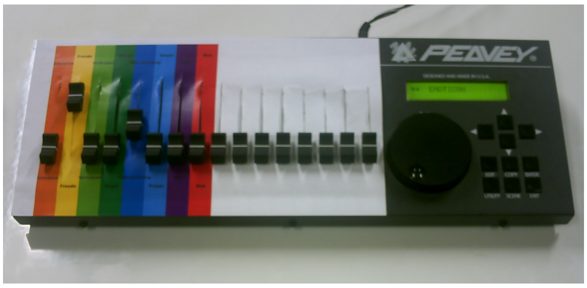

Figure 4: Analogon 2 - mixer

The Direct Control allows an arbitrary mix of all Basic Emotions whereas each single emotion can be controlled almost continuously. Although this model seems to be very intuitive the user has to be very clear about what he is doing which results in an high cognitive load.

Indirect Control

Using Russel's circumplex model as an interaction paradigm allows the user to specify vaguely which emotion the avatar experiences at the moment. Defining the states of arousal and pleasure without having to decide which exact emotion to choose leads to a reduction of cognitive load allowing to perform other tasks simultaneously.

According to the interfaces used for the Direct Control we used the graphical representation of the emotional space as a GUI and provided a prop for navigating.

When using the visualisation as a GUI which can be controlled by the mouse the user can pick a certain emotion and an emotional expression respectively by clicking on the corresponding point within the coordinate system. By holding down the mouse button one can navigate through the emotions which are mixed implicitly. Theoretically the systems enables the user to pick every possible emotion separately. Technically the system mixes the resulting emotions by averaging and normalizing the amplitudes of the individual Action Curves.

As a second approach we used the Spacemouse as a prop. This allows the user to navigate in six degrees of freedom whereas the device is mounted elastically and therefore always returns to a defined rest position automatically. This provides a haptical feedback about the position in the plane. Obviously we only use two degrees of freedom for navigating on the plane and controlling the emotion. This means that the character always starts in a neutral state. For accessing a state of high arousal the character has to go through all in-between states continuously. This leads to a very smooth flow of emotional expressions.

Constraining the input parameters to arousal and pleasure leads to a low cognitive load for the user. On the other hand it can be complicated to access a desired emotion directly and the system appears to be hardly predictable. The restriction on two parameters predestinates the concept of Indirect Control for being used not only for manual synthesis. It can also be used as basis for automatic synthesis of emotional expressions as it could be used for autonomous agents.

4 EVALUATION

All the proposed models and interfaces for controlling the emotional expressions of humanlike figures have their individual strengths and weaknesses. For a decent validation we performed an experimental evaluation. We wanted to know:

- *Which method can be used most intuitively?*
- *Are the interfaces useable?*
- *Do they behave according to the expectations of the user?*
- *Is the method easy to learn?*
- *Which is the most enjoyable to use?*

The experiment was designed as follows:
We asked 12 participants to control a virtual character interactively while a story was read to them. Afterwards they were asked to answer three qualitative paper-and-pencil questionnaires: a general one (age, expertise) and one for each control method. The respondent was presented with a continuous scale between 0 and 100.

- *Is the prototype usable?*
- *Is the usage intuitive?*
- *Is the prototype easy to learn?*
- *How much did the facial animation correlate to what you expected to see?*

The participants were between 20 and 50 years old and their expertise ranged from none to being professional puppeteers. Table 2 shows the results of the Direct Control. The percentage is a degree of satisfaction (0

Criteria	Mouse	Keyboard	Mixer
Ease of use	73%	56%	89%
Intuitivity	75%	74%	92%
Learning curve	84%	66%	92%
Expectations	67%	39%	90%
Overall score	66%	45%	93%

Table 2: Summary of evaluation results for Direct Control

means very unsatisfied and 100 means very satisfied). The variances were rather small (between 10 and 20)

Table 2 shows that using the Mixer as a prop works best for most of the participants (93% Overall). The mixer can be used right away is very intuitive and invites to play (92% Intuitivity). The reason why the keyboard performs so poorly (45% Overall) is that most users had strong difficulties using the keyboard sensitivity. So they were not able to gain the desired emotional intensities. Especially puppeteers had really fun with Direct Control and the mixer. All of them were performing the expected emotional expressions by themselves.

