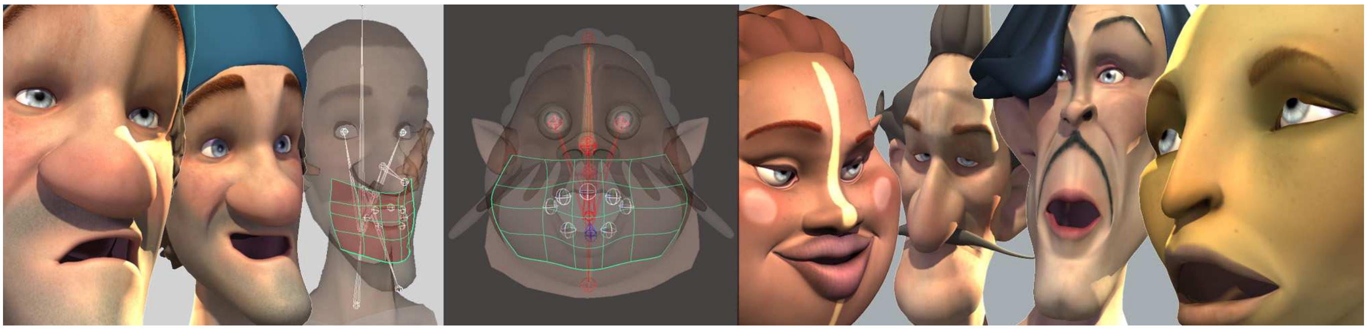
Verónica Costa Teixeira Orvalho

Reusable Facial Rigging and Animation: Create Once, Use Many



(Copyright 2005 Dygrafilms)

A DISSERTATION

in

Software (Computer Graphics)

Presented to the Faculties of the *Universitat Politècnica de Catalunya*

in Partial Fulfillment of the Requirements for the *Degree of Doctor of Philosophy*

Advisor: Dr. Antonio Susin
Tutor: Dra. Isabel Navazo

Barcelona, June 2007

Abstract

Facial animation is a serious bottleneck in any computer generated (CG) production. It is the key element to convey emotion to 3D characters. Speeding up the rigging process remains an unsolved problem, specially for film and videogames, which require high quality results. The character rigging is analogous to setting up the strings that control a puppet. Today, skilled artists manually create the facial rig to ensure the best quality in the animations; but, this is a slow, labor-intensive and costly process.

This thesis presents a portable character rigging system that integrates into current animation production pipelines. It automatically transfers the facial rig and animations created for one character to different characters, independent of their shape and appearance. It enables artists to create more lifelike facial models in less time; about 90-99 percent faster than traditional manual rigging. Characters can display complex expressions and behavior, easier and with decreased artistic effort. As a result, we dramatically reduce the time needed to create high-quality facial animations for the entertainment industry.

We studied techniques from the fields of computer graphics and computer vision, to come up with a solution to the rigging problem. Based on a generic facial rig definition and a new deformation method, our system converts 3D face models into digital puppets that experienced artists can control. The system adapts the skeleton, weights and influence objects (NURBS surfaces, lattice, etc.) from a source rig to individual face models to obtain unique expressions, and enables easy reuse of existing animation scripts. Our work differs from previous morphing and retargeting techniques, because that work was oriented towards transferring animations, while ours aims to transfer the complete facial rig, in addition to animations.

The system was validated with a series of experiments. We used models and rigs from major film and videogame companies: Electronic Arts, Radical, Dygrafilms. The results were supervised by Technical and Art Directors, who approved the quality of our rigs and animations to be used in CG productions, replacing the artist generated ones. Our proposal is: generic (the facial rig can have any type of configuration and models can be created by artists), flexible (the rig has no initial constraints), independent of the shape and appearance of the model, and enhances the freedom of the artists (does not force the use of a predefined rig). Now, motionless and inanimate faces can come to life with the help of our technology.

Acknowledgements

This long walk started five years ago, when I took the first step to follow my dreams of working within the entertainment industry. But the journey would have not been possible without the support and friendship of many people. I would like to thank them with all my heart for helping me accomplish this goal.

First, I express my deepest gratitude to my advisor Toni Susin, who rescued me from the bottom of the sea. He was always a great source of encouragement and guidance. Also Isabel Navazo, for providing academic advice.

A very special thanks to Ernesto Zacur, which helped at the early steps of my research by sharing his work on facial reconstruction with me. His invaluable feedback inspired many of the current features of the system.

