Pedro Tavares Barata Bastos

EASY FACIAL RIGGING AND ANIMATION APPROACHES

A dissertation in Computer Graphics and Human-Computer Interaction

Presented to the *Faculty of Engineering of the University of Porto* in Partial Fulfillment of the Requirements for the *Degree of Doctor of Philosophy* in Digital Media

Supervisor: Prof. Verónica Costa Orvalho

April 2015

QUALIFICAR É CRESCER.

QUADRO DE REFERÊNCIA ESTRATÉGICO NACIONAL
PORTUGAL 2007.2013

GOVERNO DA REPÚBLICA PORTUGUESA

UNIÃO EUROPEIA
Fundo Social Europeu

This work is financially supported by Fundação para a Ciência e a Tecnologia (FCT) via grant SFRH/BD/69878/2010, by Fundo Social Europeu (FSE), by Ministério da Educação e Ciência (MEC), by Programa Operacional Potencial Humano (POPH), by the European Union (EU) and partially by the UT Austin | Portugal program.

Abstract

Digital artists working in character production pipelines need optimized facial animation solutions to more easily create appealing character facial expressions for off-line and real-time applications (e.g. films and videogames). But the complexity of facial animation has grown exponentially since it first emerged during the production of Toy Story (Pixar, 1995), due to the increasing demand of audiences for better quality character facial animation. Over the last 15 to 20 years, companies and artists developed various character facial animation techniques in terms of deformation and control, which represent a fragmented state of the art in character facial rigging. Facial rigging is the act of planning and building the mechanical and control structures to animate a character's face. These structures are the articulations built by riggers and used by animators to bring life to a character. Due to the increasing demand of audiences for better quality facial animation in films and videogames, rigging faces became a complex field of expertise within character production pipelines. The demand for more quality has contributed to increase the productivity in this field but also to disperse it and deprive it of a consensus on how to build and provide the mechanics and controls of a facial rig. Today, facial rigging can be considered fragmented because it is an extremely laborious, time-consuming, disparate and sometimes frustrating process for both riggers and animators.

Aware of the problems described previously, the author presents in this thesis optimized rig approaches for facial animation that help facilitate and empower the artists workflow, by presenting a guide to place and configure the rig mechanical components and a flexible rig interface to control a character's face. The optimization is oriented to decrease the complexity of the facial rigging process for a human subject facial model as a basis for other facial styles and also for key frame facial animation, a technique which is in its peak of development because it continues to be popular, despite the existence of more recent techniques like motion capture, therefore key frame is highly prone to be optimized. The research steps carried out by the author include an extensive study of the facial rigging techniques used by several artists worldwide in this field, through literary as well as approaches posted in the web, and then the definition of facial rigging and animation approaches to help ease the job of the rigger and the animator. The approaches presented are based on generic and validated guideline proposals to build facial rig mechanics and on a flexible facial rig control system based on a multimodal interface. The approaches developed were validated initially in a pilot study with three users, followed by four detailed user experiments, respectively involving five, fifteen, twenty and twenty two users. The facial rig approaches have been approved and considered by the users to (i) ease the construction process of the mechanisms that deform a facial model and (ii) ease the use of the control interfaces that allow animators to trigger facial expressions. The results have been published in a sequence of papers which provide a coherent storyline to this dissertation, a storyline which is meant to help bring a consensus to the state of the art in facial rigging and animation and, hopefully, decrease the fragmentation found in this research and production field.

Key-words: Faces, Rigging, Animation, Interaction, Interfaces.

Resumo

Os artistas digitais que trabalham em desenvolvimento de personagens precisam de soluções otimizadas de animação facial para mais facilmente criar expressões faciais apelativas para aplicações em *off-line* e em tempo-real (i.e. filmes e jogos). Mas a complexidade da animação facial cresceu exponencialmente desde que surgiu na produção de Toy Story (Pixar, 1995), devido à maior exigência do público por animação facial com mais qualidade. Nos últimos 15 a 20 anos, empresas e artistas desenvolveram uma diversidade de técnicas de animação facial em termos de deformação e de controlo, que representam um estado da arte fragmentado em *rigging* facial de personagens. O *rigging* facial é o ato de construir as estruturas mecânicas e de controlo para animar a cara de uma personagem. Estas estruturas são as articulações construídas por *riggers* e usadas por animadores para dar vida às personagens. Devido a uma maior procura por animação facial de qualidade, o *rigging* de caras tornou-se gradualmente mais complexo nas produções profissionais de personagens. A procura por mais qualidade contribuiu para aumentar a produtividade nesta área mas, também, para a dispersar e privar de um consenso sobre como construir e disponibilizar as mecânicas e controlos de um *rig* facial. Atualmente, o processo de *rigging* facial pode-se considerar fragmentado porque é trabalhoso, demorado, disperso e, por vezes, frustrante para *riggers* e animadores.

Consciente das questões descritas, o autor apresenta, nesta tese, abordagens otimizadas de *rigs* faciais que ajudam a facilitar e fortalecer o trabalho dos artistas, através de um guia para colocação e configuração das mecânicas de um *rig* facial e de uma interface flexível de *rig* para controlar a cara de uma personagem. Esta otimização pretende diminuir a complexidade do processo de *rigging* facial de um modelo 3D humano enquanto base para outros estilos faciais e está orientada à animação facial por *key frames*, uma técnica de animação que está no seu auge de desenvolvimento por continuar a ser popular e muito recetiva a ser otimizada, apesar da existência de técnicas mais recentes como o *motion capture*. As etapas de investigação realizadas incluem um estudo extensivo das técnicas de *rigging* facial desenvolvidas por artistas de todo o mundo, através de referências literárias e de abordagens disponíveis na internet, e a definição das abordagens de *rigging* e animação facial facilitadoras do trabalho do *rigger* e do animador. As abordagens apresentadas baseiam-se em propostas de instruções genéricas e validadas para construir as mecânicas de um *rig* facial e ainda numa interface multimodal para controlo da animação facial. As abordagens desenvolvidas foram validadas inicialmente através de um estudo piloto com três utilizadores seguido por quatro experiências, respetivamente com cinco, quinze, vinte e vinte e dois utilizadores. Os utilizadores aprovaram e consideraram as abordagens como (i) facilitadoras do processo de construção das mecânicas que deformam um modelo facial e (ii) facilitadoras da construção e uso das interfaces de controlo para os animadores manipularem expressões faciais. Os resultados foram publicados numa sequência de artigos que proporcionam uma história coerente a esta dissertação, concebida para ajudar a providenciar um consenso ao estado da arte na área de *rigging* e animação facial e, esperançosamente, ajudar a diminuir a dispersão encontrada neste campo de investigação e produção.

Palavras-chave: Caras, Articulação, Animação, Interação, Interfaces.

Acknowledgments

The course of this investigation was demanding but also gratifying. It was only thanks to the support of several people that I was able to produce this thesis, of which I am proud and also recognize that I owe it not just to me but to everyone mentioned following.

To my supervisor, Professor Verónica Costa Orvalho, for accepting me as a doctoral student, for guiding me during my research and for always having a critical sense over my research progress.

To the friend and research colleague Xenxo Alvarez Blanco, Technical Director at the research group Porto Interactive Center, for sharing his knowledge in rigging and for his moral support.

To my other colleagues who I met in Porto Interactive Center for their fellowship.

I thank the president and the juries of my research proposal defense, Professors Eurico Manuel Carrapatoso, Karin Gwinn Wilkins and Rui Pedro Rodrigues, for presenting pertinent questions which were fundamental for the evolution of the research.

To Karen Gustafson for her support during my exploratory visit to Austin and to College Station in Texas in late 2011, where I met with Professors Frederic Parke and Tim McLaughlin, who contributed to help define the research pathway.

To Pedro Vieira, of the Center of Events of Faculdade de Engenharia da Universidade do Porto, for his help in requesting and configuring the equipments used in the recording of my facial expressions in late 2012.

To Professor Paulo Baeta Mendes for his vast experience in photography and for his equipment which he made available to photograph my face so that I could build and texture my 3D facial model.

To Luís Álvaro Fernandes and Sharon Lee Strover for their collaboration in the final results of this research.

To the Blender Foundation and the community of Blender digital artists for their continuous work in improving this digital tool.

To the digital artists who validated the solutions developed in this thesis: Benjamin Bays, Carlos Cidrais, Carmen Pereira, Eduardo Teixeira, Filipe Feijó, Francesco Siddi, Frutuoso Silva, Gustavo Rosa, Helder Lopes, João Alves, João Cesário, João Seabra, Jorge Ribeiro, José Maria Andrade, José Pedro Teixeira, José Ricardo Silva, Júlio Ramos, Leandro Monteiro, Luís Pereira, Nelson Gonçalves, Nuno André, Nuno Estanqueiro, Nuno Matos, Nuno Vasco, Paulo Cristóvão, Paulo Martins, Paulo Moreira, Pedro Ângelo, Pedro Mota Teixeira, Pedro Pereira, Quentin Geluyckens, Ricardo Megre, Rogério Perdiz, Rui Campos, Rui Guimarães, Rui Louro and Sarah Laufer.

To the Professors António Fernando Coelho, Artur Pimenta Alves, José Manuel Azevedo, José Torres Farinha and Maria do Amparo Monteiro for their moral support.

To my parents, Zélia Conceição and Augusto Manuel and to my brother Luís Filipe, for their unconditional and constant support.

The greatest of thanks goes to my wife, Denisa Mateus, for always advising me as only she knows how.

Agradecimentos

O percurso desta investigação foi exigente mas também muito gratificante. Foi apenas graças ao apoio de várias pessoas que me foi possível produzir esta dissertação, da qual me orgulho e também reconheço que a devo não só a mim mas a todos aqueles que menciono a seguir.

À minha orientadora, a Professora Verónica Costa Orvalho, por me ter aceite como seu doutorando, por me ter guiado durante a minha investigação e por ter tido sempre um sentido crítico sobre o meu progresso.

Ao amigo e colega de investigação Xenxo Alvarez Blanco, Diretor Técnico no grupo de investigação Porto Interactive Center, pela sua partilha de conhecimento em rigging e pelo seu apoio moral.

Aos meus outros colegas que conheci no Porto Interactive Center pelo companheirismo.

Agradeço ao presidente e arguentes da minha defesa da proposta de tese, os Professores Eurico Manuel Carrapatoso, Karin Gwinn Wilkins e Rui Pedro Rodrigues, pelas questões pertinentes que colocaram e que foram fundamentais para a evolução da investigação.

À Karen Gustafson pelo apoio prestado aquando da minha visita a Austin e a College Station no Texas no final de 2011, onde conheci os Professores Frederic Parke e Tim McLaughlin, que contribuiram para ajudar a definir a direção da investigação.

A Pedro Vieira, do Centro de Eventos da Faculdade de Engenharia da Universidade do Porto, pela sua ajuda na requisição e configuração dos equipamentos usados na gravação das minhas expressões faciais no final de 2012.

Ao Professor Paulo Baeta Mendes pela sua vasta experiência de fotografia e equipamento que disponibilizou para fotografar a minha cara para que eu pudesse construir e texturizar o meu modelo facial 3D.

A Luís Álvaro Fernandes e à Professora Sharon Lee Strover pela colaboração nos resultados finais desta investigação.

À Blender Foundation e à comunidade de artistas digitais de Blender pelo seu contínuo trabalho de melhoria desta ferramenta digital.

Aos artistas digitais que validaram as soluções desenvolvidas nesta tese: Benjamin Bays, Carlos Cidrais, Carmen Pereira, Eduardo Teixeira, Filipe Feijó, Francesco Siddi, Frutuoso Silva, Gustavo Rosa, Helder Lopes, João Alves, João Cesário, João Seabra, Jorge Ribeiro, José Maria Andrade, José Pedro Teixeira, José Ricardo Silva, Júlio Ramos, Leandro Monteiro, Luís Pereira, Nelson Gonçalves, Nuno André, Nuno Estanqueiro, Nuno Matos, Nuno Vasco, Paulo Cristóvão, Paulo Martins, Paulo Moreira, Pedro Ângelo, Pedro Mota Teixeira, Pedro Pereira, Quentin Geluyckens, Ricardo Megre, Rogério Perdiz, Rui Campos, Rui Guimarães, Rui Louro e Sarah Laufer.

Aos Professores António Fernando Coelho, Artur Pimenta Alves, José Manuel Azevedo, José Torres Farinha e Maria do Amparo Monteiro por todo o seu apoio moral.

Aos meus pais, Zélia Conceição e Augusto Manuel e ao meu irmão Luís Filipe, pelo apoio incondicional e constante.

O maior dos agradecimentos dirige-se à minha esposa, Denisa Mateus, por sempre me ter aconselhado como só ela sabe fazer.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Character facial animation has been a subject of research and production for the last four decades, since the pioneering work of Frederic Parke [Parke 1972], who built the first digital facial parametric model. Today, the entertainment industry of films and videogames thrives for greater quality facial animation for 3D characters, because the face plays a key role in transmitting a character's emotions to an audience [Orvalho 2007]. This dissertation is meant to help improve the jobs of the 3D digital artists who are responsible for preparing a face for animation and for actually animating a character's face within a character production pipeline in the entertainment industry – a process which is known in research and in the industry as facial rigging, the act of planning and building the mechanical and control structures to animate a character's face. The approaches presented in this thesis help provide a consensus to the fragmented state of the art in this field because they include (i) generic and validated guideline proposals to place and configure the mechanical components in a facial rig and (ii) a personalized rig interface to trigger facial expressions according to the user control preference. Together, these approaches help ease the laborious, time-consuming and startling process of creating, maintaining and using facial animation structures. This chapter begins by describing the problem and opportunity in this research, then it presents the solution and strategy used to overcome the problem, after that it presents the author's motivation and the contribution of this thesis to the fields of computer graphics and human-computer interaction and lastly it presents the structure of the dissertation.

1.1 Problem and Opportunity

Digital media such as films and videogames have a growing demand for better looking character facial animation. To manage the increasing demand, the character production pipelines in the entertainment industry worldwide have suffered a complexity boost in their animation setup processes in the last 15 to 20 years, since the film Toy Story (Pixar, 1995). Today, the goal of the 3D digital entertainment industry is to induce emotions to an audience through subtle and appealing character facial movements, because the face can "twist and pull into 5000 expressions" [Blum 1998]. But despite facial animation is now being further explored, an advent remains: the many hours of intensive labor carried out in a daily basis by digital artists to produce the quality facial animation seen in films and videogames like The Avengers (Joss Whedon, 2012) and The Last of Us (Naughty Dog, 2013), seen in Figure 1.1.

Figure 1.1: Examples of facial expressions in films and videogames; left: the face of the character Hulk in the film The Avengers (Copyright 2012 Marvel Studios); right: the face of the character Joel in the videogame The Last of Us (Copyright 2013 Naughty Dog).

The facial animation seen in the digital media illustrated in Figure 1.1 is only possible due to a labor-intensive and time-consuming process called facial rigging. Facial rigging is the act of planning and building the deformation mechanics and the interactive control interfaces to animate the face of a 3D character. The process of planning and building these structures is carried out by riggers, who are the artists responsible for providing to the animators with an articulated facial rig. Then the animator makes use of the controls in the facial rig to trigger the deformations in the character's face and cause it to reveal facial expressions and emotions. The job of rigging first emerged in the successful company Pixar [Price 2008] in the film Toy Story (Pixar, 1995), which is considered “the first full length feature film produced entirely using the technology of computer animation” [Henne et al. 1996]. At that time, the rigger was known as the builder, or the person who “decided on the set of controls over a character's articulation” [Henne et al. 1996]. Usually, there are only a few riggers in a character production pipeline of a film or videogame. The reason for this is that rigging is a complex job, because it involves preparing a character for animation, and only few artists have the required experience to provide to the animators with efficient interactive user interfaces to handle the diversity of facial deformations for a character. But because faces can have different morphologies and withstand a wide range of behaviors, rigging has become a labor-intensive, time-consuming and sometimes strenuous process, which is why it is only carried out by the most experienced artists. Paul Meyer, Managing Director at Luma Arcade, said recently: “producing quality animation does not happen overnight. It takes time, skill, experience” [Meyer 2013]. The fact is that the jobs of the rigging artists keep growing in complexity due to the increasing demand for better quality facial animation in films and videogames. Moreover, the constraints of short time and low budget common to these productions contribute for riggers to often produce a limited or feasible rig solution rather than an ideal one. In addition, this field lacks advanced and complete documentation and the references available do not provide validation on how to build a facial rig. Instead, they are limited to isolated cases, which generally do not encompass dealing with more complex facial behaviors. The former facts have plunged the facial rigging job into a disparate and fragmented process within character production pipelines. In fact, since Toy Story (Pixar, 1995), different facial rigging approaches have been developed by 3D digital artists worldwide for a diversity of characters. Figure 1.2 illustrates a few examples of facial rigging approaches developed by artists who have posted their facial rigs online in YouTube and Vimeo to show off their work in this field (see Sections 2.6 and 2.7 of Chapter 2 for an in-depth overview of the state of the art in facial rigging).

