Intelligent facial animation: Creating emphatic characters with stimuli based animation

Abstract:

The purpose of this PhD research is to define a novel method for real-time, stimuli based, procedural facial animation with cinematographic quality. In procedural methodologies the artist does not control the animation directly, only provides the initial parameters and defines a set of rules, which are used to generate the animation. Procedural animation has been used for body animation, while the face remains an unexplored area, where traditional techniques prevail. Conventional facial animation methods include: key frame animation, which consists on having an artist create all the (facial) poses by hand, therefore becoming an extremely time consuming task and performance-based techniques that can speed up the animation process but still require cleaning the motion data manually.

Thus, generating facial animation automatically is a demand but also a challenge, since it is very difficult to predict all possible facial movements, becoming even harder if the character is supposed to react to external stimuli.

The key contributions of this research will be: a procedural animation method, an intelligent agent that will automatically react to external stimuli, a full study of the reactions caused by these stimuli and how they affect the face. As a result there will be a reduction of the time required to create a facial animation. It will have great impact in both films industry and immersive systems such as videogames and medical systems, by reducing the time required to animate a character and creating interactive, highly empathic human-like characters.

Detailed Description:

In this thesis the main research challenges to overcome are: which are the triggers to generate a specific expression, how this expression is created and how to automatically do generate animation. Currently there is no standard to how an external event triggers an action in real life applied to facial animation. One possible case could be to pinch someone on the face. What would the facial expression of that person look like? Would he/she cry? Become angry? Both? This depends not only on personal factors, such as age and emotional state, but also the context of the event. Was it a friend? Were they fighting? Or maybe playing a game? Also, how do the context, personal factors and triggers contribute to creating the pose? Although the deformations that occur on the model, as a result of a direct interaction (i.e. the pinch), have already been properly explored, the afterwards reaction has not. This is due to the complexity of the task and lack of methods for easy animation of the aftermath.

The overall goal is to create a system capable of reacting to external stimuli and generate proper facial animations in an automatic way, thus creating a procedural facial animation method. To achieve this goal the research plan is divided into three stages:

S1: To study and understand how different stimuli lead to the creation of different poses and create a methodology based on how these stimuli affect facial movements;

S2: To create a procedural method with S1 as a basis capable of generating, with few parameters, any expression.

S3: To research intelligent agents capable of using the data of the studies performed in S1, to simulate human-like reactions and combine the agent with the results of S2

S4: To validate the whole system.

S1: Creation of a methodology to represent facial movements: The methodology will be based on voluntary and involuntary movements. To create the rules that define the methodology it is necessary to understand how psychological and physiological reactions to a stimulus affect the face. A similar study has already been made for body animation [1], but as far as it is known none exists for the face. Research will be done by analysing how the poses are created from psychological, psychiatric and neurological points of view. Regarding voluntary movements, they are the result of the cognitive functions of the brain e.g. speech or an expression. Involuntary movements are the result of the autonomous areas of the brain, which result of an illness or a psychological factor such as a nervous tic.

To better understand how to create the set of rules, the FACS [2] system will be analysed. This system can be used to animate a face even though it was developed from a psychological perspective. It defines a set of action units (basic element), which in turn are associated to sets of muscles. This system limits the creation of different poses and does not have the notion of motion as procedural animation requires. Facial movement will also be analysed from a psychological point of view. The goals of this step are:

$S_{1.1:}$ To study trigger-reaction causality pairs, initially for controlled cases and then generalize the study;

$S_{1.2:}$ Based on S1.1, identify involuntary and voluntary movements, caused by different reactions;

$S_{1.3:}$ To define a set of rules for facial movements, based on $S_{1.2}$, thus creating the facial animation methodology.

S2: Create the facial procedural animation method: The method will be based on S1.3. It will allow high-level commands based on the defined rules and will be the interface between the animator and the method. To achieve this method the current state of both, facial animation and body procedure animation will be studied and used as a layer to implement the rules (e.g. use a muscle system and on top of it create the rule system). Using this method, it will be possible to create and change facial animations. The goals are:

$S_{2.1:}$ To study full body procedural animation

$S_{2.2:}$ To create a method for procedural facial animation based on $S_{1.3}$ and $S_{2.1}$

S3: Create an intelligent agent that is capable of reacting to external stimuli. After creating the animation procedure, a layer of behaviour will be added with the purpose of creating a completely autonomous agent. Artificial intelligence methods (for instance reinforcement learning methods) will be studied and used as a basis for creating the agent. By adding an intelligent agent to the procedural method, it is possible to create a fully automatic animation agent, where the animator's role is replaced by the agent's. The agent is capable of reacting to external stimuli, passing the action to the animation procedure, which will allow creating animations on the fly. This agent will be used to automatically animate the talking head producing a human-like behaviour. The main goals are:

$S_{3.1:}$ To study artificial intelligent agents

$S_{3.2:}$ To create an intelligent head and apply to it the procedural facial animation method. This step is based on $S_{2.2}$ and $S_{3.1}$

S4: Validation. The final step is to validate the whole system through qualitative and quantitative methods. The qualitative validation will be based on empirical methods (observation and enquiries) and will be made for both artists and viewers. The tests will be the visual quality of both: a comparison between traditional animation and the system's quality and ability of the character to convey emotions based on subtle expression. The quantitative tests will also be based on the same comparison and will include: the time required to create animations, the reaction time of the character and a comparison between the muscles moved by the system and the ones a person would move in a similar situation amongst others. The goals are:

$S_{4.1:}$ To validate the system in a qualitative and quantitative way

Objectives:

The application created with the developed system will be a talking head (face model animated with speech) that automatically reacts to external stimuli, behaving in a credible way (i.e. resulting animations are understood by the viewer and resemble to what a person would do). By the end of this thesis the following advances will have been accomplished:

$A_{1:}$ To carry out research leading to the definition of a schema for external stimuli-reaction causality pairs.

A_2: To carry out research leading to a set of rules that defines facial movements.

$A_{3:}$ To define a real-time procedural facial animation method.

$A_{4:}$ To explore different models that will lead to the creation of intelligent agents capable of simulating human-like behaviour to external stimuli.

A_5: To create and evaluate a talking head prototype, which automatically reacts to stimuli.

$A_{6:}$ To validate previous advances.

$A_{7:}$ To disseminate results to the general public and scientific community in conferences such as SIGGRAPH, SIGCHI and EuroGraphics.

State of the art:

Realistic facial animation is a demanding task and although research has been made to improve and ease the creation of animations, there is currently no standard to automatically achieve this. Different facial poses can be created, without using specialized hardware, by manipulating the face's geometry, using techniques [3] such as blend shapes and FFD or a muscle based system. These poses are then used to animate the face. Creating realistic expressions is an extremely important step in facial animation, as it requires the animator to create, not only the overall shape, but all the subtleties associated with it. These subtleties are what differentiate a believable, engaging character from one that is not able to properly convey feelings to the viewer/player. [4, 5] proved that both the body and face's subtle movements positively influence the immersion levels [6] of the viewer. However, a complete study on subtle movements does not exist, thus leading to a problem when animating a face. When creating the different poses, the artist solves this problem not only with skills but also with a fair amount of taste and experience. This only makes the problem of automatically generating an expression worse.

According to Midori Kitagawa, professor at UT Dallas, in procedural animation, the animator defines a set of rules used to generate the animation. The rules can be defined by artificial intelligence (AI) algorithms, physical laws[and mathematical expressions, or any combination of these. AI can be used to automatically select a pose after a certain event occurs. Physical laws can be used to simulate the movement of a body falling. Mathematical expressions can be used to create a circular movement of an arm. Full-body procedural animation already exists and is used in commercial applications, such as Natural Motion's Euphoria [7]. Research on this matter was done in [8, 9, 10]. [11] studies how to achieve natural body animation based on voluntary and involuntary movements but no similar study was made to the face.

Some research has already been made on how to automatically animate a face. [12] describes a system capable of creating different expressions based on three parameters. [13] proposed a method based on different layers of animations. Both systems allow automatic creation of poses at a specific moment, whereas a complete facial procedural animation requires the notion of movement over time (e.g. tics). A basic procedural animation system has been developed [14], but it does not allow the manipulation of all facial muscles. [4] describes a system capable of automatically producing natural head movements (nod) synchronized with speech. [15] researched how to animate faces and facial behaviour based on text analysis. Behaviour simulation (AI) has been used for body animation however it has not been properly applied to the face even though some studies have already been made.

Research on AI has been done in the creation of agents capable of simulating human thinking and learning [16] (to some extent). [1] uses various psychosocial models to ease the modelling of human behaviour. Maxine Engine [17] analyses multimodal input (mainly speech and video) to produce an answer (speech) with an emotional state. Although this method shows promising results, it does not allow to automatically alter an animation nor does it permit the creation of any expressions.

Automatic facial animation is in its beginning and this is where this thesis will contribute decisively.

Repercussions:

Several repercussions are expected by studying the science of facial animation for procedural animation.

At a scientific level these include:

R1 - To lay the foundations for studies on virtual human reactions, leading to a deeper knowledge of how stimuli affect facial movements;
R2 - A new methodology to represent facial movements that will become the basis of new facial animation techniques and studies.