Juan Nouche and everyone at Dygrafilms, for seeing the potential of the technology at an embryonic stage and believing in me. Thanks for all the time, dedication and material so important to validate the system. I am specially grateful to Xenxo Alvarez, who dedicated many many hours to thoroughly test the application and coordinated the creation of the 3D models and animations; his suggestions have greatly improved the system. I extend my thanks to his team: Rubén Lopez, Juan Ramón Pou, Esteban Ferreño, Maria Lodeiro, Cristina Conde and Cristian Parrado.

Miguel Angel Sepúlveda for helping me narrow down the research focus and for his initial orientation.

Kelly Booth from UBC (University of British Columbia) for his hospitality (Tojo's was delicious!) and dedication in putting me in contact with so many people that later helped my research. Sid Fels for making my stay at UBC's HCT lab not only fruitful but also fun.

Jean-Luc Duprat (then at Electronic Arts), for his friendship (and Belgium beers), for organizing interesting meetings and for giving me very good advice on how to approach the entertainment industry. Fred Fowles and Crystal Wang from EA, for always having time to answer my questions, provide me with high quality test material and their constant words of motivation.

Dave Fracchia (then at Mainframe Entertainment) for sharing his insightful knowledge on the film industry, which helped me define my carrier goals, and his tips on how to improve my demos.

Juan Pablo Buscarini, who gave me the opportunity to get started in the film industry. He inspired my current path.

My friends all over the world, who reminded me from time to time that real life goes on outside the silicon chips. And my godmother Raquelita, for helping me become a better person each day.

My husband's extended family, for making me part of their lives.

My parents, for always making sure that I received the best education. My sister, because I love her.

Last, but the most, I deeply thank my husband João, for being my source of inspiration and everyday happiness. For his unconditional love and support. For keeping me going in the darkest hours. For all the days he spent fixing my clumsy writing. For giving me constant feedback to improve the research. For reviewing my work, probably as many times as I did. And much much more. Without him, everything would be different...

To João Orvalho for his endless love.

"Where we touch, the fibers merge
and intertangle.

I am no longer
certain where I
end...where
he begins..."
- Alan Moore, Swamp Thing (1990)

"When all I want is you..." - Bono (1988)

Contents

Chapter 1

Introduction

Facial Animation *is the key element to convey emotion and personality to a character. The entertainment industry requires high quality results and drives the research efforts to automate the character setup process for realistic animation of 3D characters. We studied techniques from the fields of computer graphics and computer vision, to come up with a solution that speeds up the rigging process dramatically. We propose a portable character rigging system that integrates into current animation production pipelines, enabling digital artists to create more lifelike characters in less time, when compared to traditional animation techniques. It automatically transfers the rig and animations created in one character to different characters, independent of their shape and appearance. This chapter briefly describes the main problems when preparing a character for animation, gives an overview of our solution and summarizes our research contribution.*

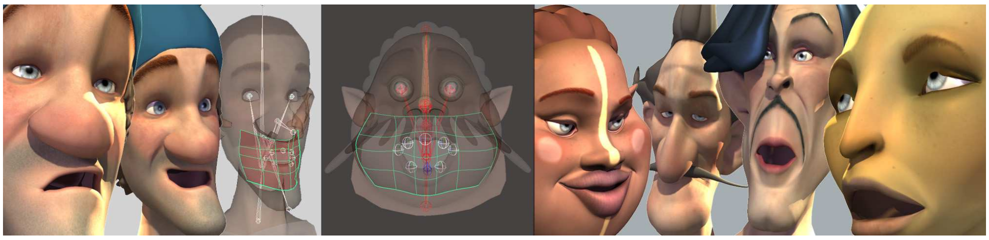


Figure 1.1: Source model rig and expressions (left); target model rig transferred from the source (middle); expressions from different target models extracted from an animation sequence (right). (Copyright 2005 Dygrafilms)

1.1 Motivation

The PhD research started in 2003 after several talks and inquiries with people in the USA digital animation film industry. We assembled a list of problems faced by film productions, and after careful analysis chose one of the most challenging: *automate the setup for realistic facial animation of 3D characters*. Facial animation is currently in great demand, because the entertainment industry requires high quality results ([Borshukov et al. 2006]), and has been an area of intensive research for the last three decades. Facial animation is crucial to convey the feelings of each character, so automating the character setup has the greatest impact on the final production output, comparing it with other unsolved problems. After the initial research stages to determine the viability of the project were completed with success, we focused on creating a prototype of the automation system to demo animation companies and professionals and get feedback and validation. The companies have showed great enthusiasm for the results and, more importantly, for the potential of the technology. From the start, we focused on finding a solution to current rigging difficulties that was simple in design, but still able to solve a hard problem.