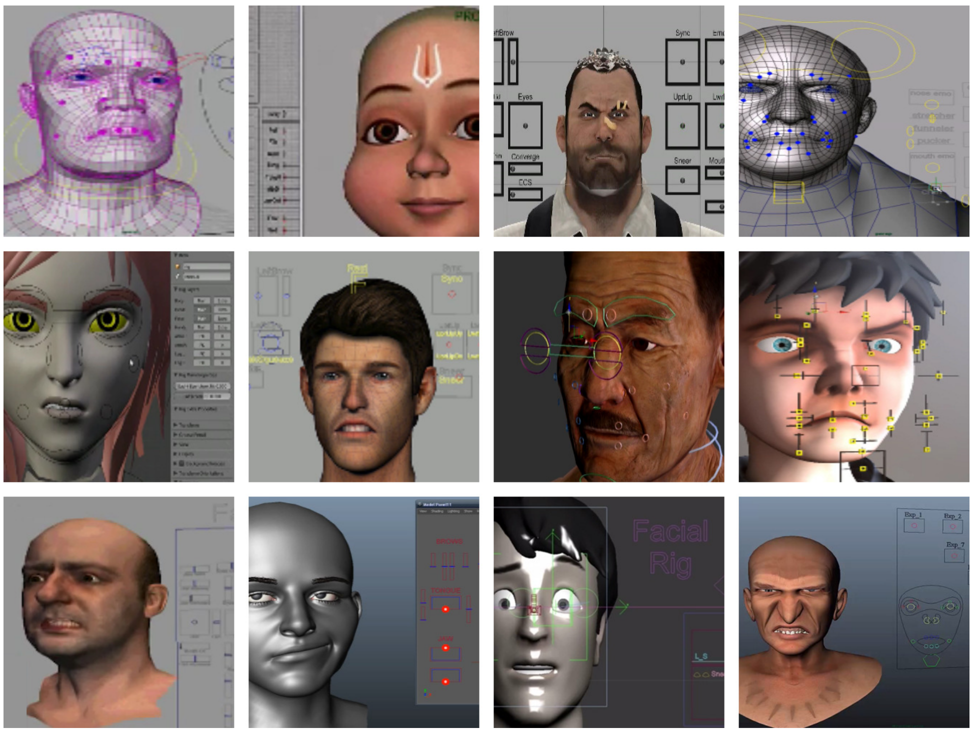


Figure 1.2: Examples of several disparate character facial rigging approaches developed recently by 3D digital artists (from 2007 until 2013); top row left to right: [Georgiev 2007, Avendra 2008, Cheung 2009, Labs 2010]; middle row left to right: [Vegdahl 2010, Neuffer 2011, Best 2011, Eren 2011]; bottom row left to right: [Vaucher 2011, Marmor 2011, Torres 2012, Kumar 2013].

The problem of the complexity and dispersion of facial rigging and animation opened up the opportunity to investigate in this research new ways to improve this field. But, this opportunity revealed difficult to make clear for the author, because facial rigging is complex and "open to significant improvements", as mentioned by computer facial animation pioneer

Professor Frederic Parke, of the Texas A&M University at College Station, during the author's exploratory visit to Austin in December 2011 under the UT Austin | Portugal program. Therefore, solutions for the facial rigging and animation process are indeed necessary, but they are also hard to uncover and to carry out efficiently. The solution and strategy devised by the author is described following.

1.2 Solution and Strategy

The initial goal of this thesis was to define a facial rigging standard, with the original title being Easy Facial Rigging: A Dynamic Industry Standard. In time, a major obstacle to this goal was identified in the review of the state of the art, which revealed that facial rigging is a vast and disparate field of expertise [Orvalho et al. 2012, Bastos et al. 2012b] (see Sections 2.6 to 2.8). Defining a facial rigging standard became impractical, because it would mean conducting dawning efforts with professional artists worldwide, when most of these artists are prevented by copyright from exposing their work and other artists are not willing to respond to research requests. Therefore, this research had to have a shift, and according to Tim McLaughlin, former Front End Supervisor at Lucasfilm Animation and Associate Professor of the Texas A&M University at College Station, it required "a focus in a particular facial style", Tim McLaughlin mentioned during the author's exploratory visit to Austin in December 2011 under the UT Austin | Portugal program. Aware of this advice, the author devised the definitive goal: define optimized character facial rigging and animation approaches that ease the complex jobs of the rigger and the animator, with a focus in the human face as a basis for other facial styles and also oriented to key frame facial animation. The reason to focus in the human face is that every facial behavior, be it of a creature or a cartoon character, can be realized as a variation of the behaviors originally found in the human face, as seen in the cartoon character created by [Harkins 2005]. The reason to focus in the technique of key frame facial animation is the fact that this technique is at its peak of growth, due to emerging recent techniques like motion capture. Still, key frame is the oldest and likely most used technique by digital artists, which makes key frame highly prone to the approaches proposed in this research. As a result, the following title was devised.

Easy Facial Rigging and Animation Approaches

The easy rig approaches are realized as scientifically validated optimized setups in facial rigging and animation carried out in a reliable and enjoyable way for riggers and animators. The purpose of the approaches is to improve the time-consuming and strenuous job of these artists in building and using a facial rig to produce subtle and appealing character facial expressions. The rig approaches are not meant to be mandatory and should not be realized as such, but rather act as scientifically validated suggestions, which are applicable to the entertainment industry of films and videogames. The advances of the suggested approaches over the state of the art occur in terms of providing objective guideline proposals to build the rig mechanics and provide flexible rig controls that together compose a facial rig. The mechanics are the underlying structures built by riggers to deform a facial model. The controls are the user interface structures built by riggers to allow the animators to trigger the rig mechanics to deform the character's face. The rig mechanical and control approaches suggested shape the method of this research, which is divided into a rig mechanical definition approach, presented in Chapter 4, and a rig control definition approach presented in Chapter 5. In terms of the facial rig mechanics, the question that is answered in the method is the following.

What if there are optimized approaches to build facial rig mechanics?

Based on the fact that the facial rig mechanical approaches found in the state of the art are disparate, the research for the optimized rig mechanics was set as the definition of generic and validated guideline proposals to place and configure the mechanical components in a facial rig to correctly deform a facial model. The proposals are considered generic due to their focus in the human face as a basis for other facial styles and validated because they are based in a set of rig construction and deformation conditions proposed and described in Section 4.1. The optimized rig mechanics are intended to help bring a consensus to the fragmented state of the art because they act as a guide to ease the job of the rigger in placing and configuring the rig mechanical components to deform a facial model. The set of rig construction and deformation conditions proposed were defined for the different facial regions of a human subject facial model. The deployment of the optimized mechanics involved studying the pre-established universal human facial behaviors in a per facial region basis using the book The Facial Action Coding System by [Ekman and Friesen 1978] and then define the rig mechanics responsible for triggering the required deformations (see Section 4.1). After the initial overall rig mechanical facial definition, a user experiment was carried out with five expert digital artists in character animation who are also college professors in the field. These users validated optimized rig mechanical approaches oriented specifically to help shed light on the rig setup of the facial behaviors known as the zipper and sticky lips effects (see Section 4.2). Zipper lips is a voluntary behavior realized as the action in effort of closing the lips together while keeping the jaw open, as in zipping the lips [Osipa 2007]. Sticky lips is the natural tendency that real-world lips have to adhere to each other toward their corners while opening the jaw [Osipa 2010]. Rig setups for the zipper and sticky lips behaviors are less common to be developed by digital artists because they are more complex to simulate. The reason for being more complex is that these behaviors occur in the lips, which is the facial region that holds the largest number of muscular activity, as categorized by [Ekman and Friesen 1978]. In terms of the facial rig controls, the question that is answered in the method is the following.

What if there is a personalized approach to use facial rig controls?

Based on the fact that the facial rig control approaches found in the state of the art are disparate, the research for the personalized controls was set as the definition of a multimodal rig interface system approach with a user-adapted interaction. The user-adapted interaction approach is meant to help provide a consensus to the fragmented rig interface controls found in the state of the art, because the common user interfaces (UIs) are unable to adapt to the preferences of animators. The reason for the lack of adaptability of the common UIs is that there is generally not enough time and budget in a production to allow the most experienced riggers to develop more advanced rigs. As a result, the animators are often presented with a feasible rather than an ideal facial rig control UI. Moreover, riggers often have to change or rebuild the control interfaces to adjust them to the requests of the animators, which is a harsh process for both riggers and animators, because these artists have to put extra effort and time to discuss modifications to the rig. It is also a difficult process for each artist individually because the rigger cannot predict or pre-define the manipulation of the rig controls and the animators often end up constrained to user interfaces with a limited control. Therefore, the option of having controls that allow for a personalized interaction is oriented to help reduce the repetitive process of rebuilding the control interfaces and make the facial animation process more efficient and enjoyable, because the interaction with the controls becomes more dependent on the preference of the animator. The personalized facial rig control approach was defined and validated in (i) an initial pilot study with three animators (see Section 3.3.2), followed by (ii) a user experiment with fifteen basic, skilled and expert 3D digital artists (see Section 5.1) and lastly in (iii) an online international user experiment with twenty long-term professional expert digital artists in character rigging and animation (see Section 5.2).

1.3 Motivation and Contribution

Character facial rigging and animation is a field which continues to be enclosed since it first emerged in the mid 90s in Toy Story (Pixar, 1995). The main reason is that, in the last 15 to 20 years, digital media of high impact has been generated primarily using commercial software such as 3dsMax or Maya. But, in the last 10 years, there has been a gradual increase in the use of Blender, an open-source 3D tool available at no financial cost and which can be further developed by any user. In the last decade, both freelance artists and companies worldwide have added Blender to their character production pipelines (e.g. Character Mill, Vivify Animation and Film Studio, Insizium, Schokolade Filmproduktion, Mox Studios, Fab Design)[1]. In March 2013, Pixar senior scientist Tony DeRose mentioned that "open-source software like Blender can do almost everything Pixar's software can do. We had a competitive advantage for ten years, but now we get more value by letting everyone contribute" [DeRose 2013]. The former facts, together with the fact that the author is an active Blender adept and entrepreneur for over five years, have led the author to implement this research in Blender as (i) a continuous investment in what the author believes, as DeRose does, is the future of the entertainment industry, (ii) the opportunity to cause a significant impact in the growing community of Blender digital artists, to help excel and sustain the community as well as leverage the use of Blender in the industry, (iii) a likely higher predisposal for validation from the Blender community, since Blender artists are not restricted to commercial constraints, (iv) the strong connection of the author with the Blender artists community, as author of Blender books and main chair of the first Blender PT Conference[2] [Bastos et al. 2014b] and (v) to help bring a consensus to the facial rigging and animation process, one of the most demanding research and production topics in the field of 3D character development, through a tool which

[1] http://charactermill.com/, http://www.vivify.co.in/, http://www.insizium.com/, http://www.schokolade.tv/, http://www.moxstudios.com/, http://www.fabdesign.info/

[2] http://problender.pt/en/conf2013

is accessible by any artist. Nevertheless, the implementation carried out in Blender is not meant as an alternative to commercial tools but rather as an added value to the industry. The purpose is not to replace but rather to enrich the professional character production pipelines with reliable and open character facial rigging and animation approaches. The open-source philosophy was perhaps more of an utopia in the past, but today it seems more feasible since it allows (i) a constantly growing and sustainable research and development (R&D) in this field and (ii) a worldwide sharing of knowledge that reaches more people more rapidly. Furthermore, although the implementation of the research was carried out in Blender, the approaches developed are likely to be applicable to other 3D tools, because they are based on (i) rigging techniques which are available in the major 3D tools and (ii) in the entire spectrum of the state of the art in this field [Orvalho et al. 2012, Bastos et al. 2012b]. The main contributions of this thesis for the fields of computer graphics and human-computer interaction are presented following.

- **optimized approaches to build the mechanics of a facial rig –** realized as generic and validated guideline proposals to build the fundamental facial rig mechanics based on a set of rig construction and deformation conditions (see Section 4.1) and on user validation of optimized rig approaches developed to animate the zipper and sticky lips behaviors (see Section 4.2). The purpose of the guidelines and of the lips approaches is to help the digital artist who is a character rigger build the deformation structures in a facial rig more clearly and with less frustration. The results of the user validation of the zipper and sticky lips rig approaches show an average score of 8,8 in a scale 0 to 10, with 43 in a total of 50 scoring slots (86%) holding scores equal or higher than 8 (see Section 4.2.5);
- **a personalized approach to use the controls of a facial rig** – realized as an initial user validation of a complete and customizable rig control interface approach based on

multiple control options provided to the animator (see Section 5.1) and on a final user validation of a user-adapted multimodal rig interface system approach based on multiple control options which are synchronized and scalable to provide a more flexible and enjoyable interaction when compared to the conventional rig interaction modes (see Section 5.2). The purpose of the multimodal interface approach is to help the rigger build a more reliable rig control interface and also help the animator to manipulate the controls in the rig UI with less frustration. The results obtained in the initial user validation of the multimodal rig approach show an average score of 8 in a scale 0 to 10 given by the 15 basic, skilled and expert users (see end of Section 5.1.5). The results of the final user validation of the improved user-adapted multimodal rig approach show an average score of 9,3 in a scale 0 to 10 given solely by the 10 long-term expert professional users of the test group of this comparative user experiment (see end of Section 5.2.6);

- **a basis for the creation of an automatic facial rig** – the rig mechanical and control approaches presented in this thesis are a basis to help deploy an add-on for Blender or a plug-in for other 3D tools to automatically generate a more friendly animation-ready facial rig.

Both the optimization process of the rig mechanics and the personalization of the rig controls are implemented in the 3D tool Blender for a human subject facial model as a basis for other facial styles. The target users of these approaches are the digital artists who are riggers and animators. The contribution for riggers is the deployment of a guide to build the mechanics and controls in a facial rig. The contribution for animators is the deployment of a user-adapted interaction to manipulate the animation of a character's face in an accessible and enjoyable way.

1.4 Structure of the Dissertation

The remaining chapters of this dissertation are outlined as follows.

Chapter 2. State of the Art

This chapter begins with brief explanations of (i) the origins and the application domain of rigging, (ii) the jobs of the rigging and animation artists within a character production pipeline, (iii) the importance for the rigger of studying the morphologies and behaviors of a character's face in order to plan and set the rig in accordance to the features of the character, (iv) the key frame animation technique as the scope of the rig methods presented and (v) the software artists use for facial rigging with a special highlight of the Blender tool, which is used to implement the method. After that, the chapter identifies and describes (i) a number of concepts planned by the rigger before and during the construction of a facial rig, (ii) the facial rig mechanical techniques and (iii) the facial rig control techniques. The chapter then presents an in-depth overview and illustration of several facial rigging approaches built by digital artists worldwide in the last 15 to 20 years for the purpose of key frame animation, including literary and web references. Lastly, it discusses the challenges of facial rigging, a field which is open to significant improvements. In this chapter, the results presented in Sections 2.1 to 2.3, 2.3.1 to 2.3.4 and 2.8 are published in A Facial Rigging Survey [Orvalho et al. 2012] (see Appendix A-2012-3) and the results presented in Sections 2.4, 2.4.1, 2.4.2, 2.5, 2.5.1 and 2.5.2 are published in Facial Rigging for Key Frame Animation [Bastos et al. 2012b] (see Appendix A-2012-2).

Chapter 3. Proof of Concept

This chapter begins by briefly stating the research problem and describing the solutions. It identifies the obstacles to overcome during (i) the optimization of the construction process of the facial rig mechanics and (ii) the personalization of the facial rig controls. It then presents two preliminary results as the framework that helped define the research pathway. A detailed behavioral study of the human face follows with a detailed construction of a human subject facial model to serve as the base to implement the easy facial rigging approaches described in Chapters 4 and 5, respectively in terms of the rig mechanics and rig controls. In this chapter, the framework results presented in Section 3.3.1 are published in A Demo of a Dynamic Facial UI for Digital Artists [Bastos et al. 2011] (see Appendix A-2011) and the results presented in Section 3.3.2 are published in A Demo of a Facial UI Design Approach for Digital Artists [Bastos et al. 2012a] (see Appendix A-2012-4).

Chapter 4. Easy Facial Rig Mechanics

This chapter presents the optimization process of the rig mechanics as generic and validated guideline proposals on how to build the rig mechanical components for the several facial regions of the human face as a basis for other facial styles. The optimized rig mechanics are deployed for the human subject facial model built and described in Chapter 3. This chapter goes into detail on how to place and configure the several mechanical components of a facial rig for a human subject facial model based on a set of construction and deformation conditions that help define how to accurately build the fundamental facial rig mechanics. Lastly, this chapter presents an optimization of the rig mechanics for the zipper and sticky lips effects, which are two complex facial behaviors essential for realistic facial animation. These were validated in a study with five expert digital artists, which is considered a relevant validation because of the users' expertise and the extensiveness of the user experiment details. In this chapter, the results presented in Section 4.1 are published in Generic and Certified Facial Rig Mechanical Approaches for Key Frame Character Animation [Bastos 2012] (see Appendix A-2012-1) and the results presented in Section 4.2 are published in Easy Character Facial Rig Mechanical Approaches to Animate Zipper and Sticky Lips Effects [Bastos 2013] (see Appendix A-2013-1).

Chapter 5. Easy Facial Rig Controls

This chapter presents a personalization approach for facial rig controls as a multimodal interface system based on a user-adapted interaction. The personalized rig controls are deployed for the human subject facial model built and described in Chapter 3. In this system the animators are provided with different sets of controls which are synchronized and scalable to ease the manipulation of the character's face. User experiments with basic, skilled and expert digital artists have been carried out to evaluate the ability of the multimodal control interface to adapt to the interactive needs of the users. The results show that a personalized interaction can ease the repetitive and frustrating process of rebuilding and using rig control interfaces. In this chapter, the results presented in Section 5.1 are published in FAC^2RIG: A Complete and Customizable Rig Interface System Approach for Facial Animation [Bastos and Fernandes 2013] (see Appendix A-2013-2) and the results presented in Section 5.2 are published in Multimodal Facial Rigging: Towards a User-Adapted Animation Interface [Bastos et al. 2014a] (see Appendix A-2014).

