IV Iberoamerican Symposium in Computer Graphics - SIACG 2009
O. Rodríguez, F. Serón, R. Joan-Arinyo, J. Madeiras, J. Rodríguez, E. Coto (Editors)

FACe! 3D Facial Animation System based on FACS

S. Villagrasa[1] and A. Susin[2]

[1]La Salle School of Engineering, Ramon Llull University, Barcelona, Spain
[2]Dept. Applied Math.- LabSiD. Polytechnical University of Catalonia, Barcelona, Spain

Abstract

In this paper we present a 3D facial animation system named FACe! It is able to generate different expressions of the face throughout punctual and combined activation of Action Units, defined by Facial Acting Coding System (FACS). This system is implemented on a 3D human head controlled by bones, riggers and skinning to deform the geometry. The bone system is implemented in order to move single or combined Action Units, so that they can implement superior layers such as expressions, phonemes, words, emotions and the synchronization of all them together.

1. Introduction

The modeling and the animation of an expressive human face is one of greatest challenges in computer graphics. Nowadays, the demand for facial animation is clearly increasing, from entertainment industry as a driven force through internet and medicine.

Why the goal of animate a face with quality and reach the limit of synthesizing a real face is so difficult to achieve? A part of the answer lies in human perception of human faces. We are able to detect the nuances of even the most basic movement.

On the other hand, the reproduction of human expression is highly complex at computer graphics, due to the great detail that must be achieved in texture, modeling, skin look and animate each of the muscles, which makes a complex work to overcome. [PW96].

Furthermore, in terms of animation, it is complex to simulate the movement of the muscles and skin, and additionally, you must to keep in mind the restrictions to play each of the facial movements.

In this paper, we present a first step towards the development of a system of multilayer 3D facial animation, simple to use by an animator, fast in production and reusable for multiple models. Recently we have different software packages that help in the work of the facial animation, such as face robot or facial studio. Although theses packages are very advanced developments the high price and the particular workflow limit its use.

We present the idea of the system called FACe!, and in particular the construction of the skeleton which is responsible for moving the areas of the face, the weight of vertex (skinning) and a correct rigger to make it functional and easy to use. This system of bones and rigger, must be consistent with each of the action units defined by FACS. [EF78].

In upper layers, our system can be defined using mark-up languages, such as phonemes or expressions, which automatically articulate the Action Units (AUs) needed to define a phoneme or an expression. Moreover, we can work in the animation, defining just which word or emotion we want to generate.

Thus, in a last layer, we can synchronize words, emotions, subtle movements of eyes and head, and even defined as an emotional state or personality given a consistent facial movement in a simple and intuitive for the animator, which will save time in production, improve the quality of facial animation and gives control of all the layers of animation defined by FACe!

One of the problems that we can expect using a system of bones is the transferring from one system to another model. There are methodologies capable of transferring successfully riggers from one model to another, such as face in motion [CS07].

Primarily, the entertainment industry has developed several technologies to assist in the production of facial animation [RP06]. The origin of the facial animation is based on the ideas of F.Parke and K.Waters [PW96]. We have from the

techniques of traditional animation with keyframing, blendshapes [JTDP03], interpolation of control parameters, muscles action based [Wat87]through performance driven systems such as motion capture (mocap) [PL06b].

It is very subtle the recreation of a perfect facial animation, which must interplay the bones, muscles, skin... If you want to recreate a real character, texture and modelling, they should be of the highest quality, and also to be recreated until the smallest detail all the facial animation. If something fails, the human eye is trained to notice that something is not real, and it will happen the phenomena named uncanny valley [Mor70], which means that it will look like as a non-human appearance.

2. FACS

In order to recreate the expressions in FACe! System, it is not based on the accurate recreation of the muscles (muscle based system) [KHS01], but in the action that is displayed in the face. To do it, it is based on the study of Dr. Paul Ekman in his Facial Action Coding System (FACS) [EF78].

The Facial Action Coding System or FACS was originally designed in the 1970s by Paul Ekman, Wallace Friesen and Joseph Hager. Given that the face has muscles that work together in groups called Actions Units (AUs), FACS teaches how to understand when these action units are triggered and give a relative score to them. Although initially designed for psychologists and behavioral scientists to understand facial expressiveness and behavior, it has also been recently adapted in visual communication, teleconferencing, computer graphics and CG animation.