To achieve the cinematographic quality we have, we had to consider lots of little details [Richie et al. 2005] to make a versatile system that can be easily integrated into different animation pipelines and work well with very dissimilar characters. Our work differs from previous morphing [Blanz and Vetter 1999] and retargeting techniques [Noh and Neumann 2001], because their work was oriented towards transferring animations, while ours aims to transfer the complete facial rig, in addition to animations. Automatically transferring the rig gives artists complete freedom to manipulate the characters: they can create new animations and not be limited by pregenerated ones.

Facial animation presents many difficulties (time, cost and complexity constraints) that limit its adoption and usefulness in different situations. [Pighin et al. 2006] discuss the research efforts and main challenges faced by some blockbuster films, and emphasize that facial puppetering and the use of non-linear rigs are still unexplored issues. Generating realistic face movements is hard, because even with current 3D software, animators cannot capture and control every detail of the face. To obtain the desired realism, traditional animation pipelines have each character separately *rigged by hand*, a very labor-intensive and time-consuming task. The character rigging process is analogous to setting up the strings that control a puppet, which in the hands of an experienced digital artist comes to life. Finding a technique that provides accurate and fast rigging remains a challenge.

what if... we want to use the rig created
for one character in other characters?

To create a **portable character rigging system** capable of synthesizing the subtle dynamics of facial expressions expected by human observers, we need to solve several *artistic* and *technological* problems. Artistically, we need to have a clear knowledge of the human familiarity and sensitivity to facial appearance, and the multitude of subtle expressive variations on the face. On the technology side, first, we have to reduce the amount of laborious work needed to create and optimize each facial rig. Second, the weight distribution used on one character will not work for others. Last, any minor artistic modification that causes the model to change in geometry (a smaller nose, a larger eyebrow) or resolution (more resolution around the eyes) leads to the restarting of most of the rigging process from scratch. These issues make it impossible to reuse the same rig in different face models; facial animation becomes a bottleneck in any CG production.

1.2 Our Solution

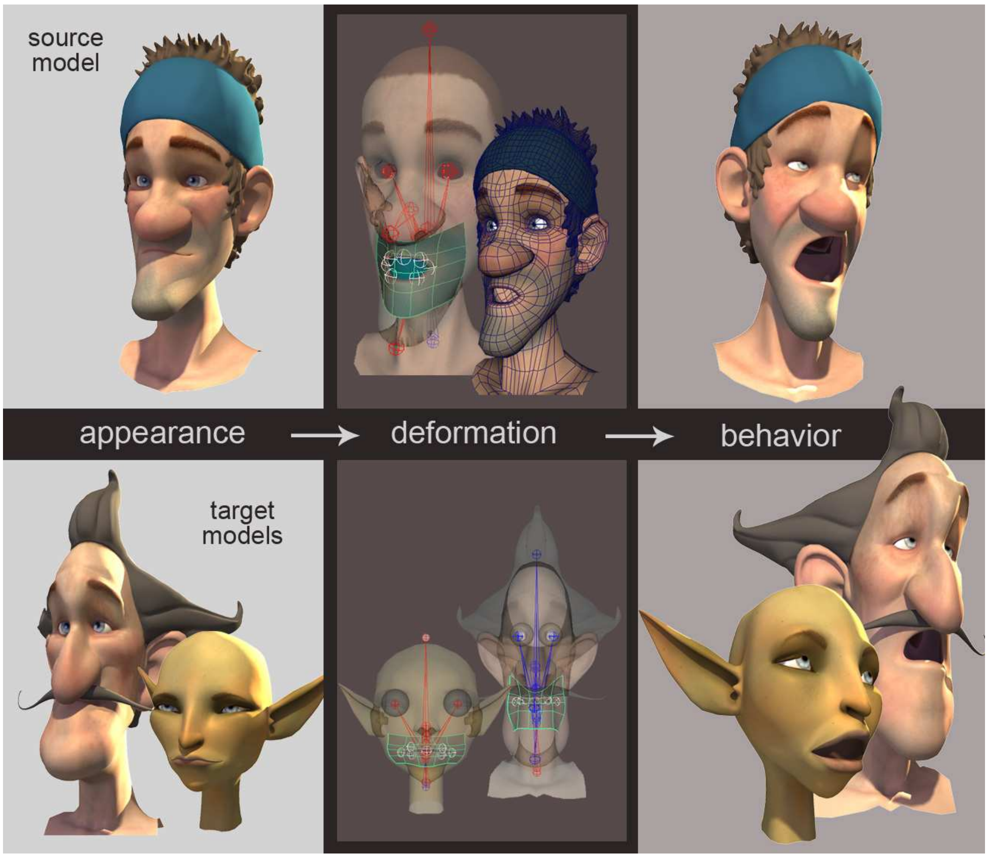


Figure 1.2: Overview: define the source and target models; adapt the source model geometry to fit the target; transfer attributes and shapes; bind the influence objects and skeleton to the target. The result is a model ready to be animated. (Copyright 2005 Dygrafilms)