Chapter 6. Trends in Facial Rig Control Design

This chapter presents a final validation with twenty two undergrad students to confirm the relevance of the personalized rig control approach presented in Chapter 5 and to try to uncover future trends in facial rig control design. It is aimed at realizing if these users welcome the benefits of each interaction mode and if they can provide extra relevant data to improve the design of the facial rig controls in each interaction mode.

Chapter 7. Conclusions and Future Perspectives

This chapter discusses the work developed in this research and presents a set of future perspectives which have potential to be further explored.

Chapter 2

State of the Art

This chapter presents a survey on character facial rigging, the act of planning and building the deformation mechanics and the control interfaces to animate a character's face. The setup of the mechanics and controls is challenging because there is not a consensus on how to build the deformation structures and the control manipulators. This chapter begins with an analogy between character rigs and marionettes. Then the application domain of facial rigging is identified. After that the facial rigging process is explained within a character production pipeline, which involves studying the morphology and behaviors of a character's face and identify the rigging and animation software that the artists use. Then a group of important concepts that are considered during the facial rigging process are presented. Following a number of facial rigging techniques are presented, divided by the rig mechanics and rig controls. Motivated by the complexity and vastness of this field, the author then presents an extensive study of the facial rigging approaches built by several artists worldwide for the purpose of key frame animation, including literary and web approaches. This chapter ends with a discussion of the facial rigging and animation challenges and open issues that led to the definition of the facial rig approaches presented in Chapters 4 and 5. In this chapter, the results presented in Sections 2.1 to 2.3, 2.3.1 to 2.3.4 and 2.8 are published in A Facial Rigging Survey [Orvalho et al. 2012] (see Appendix A-2012-3) and the results presented in Sections 2.4, 2.4.1, 2.4.2, 2.5, 2.5.1 and 2.5.2 are published in Facial Rigging for Key Frame Animation [Bastos et al. 2012b] (see Appendix A-2012-2).

2.1 Character Rig Analogy

A fine analogy for a character rig is the setup of the strings that allow to manipulate a marionette doll [Orvalho et al. 2012] (see Appendix A-2012-3). To confer animation to a digital model, a rigger builds a rig for animators to manipulate, as to confer animation to a marionette a puppeteer builds the strings to manipulate the marionette doll [Leite 2007].

Early research in computer animated characters is based on the pioneer work of Lee Harrison III, who in the early 1960s “rigged up a body suit with potentiometers” [Sturman 1998] to animate real-time figures. A marionette doll is connected to a control frame and it behaves as an articulated doll when the frame is manipulated, causing the strings to push or pull the doll's joints. A higher number of strings increases the ability to handle the doll, but it also increases the complexity of manipulation. The same concept exists in a character rig, where rig control points can be assigned to different parts of a computer representation of a model. These control points affect the areas that they are linked according to the geometric operations applied to them (either a translation, rotation and/or scaling). The control points can be directly manipulated, which is more complex to use and maintain in case there are many controls in the rig to deal with, or by means of a more user-friendly graphical user interface. The inherent correspondence between rigging a digital model and manipulating a marionette was a source of inspiration for both academia and the industry since the early days of computer animation.

2.2 Application Domain

Facial motion is varied and subtle, be it voluntary or involuntary, which results in many possible facial expressions [Badler 1995]. As seen in [Orvalho et al. 2012] (see Appendix A-2012-3), the intrinsic variation and subtly of movements confers human beings the ability to communicate their emotions and intentions [Tian et al. 2001], which is also true for digital characters, in which facial motion is a key component for their performance and authenticity [Unay and Grossman 2005]. Facial performance is considered by [Stoiber et al. 2009] as essential to reproduce synthetic models that resemble a specific person. Nonetheless, due to the daily familiarity and contact with facial expressions of other people and also our own, humans can fairly identify unnatural behavior. This daily familiarity is the reason why authors generally agree that digital facial animation is so hard to achieve, because the expressions made by 3D characters need to be convincing for the audience. When this goal is achieved, facial animation becomes a prevailing instrument for story-telling in the entertainment industry, as illustrated by [Hauser 2011] in the book The Pixar Treasures.

Within the entertainment industry, facial rigging and animation can be applied in off-line and real-time interactive systems [Orvalho et al. 2012]. Off-line systems are feature films, visual effects or advertising. These systems usually require accurate and realistic characters that are able to emphasize the credibility of the audience. In off-line applications a character's animation is first rendered and then shown on cinema or on a TV screen. Real-time interactive systems are videogames and virtual reality. These applications require a balance between being visually realistic and having its graphics rapidly processed by the computer, because character animation needs to be calculated in real-time by means of the user interaction with the technological devices of the system in order to appear natural. Figure 2.1 [Orvalho et al. 2012] overviews the evolution of facial animation since the 1970s until today.

In off-line systems, computer facial animation was first introduced by [Parke 1972] in the film Futureworld (Richard T. Heffron, 1976). But the first CG (computer graphics) animated character conveying facial emotions is Tony de Peltri [Bergeron and Lachapelle 1985]. From there on, the higher expectations for realism of the audience drive the development of new technology to be used in films. In the mid 90s, computer animation was no longer measured in seconds or minutes, with Toy Story being the first full length feature film produced using

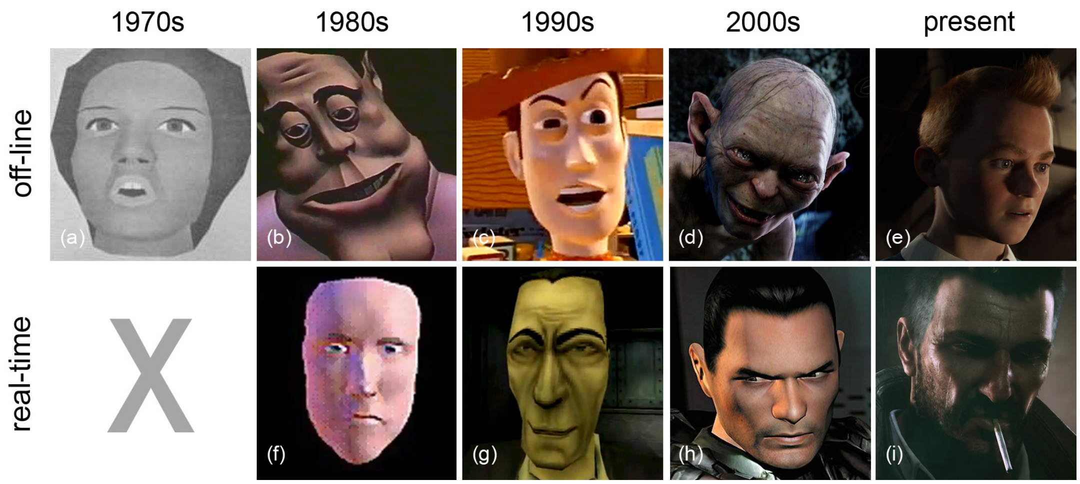


Figure 2.1: The evolution of character facial rigs and animations since the 70s until today; off-line: (a) first facial parametric model by Frederic Parke [Parke 1972], (b) Tony de Peltrie [Bergeron and Lachapelle 1985], the first animated short film to use a parameterized facial model, (c) Toy Story [Porter 1997], the first full length feature film produced entirely using the technology of computer animation [Henne et al. 1996] and the first computer animated film introducing the twelve principles of animation, (d) Gollum [Raitt 2004], a highly realistic computer animated character in a leading role in a live action film, (e) The Adventures of Tintin (2011), recent state of the art in character rigging with performance capture in an animation film; real-time: (f) Mike the Talking Head [Degraf 1988], the first real-time virtual puppetry, (g) Half-Life (1998), early facial animation on 3D videogames, (h) Doom 3 (2004), bone-based facial rigs for videogames, (i) The Samaritan Demo [Oliver et al. 2011], recent state of the art in facial rigging and animation in real-time (image retrieved from [Orvalho et al. 2012]).

only computer animation techniques [Henne et al. 1996] and the first computer animated film introducing the twelve principles of animation. In the early 00s, the trilogy Lord of the Rings (Peter Jackson, 2001) featured Gollum, a CG character with realistic expressions in a leading role in a live action film. A recent example of a fully CG animated film is The Adventures of

Tintin (Steven Spielberg, 2011). Recent films that combine real footage with CG facial animation are Rise of the Planet of the Apes (Ruppert Wyatt, 2011) and The Avengers (Joss Whedon, 2012). Current examples of films featuring state of the art facial rigging and animation are The Lego Movie (Phil Lord and Christopher Miller, 2014) [Hollywood 2014] and Transformers 4 (Michael Bay, 2014) [Failes 2014], in which the visual effects studio Industrial Light & Magic (ILM) used its proprietary procedural rigging system BlockParty [Rose et al. 2013].

In real-time interactive systems, early computer facial animation was generated using full motion video only, like in the videogame Dragon's Lair (Advanced Microcomputer Systems, 1983). A digital puppet demo shown at SIGGRAPH'88 Electronic Theater, called Mike the Talking Head [Degraf 1988], was the first system to adopt facial rigging as a necessity to include interaction in the applications. Then in the videogames The Secret of Monkey Island (LucasArts, 1990) and Wolfeinstein 3D (Id Software, 1992), respectively 2D sprites and 2.5D techniques were used. In the late 90s, videogame animation improved with the rise of the 3D age, providing full 3D character body animation for action games like Quake (Id Software, 1996). Despite technology at the time still did not support facial animation for the purpose of story-telling, facial rigging allowed basic facial expressions and speech to be included in Half-Life (Valve, 1998) using the Quake Engine (Id Tech 1), thus increasing the level of interactivity. Today, new architecture in terms of Graphics Processing Units (GPUs) includes algorithms for facial tessellation and deformation [Beeson and Bjorke 2004, Bunnell 2005] that allow to create skin wrinkles and morphing to produce real-time photorealistic facial synthesis (UDK Samaritan Demo [Oliver et al. 2011]). A current example of a videogame featuring state of the art facial rigging and animation is The Last of Us (Naughty Dog, 2013) [Simantov, 2013].

2.3 The Facial Rigging Process

Character facial animation is possible due to the process of character facial rigging. A rig is a “control structure designed to ease the animation of creatures and objects” [Schleifer et al. 2002]. [Falk et al. 2004] define rigging as “the process of taking a static, inanimate computer model and transforming it into a character that an animator can manipulate frame-by-frame to create motion”. Rigging can also be realized as “setting up a group of controls to operate a 3D model” [Orvalho et al. 2012] (see Appendix A-2012-3). In overall, rigging is a process that involves planning and setting the deformation mechanics and the control interfaces to animate a character. Rigging is necessary because it is impractical to manipulate each vertex of a 3D model individually, especially when dealing with complex facial models, which can have thousands of vertices due to the required detail of deformation. In certain characters there can even be other 3D objects interacting with the face that make it more complex to deform, an example can be a character using glasses. Depending on the complexity of the facial model and the level of control desired, a character face can have dozens of rig controls. For instance, Woody, the main character in Toy Story (Pixar, 1995), had “712 avars (controls), counting 212 in his face alone” [Henne et al. 1996], resulting roughly in 30% of the character rig being in Woody's face. Other characters in other films can have higher percentages due to their own particular facial features, which sometimes require dealing with special animation effects such as, for instance, dismantling and re-assembling parts of the face.

Figure 2.2 shows the position of the rigging stage within a professional character production pipeline. The complexity inherent to a character production pipeline requires the digital artists involved in the pipeline to carry out many hours of labor in specific tasks that require specific skills, such as the modeling, rigging and animation stages. As seen in Figure 2.2, after the pre-production stage, in which character concepts are drawn and illustrated, the modeling stage takes place, in which static 3D models are built to represent the characters. In

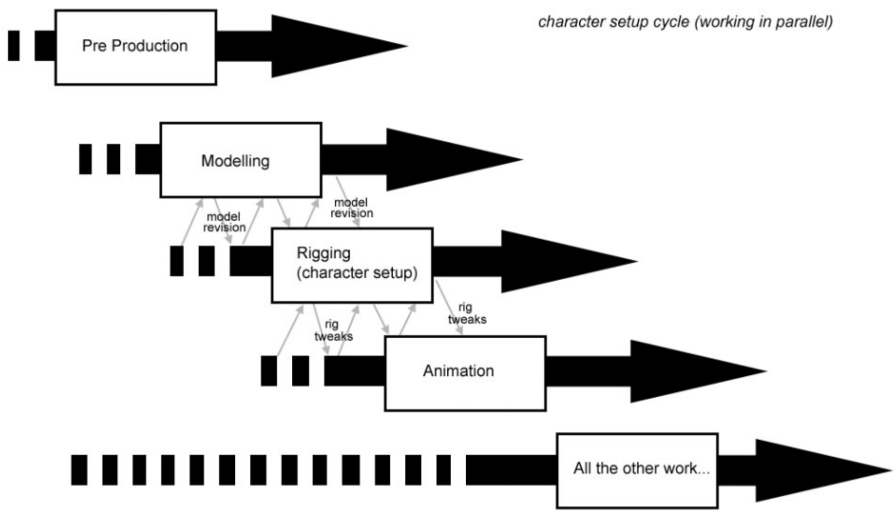


Figure 2.2: The character setup cycle used within professional 3D production pipelines (image retrieved from [Schleifer et al. 2002]).

order to animate the models, they need to be rigged in the character rigging stage, which means building a deformation structure and a control interface structure for the model. The deformation and control structure, also called the rig, can then be used by the animator to deform the character in the animation stage. But, the production pipeline is usually not linear, there is a continuous back and forth interaction between the different production departments. The rigger is positioned in-between the modeling and animation stages because the role of the rigger is to bridge these stages. For that purpose, the rigger can ask the modeler for revisions to the current model version and can also tweak the rig according to what is necessary for the animators to feel more comfortable using the rig.

The professional teams working in these pipelines are larger depending on the number of digital characters required for the production and also on the production time and budget. A large film or videogame production can have dozens to hundreds of professionals. Some of these artists have specific jobs within the production pipeline, specializing in production stages like modeling, rigging or animation. Each of these stages is further detailed in the next paragraphs.

Character Modeling

In character modeling the artist is concerned with constructing the geometry of the character to allow accurate deformations. Such task can take time to carry out, especially if the face needs to get wrinkled during animation. If that is the case, the modeler needs to add more polygonal detail using fast modeling methods such as sculpting [Raitt and Minter 2000]. If the vertices, edges and polygons of a facial model are not accurately distributed it will be harder to manipulate the model, even if the rigging of the model is carefully designed, and it will also delay the rigging process.

Character Rigging

Character rigging evolved from modeling. The artists known today as riggers were known in Toy Story as the model builders [Henne et al. 1996], who at the time modeled and also rigged characters. But since then the increasing need for characters to perform more complex facial animation for breakthrough feature films has also increased the rigging complexity. Today, rigging is realized as the specialized process of planning and setting the rig mechanics and rig controls for a character. The rig mechanics are the components of the rig that deform the facial model and the controls are the user interface that trigger the mechanics. Even with a set of efficient rig controls, a facial model can still be hard to manipulate if the rig mechanics do not deform well the model. Given an accurate facial model, the facial rig that will be attached to the model needs to be efficient both in terms of the deformation mechanics and the control interfaces. The attachment of a rig to a model is a laborious, time-consuming and sometimes frustrating process because it means preparing a character for animation taking many variables into account (e.g. behavioral study, deformation accuracy and interface control design and interaction). Rigging is a demanding and startling job even for skilled artists, with only a few becoming expert riggers during their careers, because rigging demands highly conceptual, technical as well as artistic skills. In conceptual terms, riggers need to study character anatomy and behavior in order to create correct and believable character model deformations. In technical terms, riggers need to build efficient rig deformations and user control interfaces, which animators can use to have a precise manipulation of a character. In artistic terms, riggers need to predict what the animators prefer in terms of control interaction paradigms in order to prevent carrying out multiple iterations with the animators and save production time. There are few character riggers in a production and even less with facial rigging experience.

Essentially, a rigger provides (i) an accurate deformation of a character, (ii) the necessary controls for proper animation and (iii) the ability for the rig to allow quick updates in case an undesired behavior is found or a new one is required. Unfortunately, rigging is a process little documented and there are a limited number of artistic and technical references available that shed light on the facial rigging process. The shortening of the artistic references is because facial behavioral study is vast and complex. Only a few authors have carried out significant work in this field, with the Facial Action Coding System [Ekman and Friesen 1978] being one of the most relevant references. This book teaches the reader how to "recognize and score the Action Units (AUs) which represent the muscular activity that produces momentary changes in facial appearance" [Ekman and Friesen 1978]. The shortening of technical references is due to the fact that riggers are a relative minority within the large teams of artists working in the entertainment industry, therefore they produce less tutorials. Also, some riggers are not able to share their rigs due to professional copyright reasons and only a few are able or willing to share their work. Another reason is the inherent complexity of facial rigging, which is a process hard to describe step-by-step. The book Stop Staring: Facial Animation Done Right by Jason Osipa [Osipa 2010] is one of the most relevant literary technical references and the only including advanced approaches. In the last 15 to 20 years riggers have been focusing on defining their own specific control solutions, which sometimes may not meet the expectations of the animators.

Character Animation

Character animation is carried out by animators, who use the rig controls to manipulate the 3D characters. The animation process "involves projecting personality, expressing nuance and emotion, using timing, staging, anticipation and follow-through" [Henne et al. 1996]. But usually animation is not a straightforward process and animators often require specific controls, so they iterate with the riggers to define a control manipulation that reduces their effort during animation.

