Intelligent facial animation

Creating empathic characters with stimuli-based animation

José Serra
Instituto de Telecomuniçöes - Porto
Rua Dr. Roberto Frias
4200-465 Porto, Portugal
jserra@dcc.fc.up.pt

ABSTRACT

Traditional facial animation techniques require either an artist to create and/or clean the animation by hand, which is very time consuming, or rely on motion capture techniques, where fine tuning is usually still required. Generating facial animation becomes even slower when modelling the reactions to external stimuli. Procedural methods are an alternative to these approaches as they rely on an algorithm to automatically generate a short track of animation. Thus, allowing partial automatization of the animation process, however, most research has focused only on procedural body animation, which left the face under-explored.

The purpose of this PhD research is to define a novel method for stimuli-based procedural facial animation with believable quality. To be capable of modelling the motions procedurally, a study on facial movements will be done to understand the basic motions and how they vary according to different types of characters. Creating a stimuli-based method also requires the definition of the character's behaviour, unfortunately most research has focused on emotional variation and its expressions, leaving other facial movements to be explored, e.g. tics. Hence, the key contributions of this work are: a study on facial movements; a procedural animation method; a stimuli-reaction model for facial animation. The results will allow reducing the time needed to create facial animation, having great impact in film industry and interactive systems allowing the creation of reactive and highly empathic human-like characters.

Categories and Subject Descriptors

I.2.1 [**Applications and Expert Systems**]: [Games]; I.3.7 [**Computer Graphics**]: Three-Dimensional Graphics and Realism—*Animation*; H.1.2 [**User/Machine Systems**]: [Human factors]

General Terms

Algorithms, Theory, Design, Human Factors, Performance

Keywords

Procedural Animation, Facial Animation, Agents, Realistic Behaviour

1. INTRODUCTION

Non-verbal communication plays a crucial role while interacting with others. A facial expression, a gaze or a nod can tell much more of the personâĂŹs mental state than just what is being said. Emotions are shown not only when a person is interacting with others, but also, when e.g. something unexpected happens, which can itself provide insights on the person. Non-verbal communication allows us to communicate in more meaningful and empathic ways than only verbal communication. This is true to either a real person and to a virtual human. Thus, when producing the animation for a 3D character, there is a need to generate faithful animation, whose movements match the ones of a person. This is extremely important for realistic characters and to some extent to fantasy characters.

The purpose of this PhD research is to define a novel method for procedural facial animation and combine it with artificial intelligence to obtain a stimuli-based animation method capable of producing believable facial animation. To achieve this, it is first necessary to study the facial movements of both realistic and fantasy characters, which serves as the core of the procedural method. Finally, automatic visual speech, or lip-sync, will be added to the stimuli-based animation method to create a talking head.

1.1 Motivation

This research is driven by the need to produce believable animation in a fast and easy way. In traditional animation, the artist produces an animation by manipulating the model, creating all the poses and then associating them to a specific time, allowing realistic results based on the artist's experience, however it is very time consuming. Another approach consists in generating the animation from data of an external source, such as an actor's performance, which fastens the animation process, however fine tuning is still required. Procedural animation relies on an algorithm and a set of parameters to produce the animation, allowing a reduction of the time compared to traditional techniques. Research on procedural animation has been mostly focused on the body [4], with studies on the face still in its beginnings.

Reducing the time required to generate an animation can be increased by using artificial intelligence (AI). In this case, the artist, better referred as the author, would only define the behaviour of the AI. Combining AI and facial animation has been mostly done by matching emotional states directly with facial expressions, leaving e.g. tics not simulated. Such combination would primarily benefit real-time applications, leading to the creation of more empathic characters. Automatic visual speech animation will also be combined with the stimuli-based procedural animation, which results from the desire to continue previous research done on this field.

This work is organized as follows: the state of the art, the main contributions expected to arise from this work, an overview of the expected methodology architecture, followed by a work plan with a description of the approaches used to obtain the contributions and finally a brief discussion of this thesis and the repercussions that may arise.

2. STATE OF THE ART

In procedural animation an artist controls the animation through a set of parameters that are the input to an algorithm. A complex animation is generated from short pieces of animation, either by concatenation or from blending. The algorithm can be based on mathematical or physical formulae, constraints (IK) or another approach capable of generating the continuous variation of motion. Such algorithm has to be capable of generating both the position values and the timing information of the facial rig, with neither being a directly controlled parameter. This definition was based amongst others on the ones proposed on [5, 4, 9], however these tend to be too broad and can conflict with data-driven techniques or even with rigging methods.