Paul Ekman et al. categorized facial expressions into 66 distinct Action Units. Each Action Unit represents muscular activity that produces momentary changes in facial appearance. This change in facial appearance of course varies from person to person depending on facial anatomy, e.g., bone structure, fatty deposits, wrinkles, shape of features etc. However, certain commonalities can be seen across people as these action units are triggered.

The action units in FACS are based on the location on the face and the type of facial action involved. For example, the upper face has muscles that affect the eyebrows, forehead, and eyelids; the lower muscles around the mouth and lips form another group. Each of these muscles works in groups to form action units. The action units can further be broken down into left and right areas of the face, which can get triggered asymmetrically and independently of each other.

3. The FACe! System

First, we should ask ourselves which is the objective to cover with FACe!: Maybe we want small quantities of high quality animation with a lot of time of production or, perhaps, we need large amount of animation with basic animation,

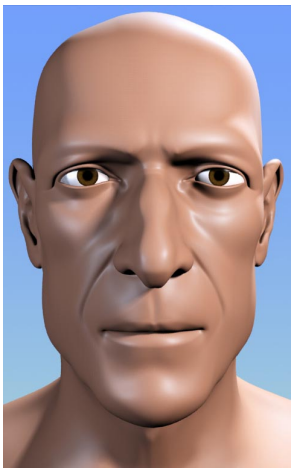

Figure 1: *Start and neutral face.*

but short time of production. FACe! system must be flexible to deal with the all set of conditions of production time. A system capable of adapting to these variables will become a first class product. Currently, it is difficult for the industry to choose between existing technologies and they create in-house technology to cover their needs.

The system that we describe in this paper is implemented and running on a facial model capable of playing efficiently all the AUs (Action Units) of the FACS standard in its layer 1, layer 2 and layer 3, as can be seen in the Table 1.

Layer 7	Sincro of all layers
Layer 6	Behaviors and Personality
Layer 5	Perceptible movements: Eye, Head
Layer 4	Words and Emotions
Layer 3	Expressions and Phonemes
Layer 2	Action Units
Layer 1	Bones Model, Skinning& Rigger

Table 1: *Multi layered system of FACe!*

Let's turn to describe more specifically the system. The first step is to obtain a correct rigger. For this, we have working under the 3DSTUDIO MAX platform for building bones system, the full-rigged and skinning. Once we have a system capable of moving parts of the face like the Action Units describe, we have worked in the module capable of activating the AUs to generate expressions (layer 3), and those phrases can activate the emotions (layer 4). This activation may be read from a database of Action Units using any system of mocap or keyframed.

If we are working in layer 4, the system only needs to read from a database how they want to move some single AU. The AU is a point that has keyframe animation. If the AU is done by mocap, then when we're playing the AU will have the

same movement that the person had been captured the AU. For this to work, it is essential that the system of bones and rigger is working to move all that AUs defined by FACS and not blendshapes expressions or phonemes, as usually done today.

For layer 4, emotions, require a harder research, since they must interpolate in time between the different expressions generated without going through state or neutral expressions, and also taking into account slow and fast movements and their combinations. The primary advantage of FACe!, it is able to generate facial animation quality, without having to rely on expensive facilities and mocap with a simple interface to use for an animator. The activation of one AU means that the character can express a certain emotional state, a certain behaviour or phonemes needed to articulate a speech.

3.1. Implementation

For 3D animation of the face, we propose to build on a system of internal bones, which means we can handle the rigging AUs necessary. The bones may move the vertices of geometry through a weight of vertices (skinning).

We propose a system of bones due to their flexibility, low computational cost, and adaptation to any objective or problem, compared with methods based on blendshapes, or muscle based simulation.

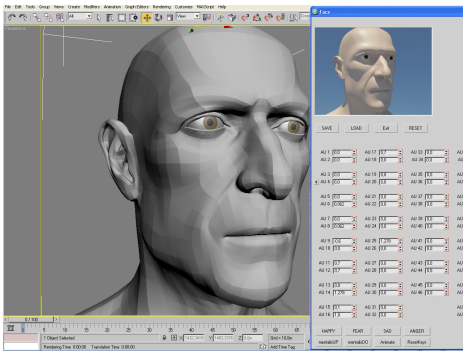

Figure 2: *3DMAX User Interface of FACe! With the controls of the layer 2, witch control all the action units separately, symmetrically or asymmetrically.*

FACS categorizes each movement on a visible unity of action. So for lower the brow, is categorized as the Action Unit 4 (Brow lowerer). With this unity of action, and with adequate controls, we don't need to know what muscle there are below the eyebrow to move (in this case would corrugator, procerus, caninus , quadratuslabii, major and minor zigomatic, frontalis, and orbicularis oris).