Vladimir Tytla, considered "the first animator to bring true emotions to the cartoon screen" [Thomas and Johnston 1981], stated that "there isn't a caricaturist in this country (US) who has as much liberty as an animator here of twisting and weaving his lines in and out" [Thomas and Johnston 1981]. Tytla was referring to the enthusiastic work that was carried out by him and his colleagues in the early days of Disney in the 1920s and 1930s. Decades later, discussing the production of Lord of the Rings and the demanding job of the animators, Matt Aitken, Digital Visual Effects Supervisor at Weta Digital, and his team reveal that "it is the work of the animators in bringing a complex digital character like Gollum to life that we consider to be Weta Digital's true strength" [Aitken et al. 2004]. These statements represent the importance of having efficient control solutions in a character production pipeline to ease the job of the large number of animators that are usually involved in a film or videogame production (e.g. "Toy Story was completed by thirty animators" [Henne et al. 1996] and today Pixar has dozens of animators).

Animators need highly interactive digital control environments in order to manipulate their characters in an efficient and enjoyable way. Usually, an animator will like to find a number of features in a rig that together make the rig easier to use. The rig being easy to use means that (i) the animator should be able to associate each facial rig control to its corresponding facial region, (ii) the control should have a high degree of precision of manipulation, (iii) the control manipulation needs to result in accurate and believable geometric deformation, therefore the mechanical components of the rig need to be triggered accordingly to the manipulation being made, (iv) certain deformations need to be automated to fasten the animation (e.g. the skin in the upper and lower eyelids can automatically follow the orientation of the eyeballs when the animator manipulates the eyeballs), (v) depending on the character there are certain mechanical and control movements that can be constrained to avoid awkward deformations, to prevent being triggered accidentally by the animator (e.g. via an excessive manipulation of a control), (vi) the controls need to work the way the animators expect them to and (vii) the rigger needs to search for a balance between the number of controls provided and the cluttering of the visual space.

2.3.1 Morphologies

Before a rigger studies the facial movements of a character, it is necessary to analyze its shape and structure, or its morphology. The study of the facial morphology provides the facial features needed for the design of its rig structure in order for the rig to deform the geometry correctly. There are two types of morphology requirements: *taxonomic* and *geometric*, which were first defined in [Orvalho et al. 2012] (see Appendix A-2012-3). The taxonomic morphology describes the shape of the anatomy and the visual styling of the face and the geometric morphology defines the type and geometric techniques that can be applied for shaping the 3D model. These types are described in more detail following.

Taxonomic Morphologies

It is hard for 3D digital artists to reproduce on the screen an immersive human-like face because of the fact that faces can adopt a great diversity of expressions, due to their anatomy and the different visual styles existent. Many authors [Miller 2004, Gorden 2005b, Miller and Grossman 2006, McLaughlin and Sumida 2007, O'Neill 2008, Diamant and Simantov 2010] alert that previously to rigging a character, it is important to understand both human anatomy as well as comparative anatomy. The understanding of anatomy becomes more evident in a character's face because even if all the facial muscles are well simulated individually, which per itself requires careful study, it is also necessary to produce the desired animation effect when different muscle actions are combined to form different expressions. The problem of the anatomic diversity has led authors to simplify the facial structure of a realistic character by separating it into anatomical [Gray 1974, O'Malley 1983, Pessa 2012], anthropometric [Farkas and Leslie 1994], and artistic form [Parks 2003]. The morphology becomes harder to realize mostly when cartoon-like characters are introduced to the 3D representation [Thomas and Johnston 1981, Lasseter 1987, Blair 1996, Williams 2001, Gorden 2005a].

As a result of the variety of expressions that a face can adopt during animation, artists try to search for methods to visually classify facial styles, but there is no formal classification. However, a pictorial vocabulary of the visual arts was proposed by [McCloud 1993] to describe a wide number of facial styles. This vocabulary is represented as a triangle in which each vertex is a style: (i) reality, (ii) language, or reality simplification and (iii) the picture plane, or how much the facial shape is abstracted. As seen in Figure 2.3, the triangle contains different examples of facial styling which change depending on their position along the edges of the triangle, as they get closer or further away from each vertex.

In the domain of 3D character representation, different taxonomies were presented by [McCloud 1993], [Ritchie et al. 2005] and [McLaughlin 2006], which are illustrated in Figure 2.4.

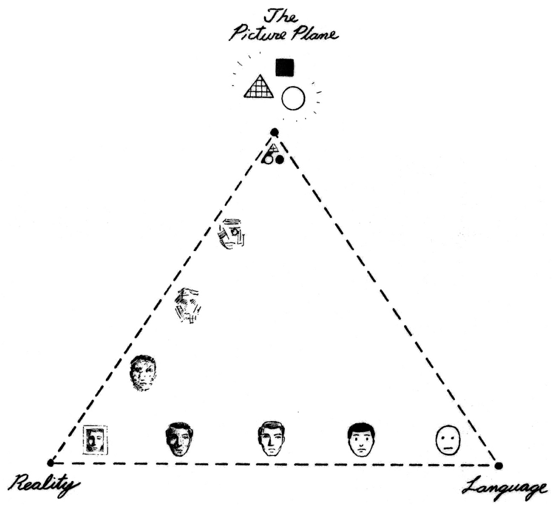


Figure 2.3: The Pictorial Vocabulary created by [McCloud 1993] to describe a wide number of facial styles divided into (i) reality, (ii) language and (iii) the picture plane (image retrieved from [Orvalho et al. 2012]).

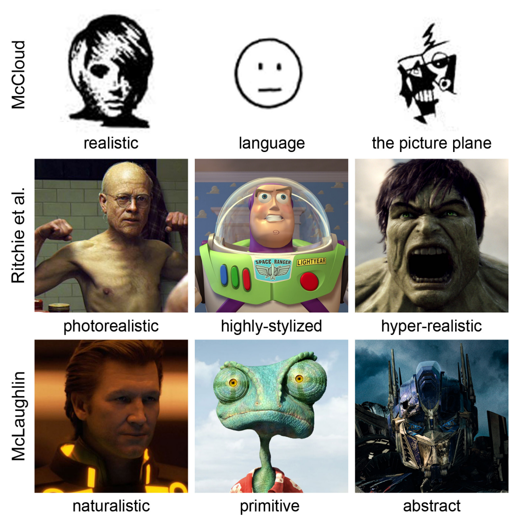


Figure 2.4: Classifications proposed for different character facial styles in groups of taxonomic morphologies; top row: set of early concepts by [McCloud 1993], left to right: representations for reality, language and the picture plane; middle row: set of more recent concepts by [Ritchie et al. 2005], left to right: Benjamin Button, Toy Story and Hulk; bottom row: most recent concepts by [McLaughlin 2006], left to right: Tron Legacy, Rango and Transformers (image retrieved from [Orvalho et al. 2012]).

As seen in Figure 2.4, three different facial styles were identified by [Ritchie et al. 2005]: (i) photorealistic, considered as the no-style type (ii) hyper-realistic, for realistic creature-like characters and (iii) highly-stylized, which represents a character that can range in scale and shape independently of the limitations of the physical world. In 2006 [McLaughlin 2006] also proposed a division of a character's visual style as: (i) naturalistic, for characters visually and behaviorally similar to real people or creatures, (ii) primitive, which is associated to cartoon-like characters and (iii) abstract, when new character forms are created based on physical elements that make sense when arranged together.

Geometric Morphologies

The geometric representation of 3D facial models has been researched for the past four decades [Parke 1972, Park et al. 2005, Sheng et al. 2006, Dickens et al. 2007, Elyan and Ugail 2007, Chang and Habib 2008, Widanagamaachchi and Dharmaratne 2008, Beeler et al. 2010, Kaiser et al. 2010]. Therefore, character modeling techniques used in production are well-known in the entertainment industry [Unay 2004, Unay and Grossman 2005, Osipa 2007, Oliverio 2007, Fedor et al. 2007, Patnode 2008]. Studying the geometric representation techniques is important to assure that a character facial model can be properly rigged so that it deforms correctly during animation. Geometric representation involves using polygons, NURBS (Non-Uniform Rational B-splines) or subdivision surfaces. Each of these techniques has its own benefits and limitations and the application context of the rig needs to be considered in order to use each technique. Figure 2.5 illustrates three facial models, each modeled using each of the geometric representation techniques identified.

High-end polygonal modeling techniques are available in major 3D modeling packages. Polygons became the modus operandi technique used by artists to model a character's face [Fedor et al. 2007, Oliverio 2007, Widanagamaachchi and Dharmaratne 2007, Patnode 2008].

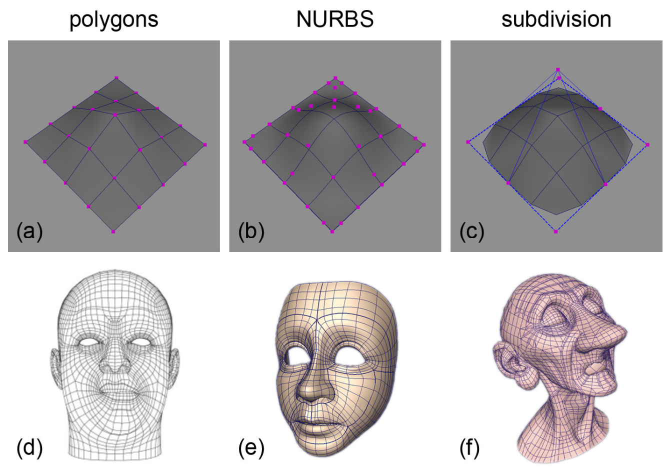


Figure 2.5: The three types of geometries used for facial modeling; top row: 3D representations of (a) polygonal, (b) NURBS and (c) subdivision; bottom row: image representations of (d) a polygonal human face by [Osipa 2010], (e) the Balloon girl face from the short film Bingo the Clown modeled using NURBS patches (Chris Landreth 1998) and (f) the main character in Pixar's Geri's Game (Jan Pinkava, 1997) modeled with subdivision surface (image retrieved from [Orvalho et al. 2012]).

The earliest attempt to determine the minimum number of polygons needed to model a realistic face established the use of 250 polygons [Parke 1972, Parke 1974]. Since then, this number has been gradually increasing due to the also gradual higher demand for better visual looking facial models and also the evolution of computer graphics hardware. The NURBS technique emerged in the early 70s from the work on curves and surfaces of Pierre Bézier. In the late 90s, [DeRose et al. 1998] highlighted that "the most common way to model complex smooth surfaces such as those encountered in human character animation is by using a patchwork of trimmed NURBS", which is also included in the major 3D tools. Despite the subdivision surface technique was introduced in the late 70s by [Catmull and Clark 1978], it was only years later in 1997 for the short film Geri's Game by Pixar, that this technique was used for the purpose of character facial modeling [DeRose et al. 1998].

The topological distribution of the geometry is also an important requirement and which is independent of the modeling technique used. Usually the topology of the mesh is optimized for deformation by preparing the edge loops of the model based on an anatomical approach. The first reference to this concept was made by [Waters 1987], who described that the "facial parameterization techniques have dealt principally with the surface characteristics of the skin and have not been concerned with the motivators of the dynamics". Recently, the artists [Gibbs and Derakhshani 2005, Unay and Grossman 2005, Cabrera 2008, Pardew 2008, Parke and Waters 2008] agree that the edge loops of the facial model should be built according to the flow and the effect of the muscle lines. The reason is to maintain their location in the face and also provide a more realistic animation, because the movement of the skin deformation will follow along with the deformation path of the muscles. Figure 2.6 shows the first edge loop representation in comparison with the current state of the art representation.

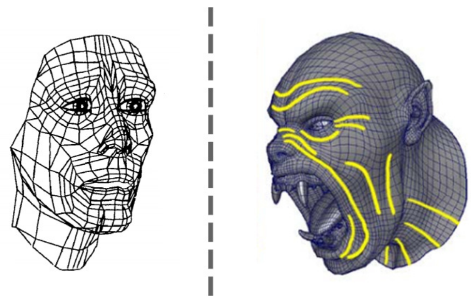


Figure 2.6: Evolution of the distribution of facial geometric topology; left: first edge loop representation by [Waters 1987]; right: current state of the art on edge loop representation by [Unay and Grossman 2005] (image retrieved from [Orvalho et al. 2012]).

2.3.2 Behaviors

The face is able to produce a wide range of behaviors per each of its facial regions. The behavioral conditions of a face make it unique, particularly in comparison to the rest of the body. The reason for the uniqueness of a face is that facial regions do not follow a hierarchy similar to that used in the body regions. In the body, the animation controls have direct parent-child relations and vice-versa. In the face, each control is independent and therefore a specific or predefined way of positioning the animation controls does not exist. In character rigging, the philosophy and the techniques used for the face are different from those used for the body, because facial animation is widely considered as the most relevant stage within character animation [Ritchie et al. 2005]. For this reason, character faces usually require more complex rigging structures in order to support the diversity and inconsistency of their facial behaviors. But these rig structures require a lot of work and time to build, which sometimes leads digital artists to not consider fundamental behaviors. Studying the facial behaviors that the character is able to perform helps to realize the character's limitations and allows the rigger to be able to prevent the rig components from adopting invalid positions which can lead to incorrect deformations of the facial skin.

As seen in [Orvalho et al. 2012] (see Appendix A-2012-3), Charles Darwin's book The Expression of the Emotions in Men and Animals [Darwin 1872] is the starting point for studying facial behavior. Later, in the mid 1970s, [Ekman and Friesen 1975] identified the six basic facial expressions, also known as the universal facial expressions: surprise, fear, disgust, anger, happiness and sadness. Not long after, [Ekman and Friesen 1978] created the book FACS, or Facial Action Coding System. This book is a manual which categorizes and scores the muscular activity of the human face into Action Units (AUs), Action Descriptors (ADs) and Gross Behavior Codes (GBCs). FACS also includes the visible facial movements responsible for producing momentary changes in facial appearance, which can be either emotional or conversational signals. Regarding the study of human facial behaviors, FACS is considered the largest and most accurate reference available. Figure 2.7 shows a number of examples of Action Units identified in FACS and the combined AUs that produce the basic or universal facial expressions. The information included in FACS allows artists to determine the different categories for each facial behavior in a reliable way and without the need to identify the facial muscles responsible for each movement. Many authors support their work with FACS [Perlin 1997, Miller 2004, Ritchie et al. 2005, Evans and Hagedorn 2007, Alexander et al. 2009, Villagrasa and Susin 2009, Athias 2010b, Li et al. 2010, Komorowski et al. 2010, Osipa 2010, Arghinenthi 2011, McLaughlin et al. 2011, Weise et al. 2011].

AU	FACS Name	AU	FACS Name	AU	FACS Name
1	Inner Brow Raiser	12	Lid Corner Puller	2	Outer Brow Raiser
14	Dimpler	4	Brow Lower	15	Lip Corner Depressor
5	Upper Lid Raiser	16	Lower Lip Depressor	6	Check Raiser
17	Chin Raiser	9	Nose Wrinkler	20	Lip Stretcher
23	Lip Tightener	10	Upper Lid Raiser	26	Jaw Drop

Basic Expressions	Involved Action Units
Surprise	AU1, 2, 5, 15, 16, 20, 26
Fear	AU1, 2, 4, 5, 15, 20, 26
Anger	AU2, 4, 7, 9, 10, 20, 26
Happiness	AU1, 6, 12, 14
Sadness	AU1, 4, 15, 23

Figure 2.7: Examples of human facial behaviors and expressions studied in FACS; top table: examples of Action Units; bottom table: the basic facial expressions represented by their corresponding Action Units [Deng and Noh 2007].

One aspect which is only briefly mentioned in FACS [Ekman and Friesen 1978] is the behavior of the tongue. The authors of FACS "hope someone will do for tongue displays, what we have done for facial movements, for we believe that some of the different tongue displays that we are not distinguishing may well have interesting and diverse psychological significance" [Ekman and Friesen 1978]. Although not analyzed in FACS, it can be necessary for the rigger to provide control over a character's tongue so that the animator can manipulate it. A helpful reference for the rigger to realize how the tongue behaves is the group of generic tongue positions identified by [Fleming and Dobbs 1998] and illustrated in Figure 2.8.

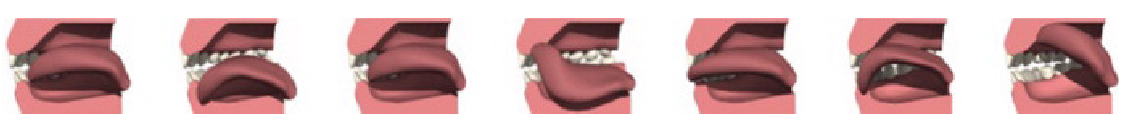


Figure 2.8: A group of generic tongue positions presented by [Fleming and Dobbs 1998].

The Artist's Complete Guide to Facial Expression by [Faigin 1990] is a book which includes a comprehensive visual index of the facial actions. This book is a fine illustrated reference for the rigger to study the role of the facial muscles, the nature of their motion and the effects of their movement on the facial skin. Figure 2.9 shows an example in which [Faigin 1990] presents the different mouth apertures and their dimension differences.