Perlin has contributed decisively to both body and facial procedural animation. In [6], Perlin extended Improv [7] to include procedural facial animation, which was based on a layered approach, where the FACS [3] were used at the lowest level to define the DoF of joints. In this work, the timing information is left as a parameter, thus being a trial and error process. Other works also allow generation of facial expressions [10, 1] from a set of parameters that use fuzzy logic to introduce small variations, however again timing information needs to be provided. Gaze and head movements simulation are also important to achieve realistic animation. Queiroz et al. [8] include these two in their system, which is controlled using a scripting language. Nevertheless, facial expressions are only based on emotional expression leaving other movements out. The best work so far on procedural animation is the one proposed by Perlin, however at its core it is based on Ekman's FACS, which limits the facial expressions that can be obtained, e.g. in mouth area, aside from being completely focused on realistic characters.

At Improve [7], Perlin and Goldberg proposed a system to generate real-time character animation that includes a behaviour engine defined through a set of rules. The behaviour engine is connected directly with the animation engine specifying the animation parameters, being one of the earliest examples of behaviour animation. Research on behaviour facial animation has mostly focused on the direct expression of emotions, such as in [2] that included a emotional motion synthesizer. There are some studies that tie other facial movements to behaviour animation and an high level control parameters, such as [12]. However, its facial animation system is based on the MPEG-4, which is more specific to human faces. Modelling behaviour already produces believable results [11] in some conditions, however combining AI with animation in a way that it is easy and predictable to define a behaviour without leading to repetitive actions is still an open question.

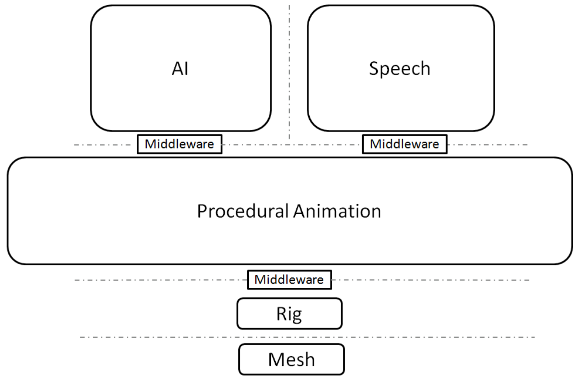

Figure 1: Hierarchical overview of the three major areas of this research, which includes AI to react to a stimulus and produce the procedural animation parameters, visual speech to create a talking head based on the analysis of audio and the procedural facial animation module that generates an animation from a set of parameters. The procedural methods lies on top of the rig requiring a middleware to translate the motion data into the different types of rigs and finally the moving the mesh.

3. RESEARCH CONTRIBUTIONS

We hope this research will result in several contributions to science. These can be summarized in:

- Study on facial movements both of realistic characters and fantasy ones.
- Procedural facial animation method capable of generating believable animation using a âĂIJsmallâĂİ set of parameters that allows a reduction of the time required to generate believable facial animation.
- AI method capable of generating a complete sequence of animation based on stimuli from a user.
- Talking head demo that combines visual speech animation, AI and procedural facial animation.

4. ARCHITECTURE OVERVIEW

Creating a talking head based on stimuli-based procedural facial animation comprises three major areas, whose relations are described in Fig. 1. Procedural animation is the core and we expect it to be based on constraints obtained from the facial animation study. This mode is independent of the rig as it lies in the animation layer, nevertheless it is still necessary to make a connector, or middleware, that will vary according underlying structure of the model. An experienced animator will generate the animation directly on the procedural animation layer. On top of this layer, the visual speech and AI layers are responsible for generating the parameters for the procedural method. On the the visual speech layer the parameters' values are generated from the analysis of audio, using a map created using machine learning techniques. The AI layer is responsible for analysing the external stimuli that might have emotional attributes, such as valence and arousal, and physical attributes such as speed and direction. The character's personality and emotions that include e.g. the physiological and psychological states are also taken into account to choose the values for the procedural layer. A casual user or a non-expert animator might control how the animation is generated in this layer.

5. RESEARCH PLAN

The research required to achieve the proposed goal was divided in several stages, where each is built on top of the previous. The steps are the following:

1 - User requirements - Study the user of the system, how artists create the animation, tools used and interest in the proposed system by the industry; Mostly done using questionnaire(s) to artists and companies. *Expected outcome*: Guidelines to define the system and allow its integration in the current pipeline.

2 - Prototyping and State of the Art - Study currently used methods and create a prototype that integrates the different subjects. *Expected outcome*: State of the art report mostly focused on intelligent procedural facial animation. A prototype with all the parts will be integrated.