Working with Action Units, the control for an animator for facial movements is simple, accurate and easy.We can also move the AUs to get the desired expressions.

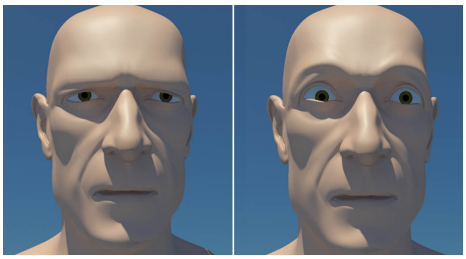

Figure 3: *Action Unit 4: Brow lowerer in maximum and minimum weights.*

Take for example that we want the expression of smile. One combination defined by FACS can activate the AU6 and AU14 corresponding to the cheek (Cheek Raiser) and the angle of the mouth (Dimple) respectively. We also implement the control of the intensity (weight, 0 to 100%) of the AUs to report degrees of activation of the action unit.

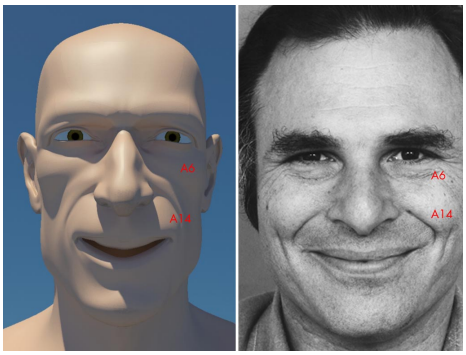

Figure 4: *Action Units 6 + Action Unit 14 combination: The result is a standard expression of smile.*

The system FACe!, allows us to be able to combine different AUs to recreate different expressions. AU6 and AU14 synthesize a standard smile, but maybe there are people who also slightly up the upper lip for generate a smile. So we must add AU6 + AU14 + AU23. Or someone could press lip and tighten slightly, which would execute the combination AU6 + AU14 + AU24.

Or if we want an exaggerated smile, get a combination of AU6 + AU12, which up the cheek and pulled up to the outside of the lips, or even some people with an exaggerated smile tight-lipped, so suggests the combination AU12 + AU20, for up the outside of the lip (AU12) and tighten the lip (AU20).

Take another example of working with FACe! To synthesize the character an expression of anger, simply instruct the system to activate units of action AU4 (lowering of eyebrows), plus the combination of AU7 (Lid Tightener) de-

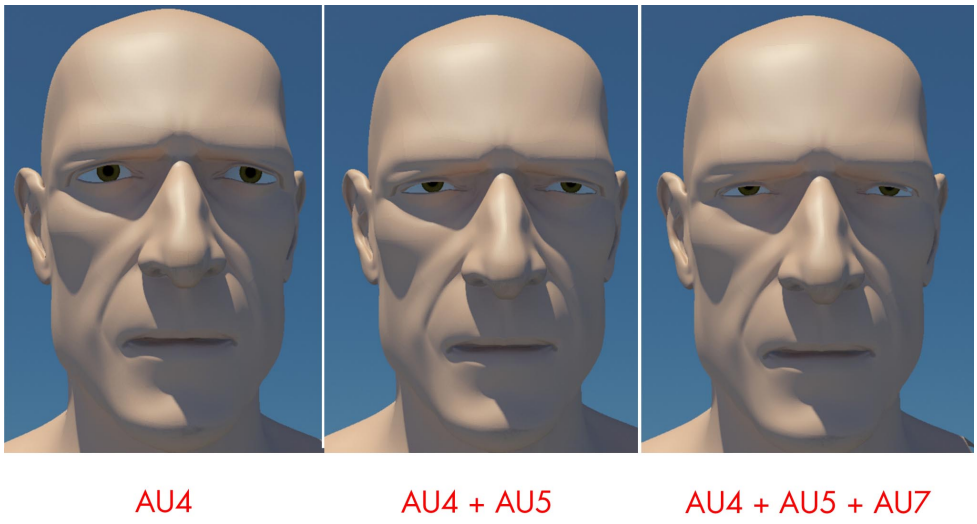

Figure 5: *Different combination Action Units for synthesizes an expression of anger.*

pending on the character or the AU5 (Upper Lid Raiser) or even the combination of all three at once, AU4+AU5+AU7.