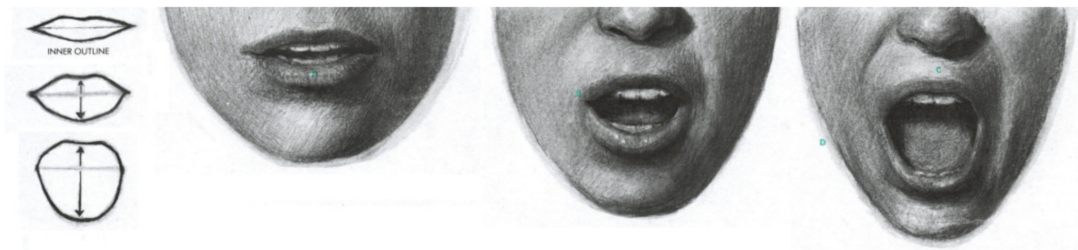
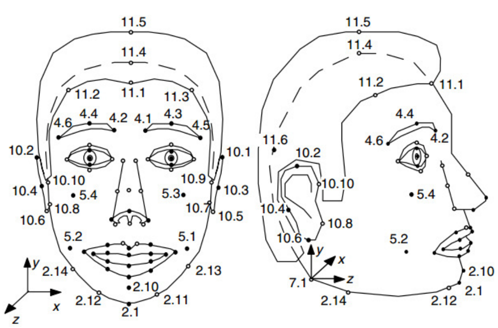


Figure 2.9: Gradual mouth opening; left: a neutral stance with the upper and lower lips slightly separated from each other; middle: the mouth slightly opened with the lower lip gradually gaining a greater distance from the center lips line; right: the mouth wide opened with the lower lip having a much greater distance from the center lips line [Faigin 1990].

The MPEG-4 facial animation specification, described by [Pandzic and Forchheimer 2002], was the first facial control parameterization to be standardized. Figure 2.10 illustrates the MPEG-4 face groups in the MPEG-4 specification.



Group	Number of FAPs
1. Visemes and expressions	2
2. Jaw, chin, inner lowerlip, cornerlips, midlip	16
3. Eyeballs, pupils, eyelids	12
4. Eyebrow	8
5. Cheeks	4
6. Tongue	5
7. Head rotation	3
8. Outer-lip positions	10
9. Nose	4
10. Ears	4

Figure 2.10: MPEG-4 Facial Animation; left: the facial feature points identified in a per facial region basis for a human face; right: different face groups each represented by a specific number of facial animation parameters (FAPs) by [Pandzic and Forchheimer 2002].

The MPEG-4 facial animation standard has been used in the entertainment industry, medicine and telecommunication. It helps define animation independently from the facial model, provided that the model is properly annotated according to the standard. It is based in three types of facial data: facial animation parameters (FAPs), facial definition parameters (FDPs) and the FAPs interpolation table (FIT). FAPs are the minimum set of parameters that MPEG-4 established for facial animation, including 84 feature points (FPs) distributed on a number of base key frames of a standard facial model posed in a neutral expression to easily reproduce facial movements, expressions, emotions and speech. The MPEG-4 standard is mainly directed towards creating realistic facial behavioral models.

2.3.3 Animation Technologies

A number of animation technologies are identified in [Orvalho et al. 2012] (see Appendix A-2012-3), such as key frame interpolation [Burtnyk and Wein 1970], geometric deformation [Sederberg and Parry 1986], physically-based methods [Platt and Badler 1981], motion capture [Waters 1987, Lee et al. 1995] and retargeting [Gleicher 1998]. The facial rigging approaches presented in this dissertation are focused on key frame interpolation due to a number of reasons: (i) this is the technique that artists most use as it continues to be popular in CG animation [Orvalho et al. 2012], (ii) it is the most accessible and affordable technique when compared to other animation techniques like motion capture [Williams 1990] because it does not require expensive equipment, (iii) the results from other animation techniques need to be adjusted using key frame, (iv) it allows artists to have a more precise manipulation of the animation [Orvalho et al. 2012] because animators assign values to the parameters of the rig controls directly in the timeline of the 3D tool with a high degree of precision and also (v) because key frame is at its peak of growth [Orvalho et al. 2012, Bastos et al. 2012b], therefore more prone to welcoming the optimization approaches presented in Chapters 4 and 5.

The term keyframing was introduced in computer graphics in the early 70s by [Burtnyk and Wein 1970]. In the mid 70s the technique was applied to skeleton structures [Burtnyk and Wein 1976] and in the mid 80s the use of splines is introduced to key frame animation by [Kochanek and Bartels 1984] to interpolate a set of spatial control points and allow temporal control by editing high-level parameters at each control point. The principles of traditional animation into 3D key frame animation are introduced by [Lasseter 1987]. Later, [Reeves et al. 1990] presented The Menv, an animation environment that uses editable spline curves of single animation parameters per time (e.g. still widely used x translation, z rotation), their animation graphical editor is included in current commercialized software. In the mid 90s, motion paths are introduced by [Snibbe 1995] as a solution for editing key frame animation in 3D, by splitting the parameters into 3D position and time, allowing the direct control of the positions. Today, rigging for key frame animation continues to be widely used by digital artists worldwide, who demonstrate their skills in character facial animation by providing their showreels in websites like YouTube and Vimeo (see Sections 2.7.1 and 2.7.2).

2.3.4 Rigging and Animation Software

The construction of a rig requires the rigging artist to use particular software, which can be different depending on (i) the application of the rig (e.g. a film or a videogame), (ii) the inherent rig creation process and (iii) the control manipulation done by the animator after the rig is ready to be used. In a commercial production of a film or a videogame, a number of software applications commonly used are the Autodesk 3D solutions (3D Studio Max [Autodesk 2014a], Maya [Autodesk 2014b] and Softimage [Autodesk 2014c]), Blender [Blender Foundation 2014a], Cinema 4D [Maxon 2014] and NewTeck LightWave [NewTek 2014]. These applications provide the necessary tools to create a character rig from scratch, including both the creation of the mechanical components that allow deforming the model as well as the manipulation components that represent the control interface for the animator.

In commercial productions companies also develop their own in-house rigging tools to make the rig creation process faster and more simple for the rigger. These tools are especially useful to ease repetitive rigging because they automatically create predefined deformation skeletons and control interfaces which may need only slight adjustments to fit the character. They can also simplify for instance the naming conventions that need to be given by the artist to the several rig components, or to export the rig, etc. These tools are proprietary of the studio which developed them, being usually built using the programming language that ships with the 3D software (e.g. Python scripting in Blender or in Maya). There are also other specific rigging tools which integrate with the main 3D production software to reduce the amount of rigging work. These were first identified in [Orvalho et al. 2012] (see Appendix A-2012-3), and are briefly described following.

Bony Face [Scriptattack and FX 2014] is a plug-in for Autodesk 3D Studio Max that automatically generates a bone based system with many rig control points which are distributed in a character's face. The main purpose of this system is to simulate how real facial muscles behave by providing detailed control.

Facial Animation Toolset [Animation Institute 2014] is a plug-in for Autodesk Maya available since 2006. It was developed to rapidly create efficient and believable non-linear character facial deformations. This toolset plug-in provides a high quality facial model and a predefined interface with controls to manipulate the movements of the character's facial regions.

Face Machine [Anzovin Studio 2014] is an automatic facial rigging tool for Autodesk Maya, based on intelligent point weighting that generates direct manipulation of the controls of the character's facial skin.

Face Robot is a tool available inside the Autodesk Softimage package. It aims at reducing the preparation time of a facial rig by providing a step-by-step wizard to the facial rigging process. It uses a soft tissue solver technology that automatically creates direct manipulation controls that simulate organic facial skin movements upon manipulation. Face Robot can be used with motion capture data or via key frame animation.

Although the solutions described previously may be faster to use if the artist has experience using them, their efficiency is highly subjective, because these tools provide predefined user interfaces which are limited to their own features. Therefore, these interfaces tend to be less flexible than the major commercial animation packages, which allow creating full featured rigs from scratch. The downside of the major commercial animation packages is that they require experience and creativity from the rigger.

2.4 Facial Rigging for Key Frame Animation: Concepts

A number of concepts are planned by the rigger before and during the construction of a facial rig for key frame animation. The rigger needs to consider these concepts to obtain an initial realization of how to integrate the rig with the facial model. But despite the rigger's anticipation, it is often necessary to overview the concepts a number of times due to momentary changes in the character production pipeline. These concepts involve the planning of the rig hierarchies, transformations, scale and pose as well as of the naming conventions applied to the several components of the facial rig. These concepts are overviewed ahead in Sections 2.4.1 and 2.4.2.

2.4.1 Hierarchies, Transformations, Scale and Pose

The rig hierarchies define the relations between the components of the facial rig structure based on their linkage and transformation sets. Each component of the rig is evaluated via its pivot, or its center of gravity (COG) [Ritchie et al. 2005]. In a character facial rig, predefined hierarchies are used mostly to control the neck and head, because these derive from the spine.

The rig components transformations (translation, rotation and/or scale) occur from each component's pivot. An example is the transformation of a character's head being directly dependent or not of the transformation of the neck. For example, the rigger can set a control rig that allows the animators to rotate the neck and either have the head inherent the neck rotation or keep the orientation of the head unchanged, like if the head moves forward or backwards keeping an horizontal line of sight. For the transformations of the rig components to be accurate, riggers set their orientations and also set coplanarities for them [O'Neill 2008]. The orientations are their individual alignments in space (evaluated in X, Y and Z) and the coplanarities are the alignments calculated between each of the rig components.

The scale of a rig is its overall dimension and also the dimensions of each rig component when compared to its neighboring rig components. If the scale of a rig component is incorrect, others dependent of it can become incorrect via their hierarchical relations, causing awkward poses of the facial rig and ultimately of the facial model. A character rig also has the benefit of allowing the character model to be scaled in order to fit the environment.

The poses are the stances that the rig is able to adopt. In a professional rig, the initial pose usually corresponds to the neutral stance of the character. The reason that the neutral pose is the first is that it represents a soft starting point for the remaining rig transformations, because the facial muscles are at rest in the neutral pose. Digital artists [Maestri 1996, Osipa 2003, Simantov 2013] defend that the eyelids and mouth of a neutral facial model should be slightly opened to ease the rig construction process in those areas. This neutral pose is especially useful to define the influence that the rig components of the eyelids and mouth have in their corresponding polygonal regions of the facial model, because these areas have protrusions which are hard to locate during the skinning process. Skinning is the process of assigning the object that needs to be deformed with the rig structure that will allow to control the object's deformation (see skinning in Section 2.5.1). The rig poses are also planned according to the deformation required per each facial region [Harkins 2005], which depends on whether the region should behave as being rigid, non-rigid or a mix.

2.4.2 Naming Conventions

It is common in facial rigging to have a large number of rig components to deal with the entire behavioral set of a character. As a result, it gets harder to keep track of the rig. It then becomes fundamental to use naming conventions to (i) properly identify and relate all the rig components – for "peace of mind and personal debugging" [Gibbs and Derakhshani 2005], (ii) easy re-editing of the facial rig even "after a period of not using it" [Gorden 2005b], (iii) keep the 3D tool optimized, (iv) save production time for tasks such as mirroring the rig components, (v) automate scripted tasks and (vi) facilitate team work when other artists use the same naming convention for debugging scripts associated to the facial rig [Unay 2004].

A naming convention is usually defined using Camel Case and/or Underscore approaches. The Camel Case method is based on writing upper and lower case characters. The Underscore method uses the flat dash underscore character to separate words. These conventions are used to (i) provide speed and organization, (ii) be universal to the digital artists involved in a production and (iii) allow faster, reusable and more compatible scripting.

A number of principles help define a naming convention: (i) writing in English to achieve a more universal rig, (ii) reduce the number of characters to ease readability and computation, (iii) avoid special characters (e.g. accentuation and spaces) as these can compromise scripting and (iv) make the distinction in the name between the different structures in a facial rig using words for the type, description, number and location of the rig component. An example to identify a rig component for the left eyebrow using the Camel Case and Underscore methods could be *ctrl_brow_L.*

2.5 Facial Rigging for Key Frame Animation: Techniques

After the initial realization of how to integrate the rig with the facial model, the rigger begins building the rig mechanics and rig controls, respectively the deformation structures and the interactive user interfaces (UIs) of the facial rig. In analogy to a car, the rig mechanics would be the engine, wheels and tires whereas the rig controls would be the steering wheel, pedals and gears. In light of this metaphor, the rigger artist would be the car mechanic and the animator artist would be the car driver. Sections 2.5.1 and 2.5.2 respectively describe the details of the different types of mechanics and controls that a facial rig can contain.

2.5.1 Facial Rig Mechanics

The facial rig mechanics are the building blocks in 3D tools that riggers build to correctly deform a character's face. The most common rig mechanics are skeletons, or bones and their articulated joints, and blendshapes, or blendshape target models (BTMs). These techniques are described in detail following.

Skeletons (Bones)

Skeletons [Burtnyk and Wein 1976, Zeltzer 1982] are articulated bone hierarchies working in a parent-child relationship. Any deformation driven by an underlying skeleton (called a bone-driven deformation) [Komatsu 1988] uses a process called skinning [Lewis et al. 2000] that links the object being deformed to the bones deforming it. Skinning involves weighting the influence that each bone has in each vertex during deformation, or assigning the mechanical components of a rig to their corresponding groups of vertices. Skinning can have a smooth [Yang and Zhang 2006] or rigid [Labourlette et al. 2005] weighting of the influence that each bone has in each vertex. In most 3D tools, the skinning values range from 0 to 1 and are colored black to white or using cold to hot colors (e.g. blue to red). A black or a dark blue correspond to the lowest skinning value (0) and a white or dark red correspond to the highest skinning value (1). In-between colors are green, yellow and orange in a colored scheme or dark to light gray tones in a black and white scheme. The bones in a skeleton are connected to each other by their joints, which the rigger distributes inside and outside the 3D character's skin to establish the hierarchies of movement of the character [Talbot 2003]. Figure 2.11 illustrates a rig setup based on bones and joints placed inside a character's face.

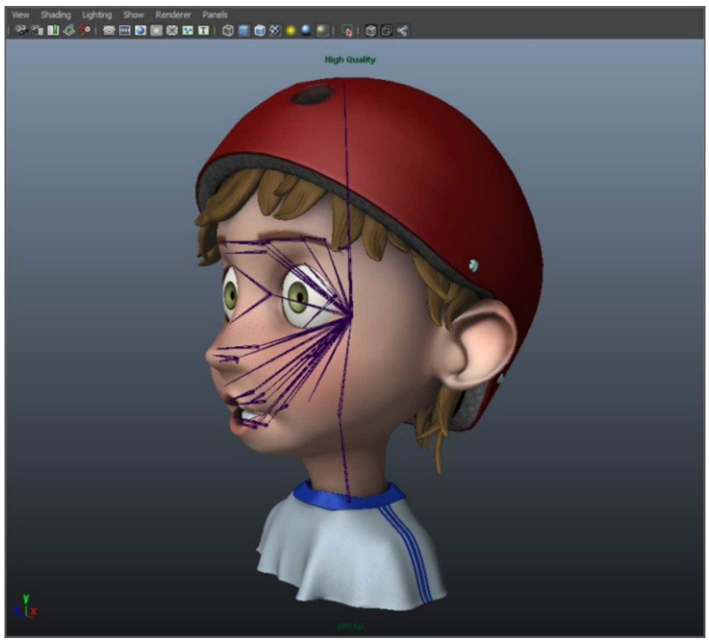


Figure 2.11: A skeleton composed of bones and their joints (colored in purple) placed inside the face of the character Nene in Maya (facial model is copyright of FaceInMotion).

During skeleton setup, the bones and joints can be moved freely in space by the rigger to position them accurately in the surface of the character's skin. Bones allow a rotation-based motion of the facial skin [Parke and Waters 2008] with hierarchically averaged skinning values. They are more suitable for arcing deformations [Ritchie et al. 2005] in larger facial regions with a higher polygonal density. Each vertex in a character's face is skinned by the rigger to be deformed only by the bones closer to it.

In a character's body, the joints act as the links between the bones to allow bending and twisting the entire skeleton. In a character's face, the skeleton needs to be highly articulated, but it is generally not necessary for it to be similar to a real person's skeleton. Only in a character's body the skeleton layout can be compared with the layout of a real person's skeleton. In the face there is no predefined guide on how to place the bones, because in reality the head is realized as the cranium and the jaw, which alone are not responsible for the large behavioral diversity of the face. Facial muscles are what drive the facial movements and deform a face. For this reason each rigger can have a subjective approach to facial rigging, because facial skeletons can be arranged in a variety of ways with the purpose to simulate the muscular activity of the face. In a character's body the bone structure follows predefined positions, because the rigger consults anatomy references that include a skeleton, to realize the physical basis and set the several bones and joints in the character's body.

For the skeleton to work properly, the rigger uses constraints [Badler et al. 1986], which allow to connect and restrict the position, orientation and scale of the bones. Constraints limit the transformations of bones using non-hierarchical relationships [O'Neill 2008]. These are setup by the rigger with the particular goal to prevent the animator from doing unexpected manipulations of the rig controls, which could cause incorrect mechanical movements and awkward deformations.

Blendshapes (BTMs)

BTMs were conceived by [Parke 1972] and popularized in the short film Tony de Peltrie [Bergeron and Lachapelle 1985]. They were improved since by [Bethel and Uselton 1989] and [Lewis and Anjyo 2010]. BTMs are duplicated facial models of an original one with equal topology [Maraffi 2003] but sculpted with a distinct shape in order to blend from the original to the duplicate via a weight parameter in the deformer. BTMs allow a per-vertex tweak of the facial geometry [Parke and Waters 2008] and are more appropriate for predefined linear skin deformations [Ritchie et al. 2005] in facial areas with less polygonal density, because generally they are more time-consuming to setup. Riggers build BTMs by (i) duplicating a 3D model, (ii) modifying its geometry to achieve a new pose (iii) and blend between the original and the duplicated geometry to achieve motion. Figure 2.12 illustrates a neutral pose and a blendshape pose sculpted to simulate a smile of the character Nene. The different duplicates of an original facial model are stored in the 3D tool to then interpolate between them, or blend them to achieve their animation, as in blending between the neutral and the blendshape pose seen in Figure 2.12.