Criteria	Mouse	Spacemouse
Ease of use	60%	67%
Intuitivity	47%	50%
Learning curve	64%	70%
Expectations	54%	54%
Overall score	57%	69%

Table 3: Summary of evaluation results for Indirect Control

Table 3 shows that the Indirect Control paradigm was hard to understand for the users. This was due to two reasons. First, the model uses two abstraction layers. The users had to listen to the story, think about the resulting emotions (first abstraction) and map them to the axes arousal and pleasure (second abstraction). Second, the Spacemouse driver turned out to cause serious trouble. The fact that the Spacemouse performs significantly better than the mouse interface in spite of the driver problems shows that the Spacemouse seems to be a reasonable device for Indirect Control.

All in all the Direct Control performs much better and is more intuitive. The Indirect Control is hard to predict and needs a lot of training. The interviews after the test and the above results show that controlling the eight Basic Emotions with a Mixer is the most intuitive and most satisfying way to interactively control the emotions (and thereby the emotional expressions) of virtual characters in real-time.

5 LIMITATIONS

At the moment the system supports only emotional expressions - neither phonems or signs nor head and eye movement have been implemented.

6 CONCLUSIONS AND FUTURE WORK

We analyzed existing classifications of emotions and deduced reasonable classifications of emotions and their expressions for an intuitive usage and interactive control by suitable tools. We developed a valid cultural spreading system to classify emotions and thus control emotional expressions interactively. We presented two different models with different devices for controlling the emotional expressions of arbitrary virtual characters in realtime. Fig. 5 shows the eight Basic Emotions in their highest intensities according to our system. Our results can be used for the recognition and synthesis of facial expressions. The synthesis can be done automatically as for virtual agents or manually as for virtual characters in computer animated movies and puppetry.

Besides integrating phonems into the repertoire it would be interesting to personalize the emotions, especially for an automatic synthesis. This could be achieved by not only using the Facial Expression Repertoire as translator between Action Units and Basic Emotions but integrating some kind of Fuzzy Logic or parameterisation for example.

Another interesting way of using our system would be to use real faces as input. Either the emotions or the performed Action Units could be computed and mapped onto a virtual character interactively.

Different participants of the evaluation suggested to mix both existing systems. The idea is to navigate roughly with the mouse and then add more emotions using the mixer or the Spacemouse.

Last but not least there is an application which is being discussed with therapists. The framework could be used as a therapeutic tool for patients with emotional perception disturbances.

7 REFERENCES

[BW10] Filmakademie Baden-Württemberg. Facial expression repertoire. http://research.animationsinstitut.de/45.0.html, 2010.

[EF78] P. Ekman and W.V. Friesen. *Facial Action Coding System: Investigator's Guide*. Palo Alto: Consulting Psychologists Press, 1978.

[HBSL04] V. Helzle, C. Biehn, T. Schlömer, and F. Linner. Adaptable setup for performance driven facial animation. In *SIGGRAPH '04: ACM SIGGRAPH 2004 Sketches*, page 54, New York, NY, USA, 2004. ACM.

[KMMTT91] P. Kalra, A. Mangili, N. Magnenat-Thalmann, and D. Thalmann. SMILE:

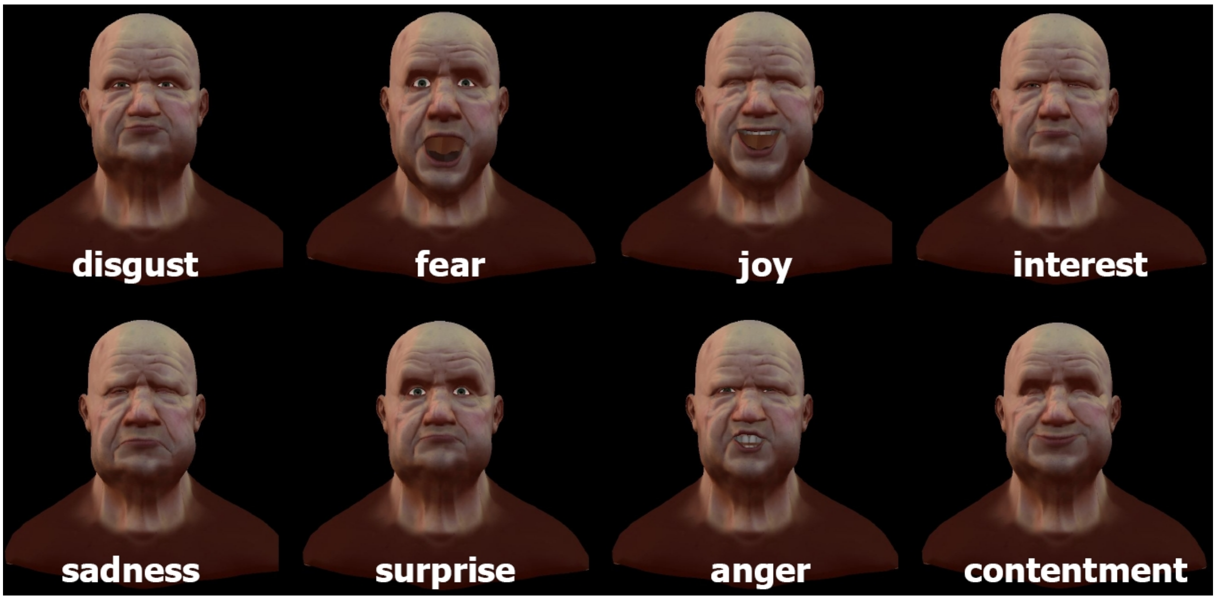

Figure 5: Eight Basic Emotions

a multilayered facial animation system. In *In T.L. Kunii, editor, Modeling in Computer Graphics*, pages 189–198. Springer-Verlag, 1991.

[KMTM+98] P. Kalra, N. Magnenat-Thalmann, L. Moccozet, G. Sannier, A. Aubel, and D. Thalmann. Real-time animation of realistic virtual humans. *Computer Graphics and Applications, IEEE*, 18(5):42–56, 1998.

[Koe02] R. Koenen. *Overview of the MPEG-4 Standard*. International Organization for Standardization, ISO/IEC JTC1/SC29/WG11 N2323, Coding of Moving Pictures and Audio, Geneva, Switzerland, 2002.

[MTM00] N. Magnenat-Thalmann and L. Moccozet. Virtual humans on stage, 2000.

[Plu62] R. Plutchik. *The Emotions: Facts, Theories and a New Model*. New York: Random House, 1962.

[Plu80] R. Plutchik. *Emotion: A Psychoevolutionary Synthesis*. Harper & Row, 1980.

[Rei00] Rainer Reisenzein. Wundt's three-dimensional theory of emotion. *Poznan Studies in the Philosophy of the Sciences and the Humanities*, 75:219–250, 2000.

[RSV09] P. Rodrigues, A. Sá, and L. Velho. Virtual emotion to expression: A comprehensive dynamic emotion model to facial expression generation using the mpeg-4 standard. In J. S. Wright and L. M. Hughes, editors, *Computer Animation*, chapter 6. Nova Science Publishers, 2009.

[Rus80] J. A. Russell. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6):1161–1178, 1980.

[Rus86] M. Russell, J. A.; Bullock. On the dimensions preschoolers use to interpret facial expressions of emotion. *Developmental Psychology*, 22:97–102, 1986.

[Rus02] J. A. Russell. Reading emotions from and into faces: Resurrecting a dimensional-contextual perspective. In James A. Russell & Josi-Miguel Fernandez-Dols, editor, *The Psychology of Facial Expression*. Cambridge University Press, 2002.

[Rus03] J. A. Russell. Core affect and the psychological construction of emotion. *Psychological Review*, 110:145–172, 2003.

[Sch07] S. Schmidt. Glaubhafte gesichtsausdrücke und animationen für echtzeitanwendungen. Master's thesis, Bauhaus-University Weimar, 2007.

[TMPT88] N. Thalmann-Magnenat, N. Primeau, and D. Thalmann. Abstract muscle actions procedures for human face animation. *Visual Computer*, 3(5), 1988.