In psychological research has determined that we have thousands of possible combinations to recreate expressions. That is why we must work on future work on a system that can automate the basic movements of each of the expressions to leave them alone to recreate the work of animator by manipulating AUs characterization of each character, and able to work more comfortably in upper layers as the emotions and speech.

One characteristic of the system is the ability to capture through mocap all AUs of a person, and also the most basic possible combinations. The training would be through a combination of phrases and expressions of the user in order to save the AUs and their combinations, keeping the user more comfortable while being captured. So, once we have the AUs and their combinations for a particular person, we can recreate again and again the facial animation of that specific person.

For the implementation of the system FACe! described in this paper also enables layer 3 (expressions and phonemes), and can instantly reproduce expressions using combinations of AUs, and thus recreate the 6 expressions bases defined by Ekman: joy, anger, sadness, fear, surprise and anger.

Another example of the system with the emotion FEAR: For achieve this we must combine for example the Inner Brow Raiser + Outer Brow Raiser + Brow Lowerer + Upper Lid Raiser + Lip Stretcher AU1 + AU2 + AU4 + AU5 + AU20.In the figure 7 we must see the Action Units that involucres in the combination for synthesize a fear expression.

3.2. Blowup of FACe!

We can think different solutions for the recreation of the AU as would be the blendshapes, muscles, bones or rigging.

The advantages one or other are obvious: The technique of blenshapes, is simpler to implement, but complex in terms of production because we have modelled each of the AUs for each AU. The problem with blenshapes is that we do not have sufficient controls to recreate nuances and subtle movements. The muscular system [KHS01] is far away for implement for computational cost and too high complexity to manipulate. And the creation of geometry with behaviour similar to the muscle, although there are several approaches, the calculation needed to successfully create a functional system for the animator is too complex.

In contrast, a system of bones and rigging, although it is very complex to create, gives us a hierarchy of bones so that depending on the movement of the bone may be deformed through proper weighing of area vertices (skinning) that we need to move of geometry. The goal of bones

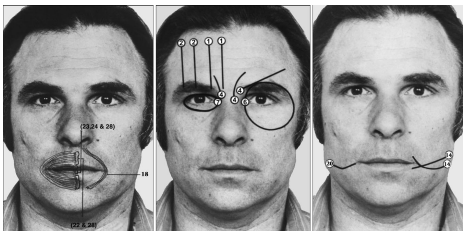

Figure 6: *FACS manual describing AU1, AU2, AU4, AU5 and AU20.*

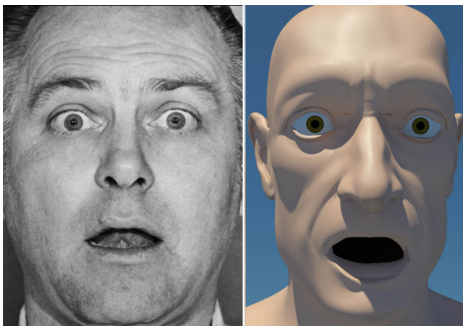

Figure 7: *Image reference of FACS for combination of FEAR described en figure 6. (right) FACe! synthesize FEAR due a combination of the action units describe in figure 6.*

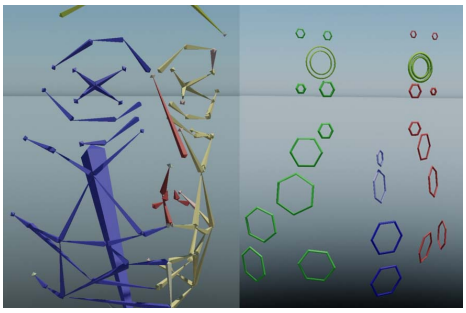

Figure 8: *Bones system and the rigged locators responsible for moving Action Units.*

and rigger, is the advantage that not having to model a vast database of heads, and a smaller set of controls. In addition, we could also use the methodology of the system face in motion [CS07] to transfer to other models Rig (retargeting) [PL06a].

Another advantage is that system is upgradeable, because if we need some movement or not originally referred to a very subtle facial movement, we can add bone relevant to the system rigged.

In the construction of the hierarchies of bones have been programmed special functions that the bones are fully chained and can also stretch to simulate elongations of the skin by scripting language. The system, bones based facial animation, finally have more control in order to generate small random movements of the head, activate randomly AU 66 and AU 65 for blink, or turn subtly with certain intensity a small AU to diversify and customize a character. For speech, the system works the same way as for the generation of expressions. There is a chain of bones which is also controlled by AUs. So we have a top level manager control which groups the phonemes in different combinations of AUs necessary to articulate the phoneme in particular.