Figure 2.12: Examples of blendshapes for the face of the character Nene in Blender; left: a neutral pose to represent an idle state for the character as a starting point for both rig and animation debugging; right: a blendshape pose sculpted to simulate a smile (facial model is copyright of FaceInMotion).

Blendshapes can be considered as predefined duplicated variations of an original character face. Riggers alter each blendshape of the facial model by modifying the positions of the vertices of the blendshape to obtain different expressions. The positions of the vertices can also be modified through sculpting the duplicated mesh, which involves altering several vertices together using a proportional influence brush modifier during the modeling process. This allows molding a face as if it was made of clay. The information stored in each duplicate face includes the variation of the coordinates of the model geometry. The different shapes are then controlled via sliders which range from 0 to 1 the influence of a given blendshape over the original neutral posed model. Animators can apply the total amount of a blendshape or only a partial value and even mix different blendshapes.

When facial animation requires the use of certain subtleties to increase realism (such as the use of wrinkles), riggers also use corrective blendshapes. These are not used as primary deformations but rather as secondary ones. They are called corrective because their purpose is to fix details which are difficult to simulate using another rigging technique.

Skeletons and blendshapes are the most common techniques because they are the ones that artists use the most. The reasons why artists use them more is that (i) they perform efficiently, (ii) they integrate well with each other and (iii) they are available in the major 3D tools (e.g. 3dsMax, Blender, Cinema 4D, LightWave, Maya and Softimage). Although adopting different names in each tool, they still work similarly. Table 2.1 shows the names used to identify skeletons and blendshapes in each major 3D tool. The overall predominant use of Maya for professional productions has made popular the names used in Maya. However, in recent years other names begin to be used more often as well, particularly the names in Blender due to the gradual increase of the Blender community of users.

3dsMax	Blender	Cinema 4D	LightWave	Maya	Softimage
Bones	Armatures	Joints	Skelegons	Skeletons	Skeletons
Morphs	Shape Keys	Morphs	Endomorphs	Blendshapes	Shape-Keys

Table 2.1. The names used by major 3D tools to identify the two primary rig mechanical techniques; top and bottom rows: respectively the names for skeletons and blendshapes.

Although used less, there are other complementary mechanical techniques for facial rigging (e.g. clusters, lattices and wires) which are briefly described next. These were not considered for the rig optimization methods presented in Chapters 4 and 5 to prevent inconsistency, because these techniques are (i) less used by digital artists [Bastos et al. 2012b] and (ii) they are not consistently available in all the major 3D tools. In contrast, bones and BTMs (i) are found in all the major 3D tools and (ii) they are the most efficient since artists use them the most for character facial animation [Bastos et al. 2012b], as seen ahead in this chapter in Sections 2.6 and 2.7.

Clusters

A cluster is a deformer with a visual handle which controls a group of vertices associated with the coordinate transformation relative to the "defined cluster origin location" [Parke and Waters 2008]. The transformations applied to the cluster handle "are also applied to the vertices in the cluster's deformation set" [Bibliowicz 2005]. Clusters are similar to joints, except they are not optimized to work in a hierarchy. They are more useful for the rigger to adjust individual groups of vertices during animation (e.g. wrinkles), which can also be achieved using individual joints.

Lattices

Lattices, or FFDs (Free-Form Deformations), were created by [Sederberg and Parry 1986] as a technique for deforming solid geometric models by editing a box-shaped flexible grid of control points that surrounds the object to be deformed like if it were a soft plastic. The object embedded also needs to be flexible in order to be sculpted by the rigger using the lattice control points. The lattice technique is more rarely used in facial rigging (see Sections 2.6 and 2.7).

Wires

The wire approach was created by [Singh and Fiume 1998], in which two curves of identical parametric range, called the base and the deforming wires, “measure the vector difference between corresponding points” [Bibliowicz 2005] on the deforming wires and apply that deformation to the vertices of the object being deformed “within the drop-off distance from the base wire” [Bibliowicz 2005]. After this technique is setup by the rigger, the animator can control the deformation caused by the wire approach via the manipulation of the curve's points, keeping their visual continuity. The wire technique is less used in facial rigging than bones and blendshapes (see Sections 2.6 and 2.7).

2.5.2 Facial Rig Controls

In order to add movement to a character's face, besides causing deformation, the rig also needs to be realized as a structure to be manipulated by the animator. The rig controls act as the user interface (UI) built by the rigger to allow animators to pose and key frame the character's face. The UI of the rig can be realized as layers of controls that allow for user interaction. The interaction is achieved with software based user interfaces and the UI rig definition can be understood as a solution within the rig’s system software.

In the mid 80s, 2D controls were used to handle 3D objects [Nielson and Olsen 1986]. In the late 80s, [Witkin and Kass 1988] introduced the concept of high-level control interfaces for key frame animation. In the early 90s, [Conner et al. 1992] and [Strauss and Carey 1992] introduced the notion of three dimensional widgets to be used as a direct manipulation interface in 3D that was later adopted by commercial packages. In the early 00s, after Pixar released Toy Story 2 (1999), direct manipulation was implemented to allow animators to easily (i) identify the controls in each facial region and (ii) manipulate them more rapidly and effectively in 3D space [Strauss 2002, Osipa 2003]. A common technique in Blender to ease the direct manipulation in a user interface is using drivers, or driven-keys in Maya, which are relationships that can be set between the controls and the mechanics. A driver allows to connect the transformation of an object to the transformation of another object without the specific need to set a parent-child relationship between the two. The object being driven is usually called the driven object, which in a facial rig could be a rig mechanical component positioned inside or in the surface of the character's face (e.g. a driven-bone or a driven-blendshape). The object driving the relation is usually called the driver object, which in a facial rig could be a rig control component positioned in a user interface (e.g. a driver-bone or a driver-control slider). Figure 2.13 illustrates a driver setup in Blender.

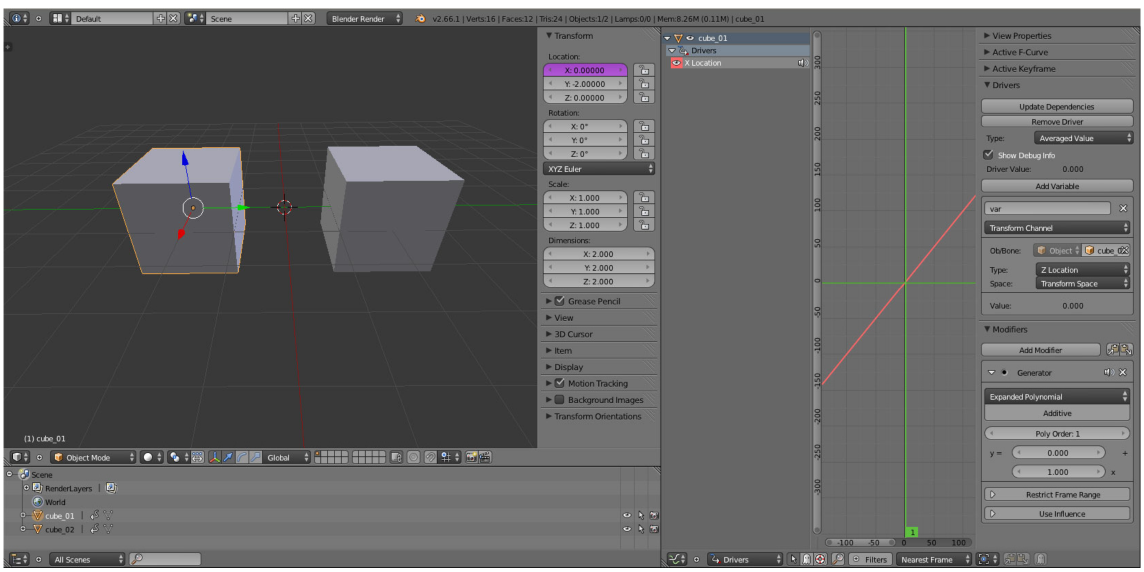


Figure 2.13: An example of a driver setup in Blender that causes the translation of a cube to affect the translation of the other cube via the use of a variable, a variable target and a modifier with a coefficient to control the influence of the driver.

In the case illustrated in Figure 2.13, the X location of the object *cube_01* is driven by the Z location of the object *cube_02*. A driver added to the X location of *cube_01* in the 3D View window of Blender was configured with a variable target being the *cube_02* and a modifier with a coefficient to control the influence, or the intensity of the driver relationship.

The approaches to the user interface for character facial rigging can be resumed in the three following user interaction modes: (i) window-based (WIN), (ii) camera-based (CAM) and (iii) viewport-based (VIP). These modes are described in detail following.

Window-based (WIN)

The controls in this type of interaction mode are located in the areas of the 3D tool interface not relative to the 3D space (e.g. the properties window in Blender or the channel box in Maya). A window-based UI provides direct input of values in different ways, for instance, [Holly 2006] proposed an UI in a separate window with buttons and sliders to select and modify the controls located in the surface of the facial model. Bredow et al. [Bredow et al. 2007] configured Maya's channel box to display multiple categorized columns of attributes. Later, a slider-based UI based on FACS was presented by [Villagrasa and Susin 2009]. In general, the window controls include buttons and 1D control knobs known as sliders, which are numerical fields manipulated horizontally from side to side. Buttons are used by animators to activate/deactivate options (e.g. reset values). Sliders are used to trigger single axis spatial transformations by dragging the mouse over the slider or entering specific values, which is possible by clicking and typing a value or copying and pasting from one slider to another. Sliders usually range from 0 to 1 (e.g. to simulate an eye blink) or -1 to 1 (e.g. to simulate an eyebrow moving upwards or downwards from a center default position). The window mode (WIN) is considered exact because (i) it presents the ability for the user to set numerical values in sliders, (ii) it provides textual information of the required facial deformation in a per slider basis, therefore not occupying the 3D space or the 3D facial model skin surface and (iii) it can be relocated in different kinds of predefined windows of the 3D tool interface. Figure 2.14 shows a list of sliders to control the facial animation of an anthropomorphic character.

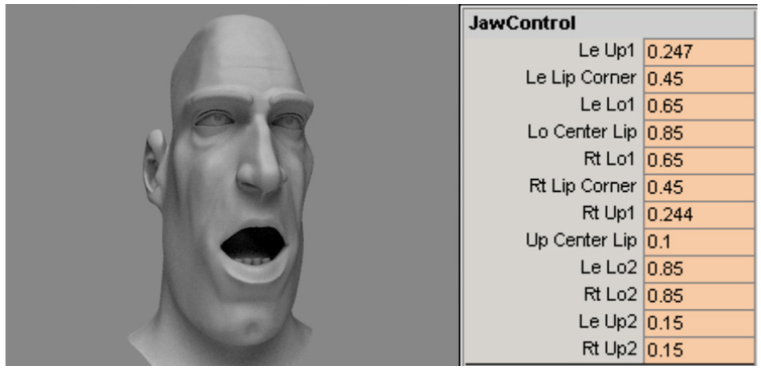


Figure 2.14: An example of a window-based UI by [Schleifer et al. 2002] with slider controls located in a window of the 3D tool.

Camera-based (CAM)

In this type of interaction mode the controls are located in the areas of the 3D tool interface relative to the 3D space (e.g. the 3D view in Blender or the perspective window in Maya). These controls are schematic and seen through a bi-dimensional camera view. They combine 1D control knobs, which can be manipulated in a single axis of transformation, with 2D square or rectangle-shaped widgets that artists usually call osipas [Osipa 2003]. An osipa holds a control knob which can be manipulated in two transformation axes, both horizontally and vertically. These sets of controls are constrained to a square-shaped delimiter visible in the viewport. The square-shaped delimiter prevents opposite behaviors from being combined and also acts as a boundary to give the animator a sense of the control limits [Osipa 2007, Neale 2008]. The camera mode (CAM) is considered intuitive because (i) it presents a schematic view of the controls, providing an easier perception of the boundaries of control due to each manipulator being constrained to a delimiter, (ii) each delimiter of each control in the scheme can be relocated and rescaled if configured to do so and (iii) it allows the user to deform the facial model independently of the angle that the character is being observed from. Figure 2.15 shows the Digital Emily Project by [Alexander et al. 2009], which used the same technique of a 2D constrained space with an anthropomorphic control UI.

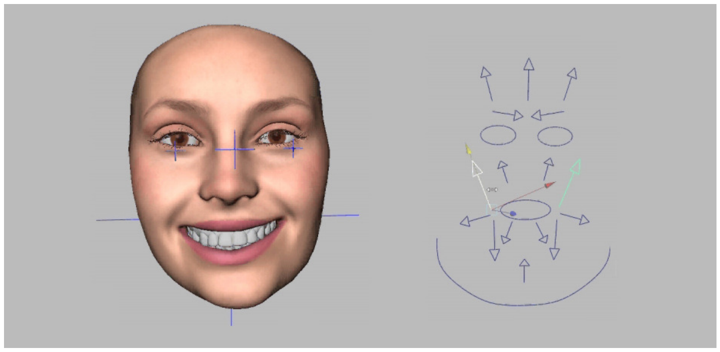


Figure 2.15: An example of a camera-based UI by [Alexander et al. 2009] with 2D controls located side-by-side to the facial model in the viewport.

Viewport-based (VIP)

The controls in this type of interaction mode are also located in the areas of the 3D tool interface relative to the 3D space (e.g. the 3D view in Blender or the perspective window in Maya). The main difference is that they are not confined to a bi-dimensional camera view. The spatial manipulation of these controls, called gizmos, is not constrained to any visual boundaries realized as interface delimiters, therefore they can be manipulated freely in all axes of transformation in case this is necessary. It is common to use 3D objects as a control gizmo to manipulate a rig in these systems. These objects are called gizmos because they are three dimensional widgets which can move in each of the three axes of transformation. A viewport-based UI is presented by [Komorowski et al. 2010] where the controls can be linked to the face behaviors with a 3D space volume. Other artists combine high level viewport controls to manipulate the major muscular groups and low level viewport controls for the more subtle facial movements. Gizmos are usually placed on the surface of the facial model skin to allow for a detailed animation, even to the level of creating specific wrinkles in the face. These controls can be placed by the rigger in the surface of the skin as well as inside the character's facial skin (e.g. inside the mouth to control the tongue). The viewport mode (VIP) is considered straightforward because (i) the deformation triggered by manipulating a gizmo is directly acknowledged by the user, due to the gizmo being located on the surface of the facial model or nearby the facial region that it is assigned to control, (ii) it provides more concentrated transformation, since each gizmo can hold more facial behaviors and (iii) the manipulation can feel more natural due to the gizmos being less constrained, as objects tend to be in the real world. Figure 2.16 illustrates two approaches of viewport controls.

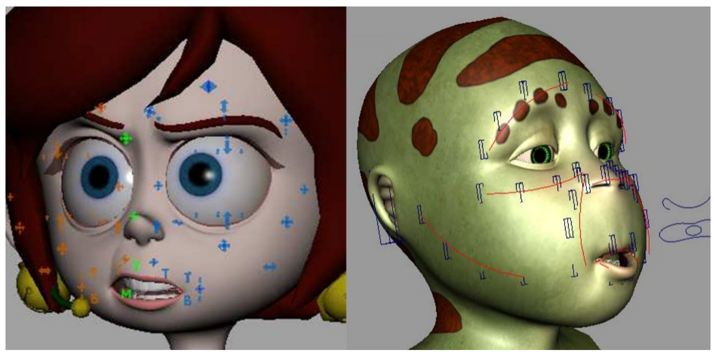


Figure 2.16: Examples of viewport-based UIs; left: gizmos located over the skin of a character by [Komorowski et al. 2010]; right: gizmos and curves located over the skin of a character by [Grubb 2009].

Each type of control has benefits and limitations. For instance, window-based sliders are ideal for single facial skin deformations such as the FACS isolated behavior that indicates the chin raise, the AU 17 - Chin Raiser [Ekman and Friesen 1978]. Camera-based controls like osipas allow the rigger to mix opposite behaviors which should not occur at the same time (e.g. a smile for one direction of the manipulator and a frown for another manipulation direction). Viewport-based controls allow a free interaction in space and can hold a particular behavior per each spatial direction. An example can be to assign to a single gizmo the jaw behaviors roll (open/close), yaw (side to side) and thrust (frontwards or backwards), being respectively assigned to the gizmo's Z, X and Y translations (given a Z-up default coordinate system [O'Neill 2008]).

2.6 Facial Rigging for Key Frame Animation: Literary Approaches

This section presents a detailed historical overview of the literary facial rig approaches built by several artists for key frame animation since [Maestri 1996]. It begins with references that encompass the entire face and then presents approaches in a per facial region basis. These include facial rig mechanics and controls that the artists developed in the last 15 to 20 years.

2.6.1 General Anatomy

Approaches that encompass the entire face use either one or a mixture of rigging techniques. For instance, Harkins [Harkins 2005] distributes bones over the face of a stylized character in Maya and also builds different blendshape heads of the character, as seen in Figure 2.17. The combination of bones and blendshapes allows a greater diversity of control over the different behaviors of the character's face because bones can be used in the form of control gizmos and blendshapes can be triggered by sliders.