We narrow our research to the rigging of 3D animation characters for films and videogames, solving the synchronization and realism problems, the reusability of facial components, together with the real time response. However, the results from our project can benefit other industries as well.

We propose a system that *given a face model, analyzes it and creates a rig ready to be animated.* Our solution starts with two 3D face models: the source model, which is rigged and includes a set of attributes (skeleton, influence objects, shapes and animation scripts), and the target model, which doesn't have a character rig associated to it (see figure 1.2). Then, the system automatically transfers the rig and animations from the source to a target model, even if they have different geometries or descriptors (one can be defined as a polygonal mesh and the other as a subdivision surface). The method follows the principals of a non-linear warp transform [Bookstein 1989], and uses facial features landmarks, resulting in an efficient deformation technique. Our solution can build models with underlying anatomical behavior, skin, muscle and skeleton, for human, cartoon or fantastic creature heads. It allows autonomous and user controlled facial features to move naturally, smoothly and realistically.

The system enables artists to create more lifelike characters in less time; it is 90-99 percent faster than traditional manual rigging. The tests showed that artists can complete in one hour, tasks that before took them one or two weeks of work; something like changing the weights, modifying an animation control position or transferring animations between characters can be achieved “instantly”.

1.3 Contribution

Current rigging techniques are slow and expensive, because they rely on traditional hand setup. Alternative systems like Eyetronics[1], Universal Capture [Borshukov et al. 2005] and Image Metrics[2], combine different technologies, like image analysis and optical flow, to achieve amazing results, but still require too much time to setup. Also, applying changes and propagating the modifications to different characters is hard and needs cleaning up afterwards, making these approaches unsuitable to be used in films and videogames with lots of characters, at least in a short term.

A major benefit of our solution is that the system is very flexible: the deformation and transfer methods are general and can be used with any type of facial rig setup and 3D model topology. Artists can define their own rig and then quickly apply it to different models, even with disparate proportions and appearance (human, cartoon or fantastic). This allows easy integration in different animation pipelines and reusability in different projects. The key contributions of this dissertation to the field of computer graphics and the entertainment industry are:

- a **generic geometric deformation method**: based on facial features landmarks, deforms and reshapes the source face model surface (polygonal mesh, NURBS, subdivison) into the target surface;
- a **facial rig transfer method**: based on the generic deformation method, relocates and adapts all attributes (texture, weights, joints, influence objects) of the source model rig into the target model. The method allows reusing a rig created for one model in different characters;

[1] http://www.eyetronics.com
[2] http://www.imagemetrics.net

- an **expression transfer method**: based on different key poses of the source model, computes the same poses in the target model, but adjusted to its shape and proportions. These poses are translated into shapes and can be used for blend shape animation;

- an **animation transfer method**: adapts animation curves from the source model to different characters, adjusting the scale, rotate and translate value to the shape and proportions of the target model.

As a result, film and videogames studios will be able to animate characters that before were devoid of expression, because of the time it takes to manually create a rig. Suddenly, all those secondary characters can come to life with the help of our technology.

1.4 Outline

The remaining chapters of the dissertation are organized as follows:

Chapter 2 describes the use of facial animation in the context of visual communication, focused on the entertainment industry. It details the character rigging process and the different components that constitute a facial rig. The chapter ends with a description of the role of the character Technical Director.

Chapter 3 discusses different methods related to facial synthesis: physically-based, geometric deformation, keyframe animation, performance-driven and retargeting. It briefly describes facial parametrization, MPEG-4 standard and Facial Action Coding System (FACS). Last, the chapter lists existing animation applications.

Chapter 4 describes our solution for reusing the same facial rig in different face models. It describes the algorithms we developed and the system architecture.

Chapter 5 shows the results and data obtained from tests with several film and videogame companies. It discusses the visual and timing validations performed, comparing the automatic output of our system with the results manually created by an artist. It also summarizes our conversations with people from the entertainment industry.

Chapter 6 discusses our work, its implications and future trends.