3 - Facial Animation Study - Understand the facial motion and its limits according to various mental states and animation styles. This will be done using MoCap and from the analysis of motion in films and videogames. *Expected outcome*: Constraints and vector field for the facial motion.

4 - Procedural Facial Animation - Propose a procedural facial animation method capable of generating both new poses and facial motion from a *small* set of parameters. The control method used by the animator also has to be defined, taking into account how the animation is created, either through blending or concatenation. The starting point for this part is the work proposed by Perlin [6].*Expected outcome*: Prototype that demonstrates the procedural method is controlled and how the animation is created.

5 - Artificial Intelligence - Create the behaviour facial animation system, by choosing the best fitted decision making system and then adapt it to generate the appropriate parameters for the procedural facial animation system. It is also necessary to study the best way to control such a system in a user-friendly way. *Expected outcome*: Prototype capable of handling external stimuli and generating the appropriate animation.

6 - Intelligent Talking Head - In this task, a coarticulation model will be implemented to generate realistic facial animation, possibly based on step 4. After this, a talking head that combines intelligent facial animation with visual speech will be implemented. *Expected outcome*: Prototype that combines all the previous technologies.

Validation is also required after steps 4, 5 and 6, as it is extremely important to continue work on top of proven work. Validation will be based on subjective questionnaires and quantitative tests to compare the time required to create the animations with and without the proposed method.

6. DISCUSSION

The completion of this research will contribute primarily to the field of computer animation, but it will also have repercussions in artificial intelligence. It will lead to a deeper knowledge on how to generate facial animation that results from external stimuli. By allowing a significant reduction in the time required to create facial animation, it will have great impact not only in casual videogames and films, but also other interactive applications such as dialogue systems and in serious games. We hope to introduce an alternative to the traditional facial animation techniques that will become the basis of new methods and studies. Thus, allowing both experienced and non-experienced authors to create believable and empathic facial animation.

7. ACKNOWLEDGMENTS

I would like to thank: Verónica Orvalho and Miguel Sales Dias for supervising my research, Instituto de Telecomunições for hosting my PhD and the Portuguese Foundation for Science and Technology (FCT) for fully funding the PhD research under the scholarship: SFRH/BD/79905/2011.

8. REFERENCES

[1] T. D. Bui, D. Heylen, M. Poel, and A. Nijholt. Generation of facial expressions from Emotion using a Fuzzy Rule Based System. *IN: Lecture Notes in Artificial Intelligence 2256*, pages 83 – 94, 2001.

[2] A. Egges and N. Magnenat-Thalmann. Emotional Communicative Body Animation for Multiple Characters. In *First InternationalWorkshop on Crowd Simulation*, pages 31–40, 2003.

[3] P. Ekman and W. Friesen. Facial Action Coding System. *Consulting Psychologist Press*, 1978.

[4] M. Gillies and B. Spanlang. Comparing and Evaluating Real Time Character Engines for Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 19(2):95–117, 2010.

[5] M. Kitagawa. Procedural animation. http://atec.utdallas.edu/midori/Handouts/procedural_animation.htm, 2011.

[6] K. Perlin. Layered compositing of facial expression. In *ACM SIGGRAPH 97 Visual Proceedings: The art and interdisciplinary programs*, pages 226–227. ACM Press, 1997.

[7] K. Perlin and A. Goldberg. Improv: A system for scripting interactive actors in virtual worlds. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques - SIGGRAPH '96*, pages 205–216. ACM Press, 1996.

[8] R. B. Queiroz, M. Cohen, and S. R. Musse. An extensible framework for interactive facial animation with facial expressions, lip synchronization and eye behavior. *Computers in Entertainment*, 7(4):1, 2009.

[9] H. Van Welbergen, B. J. H. Van Basten, A. Egges, Z. M. Ruttkay, and M. H. Overmars. Real Time Animation of Virtual Humans: A Trade-off Between Naturalness and Control. *Computer Graphics Forum*, 29(8):2530–2554, 2010.

[10] Y.-L. Xue, X. Mao, Z. Li, and W.-H. Diao. Modeling of layered fuzzy facial expression generation. pages 243–252, 2007.

[11] G. N. Yannakakis. Game AI revisited. In *Proceedings of the 9th conference on Computing Frontiers - CF '12*, page 285. ACM Press, 2012.

[12] V. Zammitto, S. DiPaola, and A. Arya. A methodology for incorporating personality modeling in believable game characters. In *Proceedings of Game Research and Development - Cybergames*, 2008.