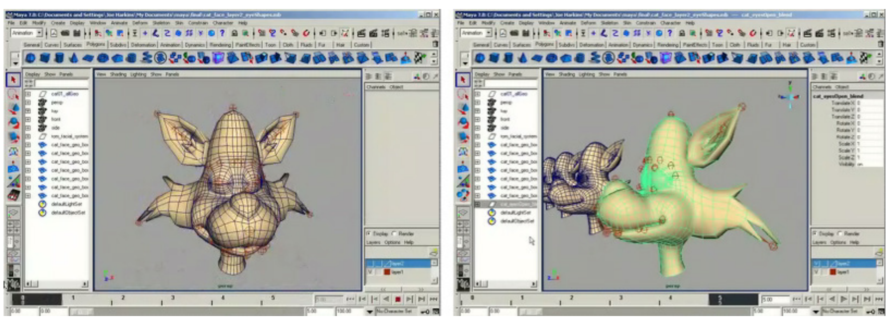


Figure 2.17: Approach using bones and blendshapes by [Harkins 2005]; left and right: respectively bones placed in the character's face and multiple blendshape heads.

The artist Neale [Neale 2008] builds a web-shaped facial skeleton for a realistic character in 3dsMax. The intersection areas of the bones are connected to gizmos via fall-off properties. When manipulated by the animator, the gizmos stretch the surrounding bones which in turn deform the skin (see Figure 2.18).

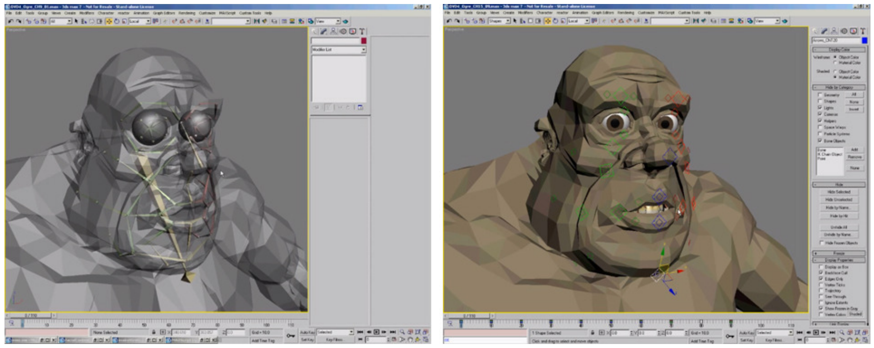


Figure 2.18: Web-shaped facial bone structure approach by [Neale 2008]; left: x-ray view of the underlying bones; right: the viewport control gizmos.

Athias [Athias 2010b] adds curves to a bone-based facial rig of a stylized character in Maya. The curves are spread over the facial regions using the guidelines in FACS [Ekman and Friesen 1978] and are added as influence objects to the facial skeleton, allowing the animator to deform the model using the vertices of the curves as gizmos. The use of curves in facial rigging can be referred as curve-based facial rigging because it relies on the animator manipulating the curves.

In [Wavefront 2003] BTMs are designed to resemble the facial muscles in a realistic face in Maya. Maraffi [Maraffi 2003] also separates the expressive areas of the face into BTMs in order to mix them to create a variety of facial expressions.

Ritchie et al. [Ritchie et al. 2005] resort to consulting FACS [Ekman and Friesen 1978] to decide whether to use bones or BTMs for each facial region in a realistic character in Maya. Bones are used in areas with a curved motion and BTMs in areas with a linear motion, because bones have a natural rotation effect and BTMs behave with a linear transition.

According to Kazuo and Pinheiro [Kazuo and Pinheiro 2011], BTMs can be used for general basic deformations and bones for more precise local deformations (see Figure 2.19).

To better see viewport controls, some artists [Harkins 2005, Athias 2006, O'Neill 2008] use two cameras. One is parented to the rig controls and aims directly at them, allowing the animator to handle the controls through it. The other is parented to the root bone of the facial skeleton and aims at the character's face, allowing the animator to see the facial deformations.

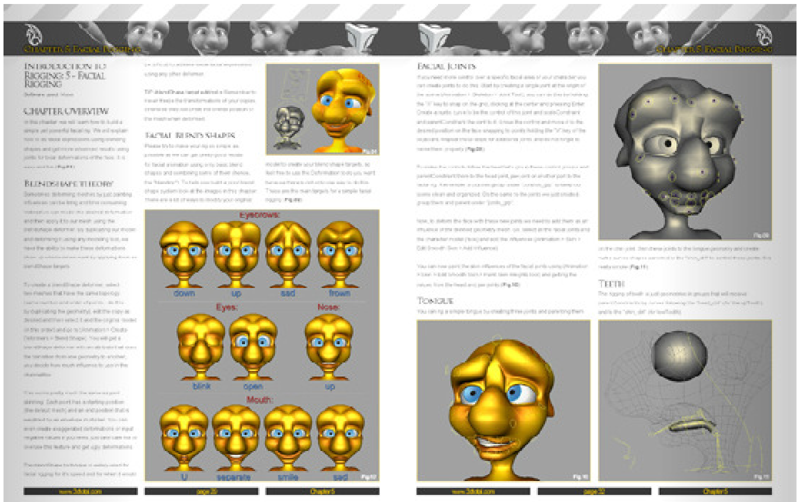


Figure 2.19: Approach using global and local deformations by [Kazuo and Pinheiro 2011]; left: BTMs for global deformations; right: joints for local deformations.

Ritchie et al. [Ritchie et al. 2005] and Athias [Athias 2006] take advantage of the concept of using two cameras and create curve-based interfaces in Maya positioned side-by-side to the face of the character and resembling the character's facial features (see Figure 2.20). Both these camera-based interfaces are optimized by locking and hiding unneeded parameters and also adding limitations to the motions of their 2D controls.

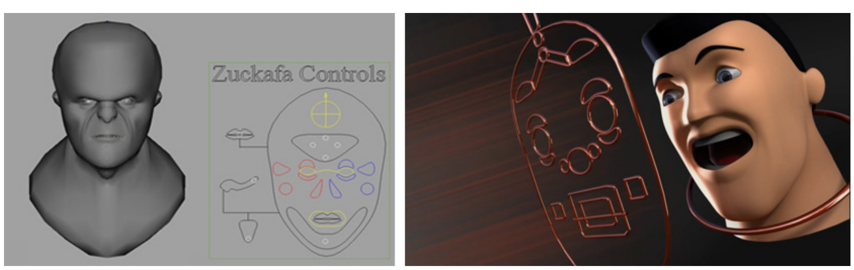


Figure 2.20: Camera-based UI approaches located side-by-side to the characters' faces; left: the Zuckafa 2D control interface by [Ritchie et al. 2005]; right: a 2D-based interface by [Athias 2006].

Holly [Holly 2006] creates a window-based user interface in Maya, seen in Figure 2.21, which holds buttons and sliders. The buttons ease the selection of the different controls laid in the facial skin of the stylized character and the sliders trigger the facial deformations.

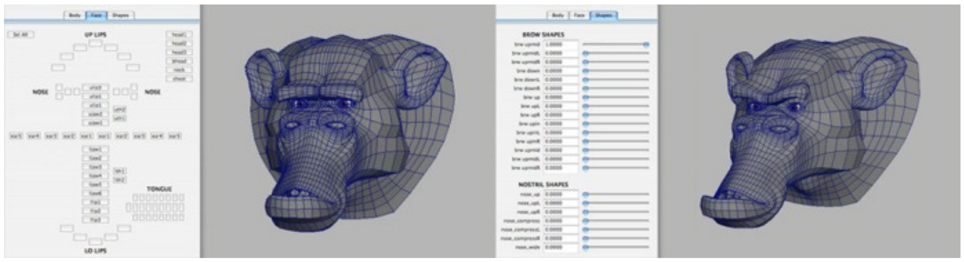


Figure 2.21: Window-based user interface approach by [Holly 2006]; left: buttons for display of controls; right: slider control knobs to handle facial behaviors.

Neale [Neale 2008] sets a camera interface in 3dsMax with sliders and osipas to animate a realistic character face. Similar controls are used by the artist Skonicki [Skonicki 2008] to animate the face of a realistic character in Maya for videogame animation. Highlights of both these approaches are shown in Figure 2.22.

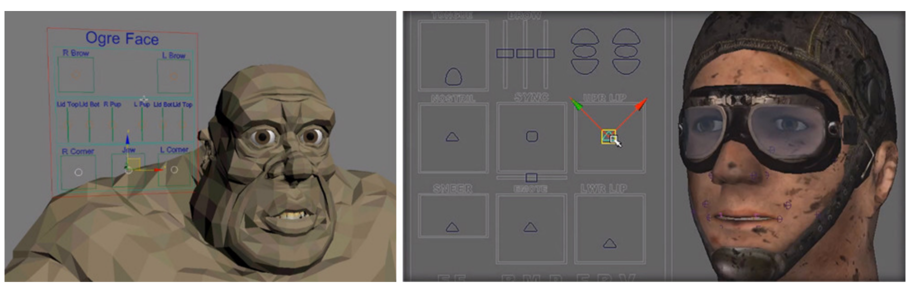


Figure 2.22: Camera-based approaches based on sliders and osipas; left: the Ogre facial UI by [Neale 2008]; right: a 2D rig interface for videogame animation by [Skonicki 2008].

Villagrasa and Susin [Villagrasa and Susin 2009] set a multilayered facial bone system for a realistic character in 3dsMax. It allows moving single or combined groups of muscles defined in FACS [Ekman and Friesen 1978]. The location of each bone is based in the FAPs of the MPEG-4 specification [Pandzic and Forchheimer 2002]. The interface is composed of many window sliders that allow controlling the intensity of the muscular activity triggered by each bone. The multilayered facial bone system presented by [Villagrasa and Susin 2009] is one of the few systems exploring in great detail the muscular activity in FACS but the user interaction is solely window-based and therefore cluttered with controls.

2.6.2 Eyebrows

The motion of a character's eyebrows can vary a great deal, especially if modeled separately from the main facial model. For instance, [Harkins 2005] configures a character facial rig with separated eyebrows using a script that attaches the polygons in the eyebrows to the facial skin. For most characters it is sufficient to mix the most commonly used techniques. Maestri [Maestri 1999] recommends using BTMs and Schleifer et al. [Schleifer et al. 2002] propose a generic group of five BTMs for the eyebrows, which are seen in Figure 2.23.

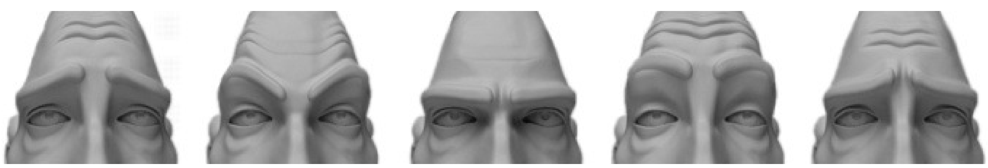


Figure 2.23: Blendshape examples for eyebrow control in a stylized character by [Schleifer et al. 2002].

Holly [Holly 2006] creates two solutions for the eyebrows of a stylized character in Maya: (i) a bone hierarchy set along the length of each eyebrow and (ii) a ribbon of polygons placed along the geometry of the eyebrows with joints locked to each polygon, the ribbon is then skinned to an existing head bone so that it follows the motion of the head keeping the independent movement provided by each joint.

Osipa [Osipa 2010] builds BTMs for a realistic character that support outward, inward and squeeze motions. These are illustrated following in Figure 2.24.

Figure 2.24: Blendshape examples for eyebrow control in a realistic character by [Osipa 2010].

Athias [Athias 2010a] distributes single joints as gizmos in the air-suspended eyebrows of a stylized character in Maya which is seen in Figure 2.25. To prevent placing each joint by hand, a script creates a joint per each selected vertex by (i) identifying the selected vertices, (ii) creating a cluster on each vertex, (iii) clearing the selection, (iv) creating the joints, (v) point constraining the joints to the clusters and lastly (vi) deleting the clusters and constraints.

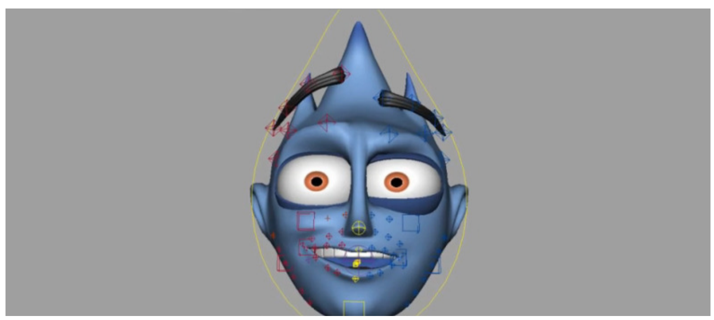


Figure 2.25: The air-suspended eyebrows of a stylized character rigged using a distribution of joints by [Athias 2010a].

2.6.3 Eyelids

Maestri [Maestri 1999] recommends using BTMs as a straightforward approach to blink the eyelids. But in characters with rounder eyeballs the linear motion provided by BTMs is contrary to the arcing motion of their eyelids. As an alternative, Maraffi [Maraffi 2003] sets one bone per each eyelid with the root joints centered in the eyeball and each tip joint located in the edge of its eyelid, as illustrated in Figure 2.26. Maraffi's [Maraffi 2003] setup allows to rotate each bone to open, close and shape the eyelid around the eyeball.

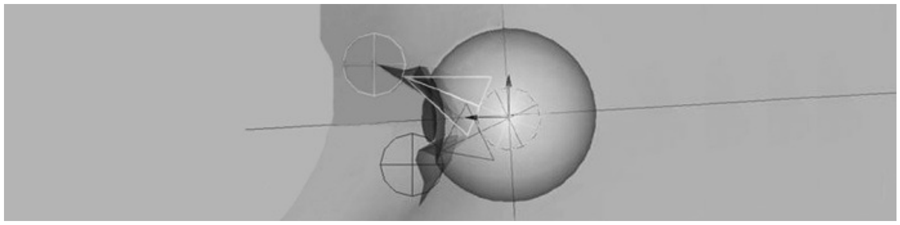


Figure 2.26: An early approach to control the upper and lower parts of the eyelids using bones by [Maraffi 2003].

Schleifer et al. [Schleifer et al. 2002] create two attributes in Maya to control the eyelids of a stylized character for the film Ice Age (Blue Sky Studios, 2002). The first produces a fast blink and the second controls the position of the blinking line along the eyeballs (see Figure 2.27). For more realistic characters, driven-keys are used instead to connect the orientation of the eyeballs to four BTMs that modify the shape of the corners in the eyelids. For instance, if the eyeball aims upwards and to the right, a BTM of the upper right corner eyelid is triggered.

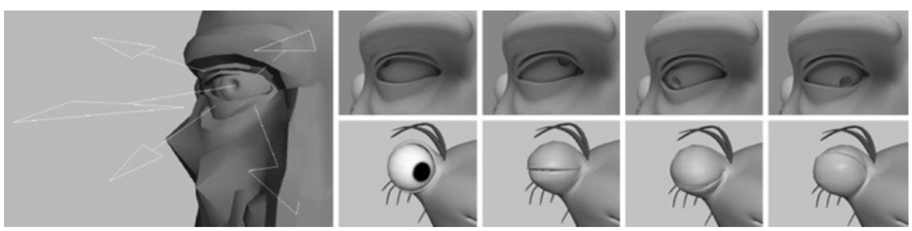


Figure 2.27: Approaches to control the eyelids by [Schleifer et al. 2002]; left: eyeball sight vectors; top right: realistic eyelid movements driven by the direction of the eyeball vectors; bottom right: stylized eyelid movements around spherical-shaped eyeballs.

Other artists [Gorden 2005b, Athias 2006] use BTMs for blinking but with the advent of causing a linear motion of the eyelids skin. To fix this issue, Unay and Grossman [Unay and Grossman 2005] add new poses in-between the first and last BTMs to obtain a more realistic curved motion of the eyelids, which is proportional to the number of in-betweens used.

In the film Surf's Up (Sony Pictures Animation, 2007), [Bredow et al. 2007] control the eyelids of stylized characters using one joint per each polygonal edge in the eyelids skin that is closer to the eyeballs. Using one joint per each polygonal edge generates many gizmos along the edges of the eyelids but it also allows the animators a greater control "to achieve any desired pose" [Bredow et al. 2007].

Athias [Athias 2011] creates a sculpt deformer per each eyeball in a realistic character in Maya. The orientation of the deformer is connected to the orientation of the eyeballs via the connection editor so that the skin in the eyelids is pushed more or less according to the rotation of the eyeballs.

Vasconcelos [Vasconcelos 2011] adds an inverse kinematics (IK) constraint to a realistic eyelid bone setup in Blender. This inverse kinematics approach is similar to that of Maraffi's [Maraffi 2003], but in this case smaller bones are added in the extremities of the deforming bones in order to be used as motion targets. The translation of these smaller bones is then controlled via scaling a controller bone located in front of the eyes which acts as a scalable gizmo with a limited transformation to prevent it to be scaled more than necessary.

2.6.4 Eyeballs and Pupils

Rigging the eyeballs of a 3D character is usually a straightforward and appealing task due to (i) their spherical-like shape, (ii) their fairly limited behaviors and (iii) the control feedback of the eye gaze. Early approaches to rig the eyeballs relied on driving the coordinates of texture maps assigned to them, which was less appealing in terms of interaction.

The most popular approach is likely the use of target objects located in the viewport to which the eyeballs point at via aim or inverse kinematics (IK) constraints. To control the eye gaze the animator only has to move the targets in 3D space. To track the orientation of the eyeballs during animation, Maestri [Maestri 1999] used cone-shaped wireframe objects attached to the eyeballs that mimic their orientations. These cones act as flashlights to help realize where the character is looking at in space. Maestri [Maestri 1999] improved further on this approach using a helper object per each eyeball and a master helper object for both eyeballs. The improved setup allows the animator to re-orient each eyeball individually by moving its corresponding helper object and also re-orient both eyeballs using the master helper object.