3.2.1. Layer One

The Bones System

If we take a look at the table 1, we must see the layers that compose the full system of FACe! We have to build the structure of bones, which is based on the underlying muscles beneath the skin. We have not already implemented all the muscles, but only those who have special relevance in the Action Units.

The complexity of this phase is to create a structure of bones be able to play until the 66 Action Units of FACS. We comment on what is most important:

It can be seen in figure 8 , which the bones colored in blue represent a related structure, that is, if placed in a particular bone, such as the cheek, this move to the rest of the structure to simulate the movement of the skin.

To achieve this purpose, we developed a hierarchy of bones efficiently compensated. When you move a bone, you are applying weight corrections to the rest of the structure so that the movement is damped or scaling depending on the area of the face of interest to move.

Here we have encountered problems in terms of consistent movement in the hierarchy of bones, because when the hierarchy has made in his motion angles, the implementation not working well. We had to apply corrections to the 3D rotation doing the calculation with quaternions and functions slerp (shorthand for spherical linear interpolation) [Sho85], all under the language maxscript.

The bones system has been created with 3DSTUDIO MAX using wire parameters and a complete rigger.

The Rigger

For Rigger, it is necessary to control the structure of bones, because and animator can not move a bone directly for unmanageable reasons.

It is worth to emphasize that the points of riggers and locators that control the bones, are analyzed in order to move singularly or in combinations for each of the Action Units defined by FACS. These Locators can also be capture, if desired, by mocap. So the points placed in a system of mocap, which is actually working, will be a direct translation of the capture of points to the Rig. The entire skeleton will be able to move directly Units Action and will not get direct information from the capture but in action units controllable and easy to manipulate.

As seen in figure 8, locators individually do not coincide with the Action Units, for create the majority of Action Units must drive multiple locators and have precise control of skinning. These locators, are controlled by small pieces of code at the bones. He also is scheduled to have some influence with locators to other locators.

The basic and key points of the face, the feature points,

we have relied on the key points made by VICOM systems and also in the FAPS of the standard MPEG-4 [PF03].

The FAPs are based on the study of minimal perceptible actions and are closely related to muscle action. In fact to order to define face animation parameters for arbitrary face models, MPEG-4 specifies 84 feature points located in relevant somatic places of the human face. Our system is not based on mpeg-4 but it is a good source of knowledge.

The Skinning

As a final part, we have to work with a skinning suitable for the deformation of the geometry and thus simulate the behaviour of the skin shifts to the bones move.

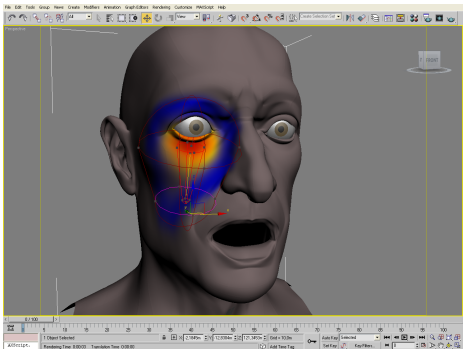

Figure 9: *Skinning with skin modifier of 3DSTUDIO MAX.*

As we see in figure 8, the weighing of the vertices must be performed bone to bone, through configuring the range of the envelopes for all the vertices of the geometry. For certain vertex it is necessary not only to use the envelopes of skin modifier of 3DSTUDIO MAX but a lot of time is needed to handle a weight for all vertices so that the deformation of the geometry is correct.

3.2.2. Layer TWO

Action Units User Interface

As we can see in Figure 9, as the controls of Action Units, for Layer 3, we have developed an interface from which any animator can control each of the action units. In this interface, we can control the intensity of each Action Unit.

From the user controls we can decide whether the right or the left part of the face we want to move, and movements such as compression, such as frowning, in which control the Action Units in particular, we can gather the proper Locators AU to achieve that.

4. Conclusions and Future Work

As future work, we are currently developing an interpolation between expressions in order to give the character the emotions, and a convincing dramatic performance.