At a technological level:
R3 - A new method for procedural facial animation with cinematographic quality, allowing the creation/changing of animations in real-time, based on an intuitive language;
R4 - A new AI algorithm for facial animation based on stimuli that allows the creation of more interactive characters.

At a clinical level:
R5 - Include the developed technology into the FCT/Austin LIFEisGAME project, which leads to the creation of agents that allow the children to see which actions cause the pretended expressions is a novelty within this area.

At an industrial level, film companies, such as Pixar and DreamWorks Animation, and videogame companies, such as Rockstar, Electronic Arts and Square Enix, will be the ones that benefit the most as direct result of the reduction of production times and the quality of the animation.

Ref

[1] You, J., Katchabaw, M. 2010. A FLEXIBLE MULTI-MODEL APPROACH TO PSYCHOSOCIAL INTEGRATION IN NON PLAYER CHARACTERS IN MODERN VIDEO GAMES. In Futureplay '10 Proceedings of the International Academic Conference on the Future of Game Design and Technology. ACM, New York, NY, 17-24.
[2] Ekman, P., W. Friesen V. 1975. Unmasking the Face. CA: Consulting Psychologists Press. Palo Alto.
[3] PARKE, F. I., AND WATERS, K. 2008. Computer Facial Animation(Second Edition). A. K. Peters. Chapter 5
[4] Yan, J., Agada, R. 2010. Life-Like Animated Virtual Pedagogical Agent Enhanced Learning. Journal of Next Generation Information Technology. INSPEC, United Kingdom and Scotland, 1, 2, 4-12.

[5] Prendingera, H., Morib, J., Ishizukab, M. 2005. Using human physiology to evaluate subtle expressivity of a virtual quizmaster in a mathematical game. International Journal of Human-Computer Studies. ACM, New York, NY, 62, 2, 231 – 245.
[6] Brown, E., Cairns, P. 2004. A grounded investigation of game immersion. In CHI '04. ACM, New York, NY, 1297-1300
[7] Natural Motion 2011. Euphoria. http://www.naturalmotion.com/euphoria. Last visited June 2011
[8] Horswill, I.D. 2009. Lightweight Procedural Animation With Believable Physical Interactions. In IEEE Transactions. 1, 1(Computational Intelligence and AI in Games), 39-49.
[9] Cassell, J., Vilhjálmsson, H. H., Bickmore, T. 2001. BEAT: the Behavior Expression Animation Toolkit. In SIGGRAPH '01 Proceedings of the 28th annual conference on Computer graphics and interactive techniques. ACM, New York, NY, 477-486.
[10] Faloutsos, P., van de Pane, M., Terzopoulos, D. 2001. The Virtual Stuntman: Dynamic Characters with a Repertoire of Autonomous Motor Skills. In Computers and Graphics vol. ELSEVIER, Amsterdam, 25, 6, 933-953.
[11] Chi., D., Costa, M., Zhao, L., Badler, N. 2000. The EMOTE Model for Effort and Shape. In SIGGRAPH '00 Proceedings of the 27th annual conference on Computer graphics and interactive technique. ACM, New York, NY, 173-182.
[12] Xue, Y., Mao, X., Li, Z., Diao, W. 2007. Modeling of Layered Fuzzy Facial Expression Generation. In Digital Human Modeling. Springer, Berlin, 243-252.
[13] Perlin, K. 1997. Layered Compositing of Facial Expression. Layered compositing of facial expression. In SIGGRAPH '97 Visual Proceedings. ACM, New York, NY, 226 - 227
[14] The Expression Toolkit 2004. http://expression.sourceforge.net/. Last visited June 2011
[15] Slinger, P., Etemad, S. A., Arya, A. 2009 Intelligent Toolkit for Procedural Animation of Human Behaviors. In Proceedings of the 2009 Conference on Future Play. ACM, New York, NY, 27-28.
[16] Johnson. W. L., Rickel, J. W. 2000. Animated Pedagogical Agents: Face-to-Face Interaction in Interactive Learning Environments. In International Journal of Artificial Intelligence in Education. 11,47-78
[17] Baldassarri, S., Cerezo, E., Seron, F. J. 2008. Maxine: A platform for embodied animated agents. In Computers & Graphics, ELSEVIER, Amsterdam, 32,4, 430– 437.

MAP-I 2011-12

Thesis Acceptance Declaration

This document serves as acceptance statement of the thesis proposal entitled as "intelligent facial animation: Creating emphatic characters with stimuli based animation" supervised by Prof. Verónica Orvalho and co-supervised by Prof. Miguel Sales Dias, by the MAP-I student José Mário Figueiredo Serra

2nd March 2012

José Mário Figueiredo Serra

José Mário Figueiredo Serra

Verónica Costa Orvalho

Miguel Sales

Miguel Sales Dias