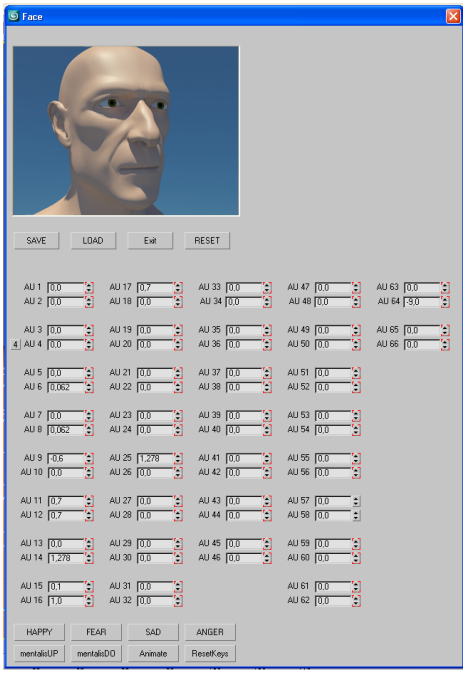

Figure 10: *User Interface Layer for an animator.*

The functionality of wrinkles is not yet implemented and is as a line of future work, which we can solve via deformation geometry or with normal mapping and masks.

Another functionality is drive de animation by a behaviour driven (layer 6), that animate the character with a certain personality and motivation.

With this work, we have presented the base system for a 3D virtual character able to offer dramatic impact, with a simple and easy interface for the an animator. One benefit is less time production for generate high-quality animation, and be able to reuse over and over again the same gestures of an actor, if we have the AUs and their combinations stored in a database.

In addition, with FACe! We can make a face expressing cartoon or fantastic head as it would a human, which is one of the requirements that the industry is the computer animation. Humanize a character not human, animal or a fantasy, always guaranteed an emotional link with the spectator or gamer.

5. Acknowledgements

This work was supported by the University La Salle- Ramon Llull and LabSID of Universitat Politècnica de Catalunya. Second author is partially granted by TIN2007-67982-C02-01.
Thanks to Ivan Costa for the face model.

We would like also to thank to Judith Clares, Carlos Villagrasa , Emiliano Ruiz and Antonio Pastor for their help and support.

References

[CS07] COSTA V., SUSIN A.: Fast and reusable facial rigging and animation. In *SIGGRAPH '07: ACM SIGGRAPH 2007 sketches* (New York, NY, USA, 2007), ACM, p. 63.

[EF78] EKMAN P., FRIESEN W.: *Facial Action Coding System: A Technique for the Measurement of Facial Movement.* Consulting Psychologists Press, Palo Alto, 1978.

[JTDP03] JOSHI P., TIEN W. C., DESBRUN M., PIGHIN F.: Learning controls for blend shape based realistic facial animation. In *SCA '03: Proceedings of the 2003 ACM SIGGRAPH/Eurographics symposium on Computer animation* (Aire-la-Ville, Switzerland, Switzerland, 2003), Eurographics Association, pp. 187–192.

[KHS01] KÄHLER K., HABER J., SEIDEL H.-P.: Geometry-based muscle modeling for facial animation. In *GRIN'01: No description on Graphics interface 2001* (Toronto, Ont., Canada, Canada, 2001), Canadian Information Processing Society, pp. 37–46.

[Mor70] MORI M.: The Uncanny Valley. *Energy 7*, 4 (1970), 33–35.

[PF03] PANDZIC I. S., FORCHHEIMER R. (Eds.): *MPEG-4 Facial Animation: The Standard, Implementation and Applications*. John Wiley & Sons, Inc., New York, NY, USA, 2003.

[PL06a] PIGHIN F., LEWIS J. P.: Facial motion retargeting. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Courses* (New York, NY, USA, 2006), ACM, p. 2.

[PL06b] PIGHIN F., LEWIS J. P.: Introduction. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Courses* (New York, NY, USA, 2006), ACM, p. 1.

[PW96] PARKE F. I., WATERS K.: *Computer facial animation*. A. K. Peters, Ltd., Natick, MA, USA, 1996.

[RP06] RADOVAN M., PRETORIUS L.: Facial animation in a nutshell: past, present and future. In *SAICSIT '06* (Republic of South Africa, 2006), South African Institute for Computer Scientists and Information Technologists, pp. 71–79.

[Sho85] SHOEMAKE K.: Animating rotation with quaternion curves. In *SIGGRAPH '85: Proceedings of the 12th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 1985), ACM, pp. 245–254.

[Wat87] WATERS K.: A muscle model for animation three-dimensional facial expression. In *SIGGRAPH '87: Proceedings of the 14th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 1987), ACM, pp. 17–24.