The target-based approach has been widely used since and improved by many artists [Graft et al. 2002, Gorden 2005b, Harkins 2005, Athias 2006, Maestri 2006, Athias 2010a, Athias 2011]. For instance, Athias [Athias 2006] parented the eye controls to a head bone to keep the eyeballs looking straight in case the orientation of the head changes. Harkins [Harkins 2005] used a setup of two bones to deform each eyeball of a stylized character in Maya. While one bone allows scaling the eyeball the other could stretch it out of the eye socket. Arrow-shaped color-coded gizmos are then used together with lines of sight to keep track of their locations in space. The line of sight begins in the center of its corresponding eyeball geometry and ends in its respective control gizmo, as seen in Figure 2.28.

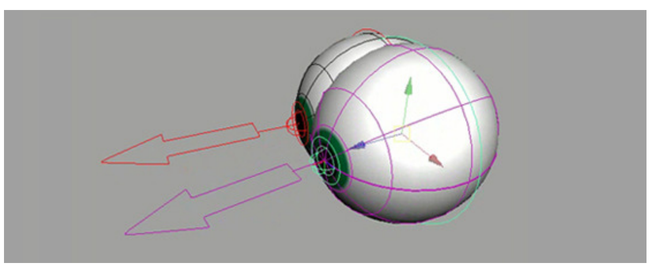


Figure 2.28: Arrow-shaped gizmo control knobs for the eyeballs and their respective lines of sight by [Harkins 2005].

Neale [Neale 2008] created point helpers instead of bones to control the eyeballs of a realistic character in 3dsMax. Each eyeball is parented to a point helper located in the center of the eyeball and parented to a head bone. Both helpers are then aim constrained at two color-coded text curve gizmos located in front of the eyeballs in space (see Figure 2.29).

Athias [Athias 2011] used bones to control the eyeballs of a realistic character in Maya.

Figure 2.29: Viewport text curves placed in front of the facial model to control the orientation of the character's eyeballs by [Neale 2008].

The bones have inverse kinematics (IK) handles that begin in their root joints and end in their tip joints. Each IK handle is point constrained to curve-shaped gizmos.

When rigging the eyeballs it is often also necessary to rig the pupils. In realistic characters the pupils only need to dilate or compress but that single visual effect can make a difference and it continues to be used in character production (e.g. recently in the videogame The Last of Us [Simantov 2013]).

In stylized characters the pupils can extend with the squash and stretch of the eyeball. BTMs are a usual approach to handle pupil dilation and compression. The artists [Maestri 1999] and [Vasconcelos 2011] used BTMs to scale the vertices in the pupils inwards and outwards to respectively reduce or enlarge the size of the pupil. Using BTMs to scale the vertices in the pupils is fast but it generates two BTMs per each pupil, which can become cumbersome for an animator to control in case the character has more than two eyeballs. A solution is to drive the BTMs using the transformation of controller objects (e.g. driven-keys linking the object's positive and negative scale values to the dilation and compression of the pupils, respectively). Neale [Neale 2008] manipulates the pupils of a realistic character in 3dsMax using a window-based slider with a numeric input field in the properties panel in the interface of 3dsMax. The slider is displayed only if the eyeballs viewport gizmo is selected.

2.6.5 Nose

The behavior of the human nose is confined to a limited muscular activity defined in FACS [Ekman and Friesen 1978]: sniff, snarl and nostrils dilation or compression. BTMs can be an adequate solution to rig the nose since they can be combined with other deformers or bone-based approaches and be driven by sliders, osipas or gizmos.

Graft et al. [Graft et al. 2002] manipulate the nostrils of a realistic character in Maya by placing a sculpt deformer inside each nostril. The animator can add positive or negative scale values to the deformer to respectively increase or decrease the size of the nostril.

A bone-based approach is presented by Holly [Holly 2006] in Maya to handle the nostrils of a stylized character illustrated in Figure 2.30. A cross-shaped bone layout is set inside the nostrils with the tip joints located in the nostrils' extremities to deform their opposite corners. To prevent the animator from selecting each joint, an osipa is positioned above each nostril in the viewport to translate each joint depending on the translation of the osipa's handle. Then a number of BTMs are also built to flare, expand, compress and pull the nostrils.

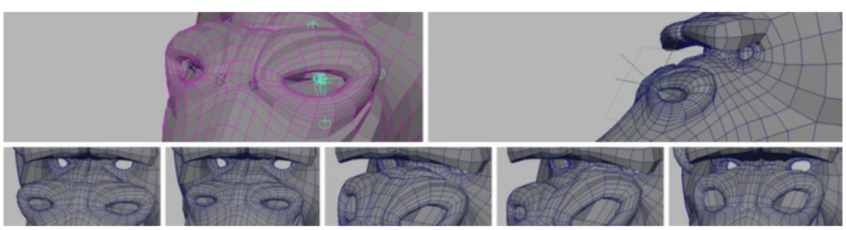


Figure 2.30: Approaches to control the nose by [Holly 2006]; top left: crossed bones to control the nostrils; top right: nostril handle constrained to a square-shaped delimiter; bottom: BTMs for the character's nostrils.

Holly [Holly 2006] manipulates the highly protruded nose of the former character using a curve placed along the length of the nose and a wire deformer skinning the curve to the geometry. Clusters are then assigned to the vertices of the curve, allowing the animator to smear the nose by manipulating each cluster as control gizmos (see Figure 2.31).

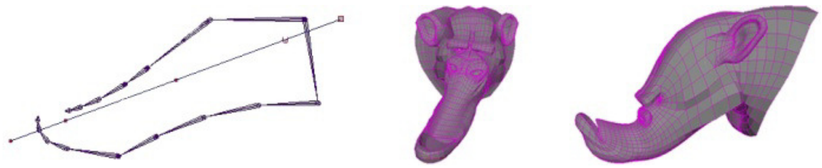


Figure 2.31: Approach to control a protruded nose by [Holly 2006]; left: bones and control curve of the protruded nose rig; middle and right: resulting flexible deformations.

Athias [Athias 2010b] sets a single joint to freely push and orient the nose of a stylized character in Maya. Skeletons are used again by [Athias 2011] to control a realistic character, with the root and end joints positioned respectively inside and outside the nostrils. Curves are placed as gizmos and linked to the bones via Maya's connection editor (see Figure 2.32).

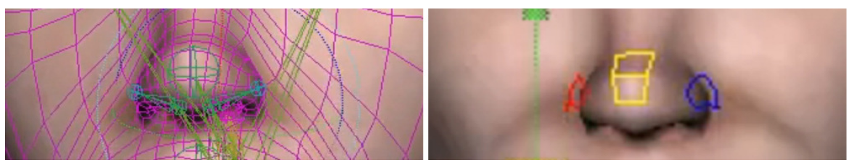


Figure 2.32: Approach to control a human nose by [Athias 2011]; left: bones positioned to deform the nostrils; right: gizmo curves positioned to manipulate the nostrils' bones.

2.6.6 Ears

Ear movements are "extremely rare in humans" [Pelachaud et al. 1994] and little noticeable. In fact, FACS [Ekman and Friesen 1978] does not consider any muscular activity for the human ear and typically only hyper-realistic or stylized characters have their ears rigged, because these can extend, bend, roll and/or twist. A common and straightforward solution to deal with such behaviors is using a bone hierarchy placed along the length of the ear. Gorden [Gorden 2005b] used this approach in LightWave to control the ears of a stylized character seen in Figure 2.33. The bones along the hierarchy have the same orientation of the root bone and define a coplanar alignment of the entire hierarchy to keep the pitch as the major rotation axis. The result is a consistent deformation of the ears when the bones are rotated.

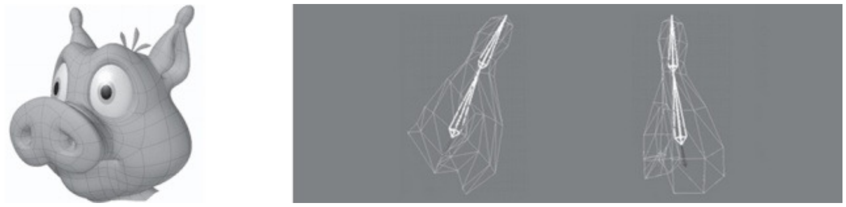


Figure 2.33: Bone hierarchies positioned along the ears by [Gorden 2005b]; left: morphology of the character's ears; right: bone hierarchies to deform and control the ears.

2.6.7 Cheeks

Skeleton-based approaches such as that presented by [Maraffi 2003] rely on a distribution of bones along the muscles of the cheeks. Otherwise, the animation of the cheeks is commonly achieved using a reduced number of BTMs based on the muscular activity found in FACS [Ekman and Friesen 1978]. These include triggering behaviors such as the cheek raise, lower, inflate or compress.

Miller [Miller 2004] creates a BTM to adjust the deformation of the skin in the cheeks of a hyper-realistic character in Maya. This BTM simulates the stretching of the skin in the cheeks during the opening of the character's mouth, making the face look more realistic. The BTM is triggered using driven-keys when the mouth opens.

Holly [Holly 2006] controls the cheeks of a stylized character in Maya according to a smile. A distance measurement system drives the cheeks motion whenever the animator manipulates single joints to which the corners of the character's mouth are smoothly skinned. If a joint reaches the proximity of its corresponding cheek, a cluster in the cheek's vertices is translated backwards towards the ear, resulting in a cheek bulge.

Athias [Athias 2011] sets BTMs to raise the cheeks of a realistic character in Maya. When the character smiles the BTMs are triggered via a distance node linking locators placed in the corner of the lips and point constrained to clusters handling the vertices in that region. Moving the clusters upwards activates the BTMs controlling their corresponding cheeks.

2.6.8 Jaw

Rigging a character's jaw is an appealing task due to (i) the challenge of correctly deforming the facial skin as the jaw opens and (ii) the appealing animation feedback obtained. Artists use mostly bone-based approaches to control a character's jaw because bones deform the skin in an arcing motion, therefore causing a more natural rotation-based movement of the jaw.

Early approaches to rig the jaw of animated characters were made by Maestri [Maestri 1999] in 3dsMax via BTMs. But many BTMs are required to support the movements of the jaw (e.g. move sideways, thrust and clench) and the penalty is a great consumption of setup time. As an alternative, Maestri [Maestri 2001] positioned a bone inside the geometry of the lower part of the face and smoothly skinned the nearby vertices to the bone. If skinned precisely, a single bone can handle the entire behavioral set of the jaw, causing less distortion of the facial skin.

Miller [Miller 2004] introduced an innovative rig approach for the jaw of a hyper-realistic character in Maya. It involves using upper and lower jaw bones, respectively smooth skinned to the upper and lower parts of the face. The upper and lower jaw rig setup provides an independent manipulation of the upper and lower parts of the head and a rapid control tweaking of each part to the other.

When rigging the jaw using bones, placing accurately the root and end joints helps the 3D tool do a more efficient default distribution of the skinning weights, which saves the artist time and helps achieve a more accurate anatomical movement. Generally, the root joint of the lower jaw bone is placed below the character's ear [Miller 2004], which is near the pivot point from where the lower jaw rotates.

Neale [Neale 2008] built two viewport approaches to control a jaw bone rotation for a realistic character in 3dsMax: (i) the translation of an osipa's handle or (ii) the translation of a gizmo located slightly in front of the jaw.

Athias [Athias 2010b] used a bone to deform the jaw of a stylized character in Maya and controlled it using a box-shaped curve gizmo aligned to the end joint of the bone. The box's transform pivot matches the root joint which in turn is orient constrained to the box. The artist Vasconcelos [Vasconcelos 2011] also uses a single bone to control the jaw of a realistic character in Blender. The single bone is then used to set a number of poses illustrated in Figure 2.34, which are then recorded in separate blocks of animation for the animator to trigger and mix individually.

Figure 2.34: Different poses for the jaw bone by [Vasconcelos 2011]; left to right poses: neutral, drop, clench, sideways, thrust and backwards.

2.6.9 Lips

The number of expressions of the lips varies according to the level of realism required. A smile or a frown are examples of expressions, but there are also visemes. The visemes are the visual representations of the phonemes, which are the smallest units of speech [Granberg 2009]. Minimum sets of visemes are usually built using BTMs because they allow a precise control over every vertex in the lips of the facial model.

Maestri [Maestri 1999] highlights that any BTMs used for the smile and frown expressions need to be built separately for both the left and right lip corners, since a smile can be asymmetrical. Graft et al. [Graft et al. 2002] define the BTMs seen in Figure 2.35, which support the visemes for the main vowels and consonants in the English language, based on those set originally by Blair [Blair 1996]: AI, E, O, U, CDGKNRSYZ, MBP, LThD, FV.

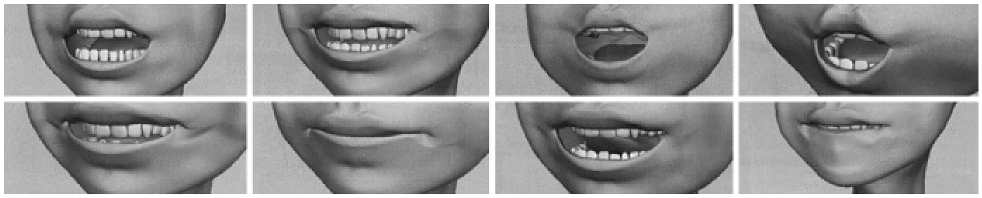


Figure 2.35: Examples of BTMs to simulate the main vowels and consonants in the English language by [Graft et al. 2002].

Schleifer et al. [Schleifer et al. 2002] rig the lips of stylized characters for the film Ice Age (Blue Sky Studios, 2002) by mixing BTMs with expression driven joints located along the length of the lips. Miller [Miller 2004] emphasizes that a generic lips rig should include joints as gizmos, since the lips motion is mainly rotation-based and prone to a direct manipulation, which is more adequately provided by joints.

Gorden [Gorden 2005b] creates BTMs for the lips of a stylized character in LightWave, some of which are mouth expressions and visemes relative to the phonemes A, E, S, W, O, M, L, F and TH. The main benefit of the former group of BTMs is mixing them to achieve a fairly large range of combinations. Athias [Athias 2006] also uses BTMs, seen in Figure 2.36, to control the lips of a realistic character in Maya.

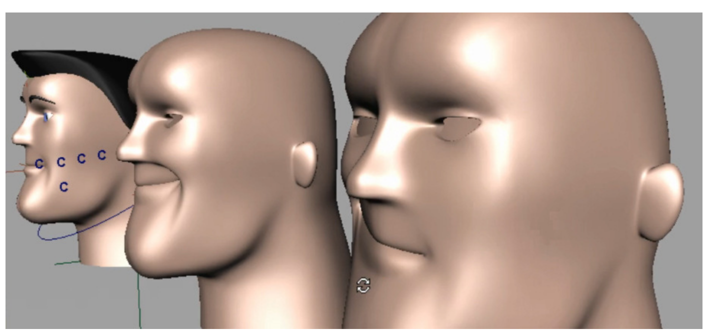


Figure 2.36: Examples of BTMs to simulate realistic looking lips by [Athias 2006].

O'Neill [O'Neill 2008] uses a mapping chart to build a group of BTMs for the mouth expressions of realistic characters. In the industry, these charts help convert phonemes to visemes to keep their number reduced. Granberg [Granberg 2009] proposes a set of thirteen visemes to achieve an overall control of the phonemes in the English language, which are nearly forty [Fleming and Dobbs 1998]. The set, illustrated in Figure 2.37, is based in the archetypes originally used by Disney animators.

Figure 2.37: The Disney's archetypes-based visemes by [Granberg 2009].

To control the lips of a realistic character in Maya, Athias [Athias 2011] places curves along the upper and lower lips of a duplicate of a character's head. The duplicate head modifies the original character's head via a BTM and the vertices of each curve are controlled by the animator using clusters.

An example of a more complex and more realistic lips setup is the simulation of the sticky lips behavior. Sticky lips is the tendency of real-world lips to stick to each other toward the corners of the mouth when the mouth opens and closes [Osipa 2010]. Osipa's [Osipa 2010] approach to rig sticky lips produces the results seen in Figure 2.38. The sticky lips behavior is not very often implemented because (i) usually only realistic characters require it, (ii) it is a behavior more noticeable for close-up shots and (iii) it is rather difficult to implement.

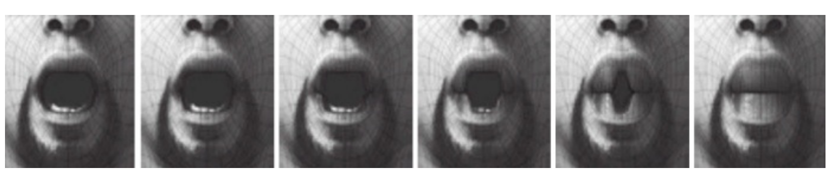


Figure 2.38: The deformation sequence of the sticky lips approach by [Osipa 2010].

2.6.10 Tongue

A tongue rig is usually not required for characters seen from a distance, because at a far range the tongue movements are nearly imperceptible to the observer. But in close range production shots, the motion of the tongue adds realism to facial performance.

Bone-based approaches have been used as the primary technique to achieve different tongue positions. For instance, Graft et al. [Graft et al. 2002] control a realistic character's tongue in Maya by positioning the root joint in the back of the tongue and the tip joint in the front. The bones are attached to a curve in an inverse kinematics (IK) spline system and then the curve's vertices are attached to two clusters that act as gizmos to control the tongue.

Instead of using IK, Bibliowicz [Bibliowicz 2005] sets a forward kinematics (FK) bone hierarchy inside the tongue of a realistic character in Maya. It divides the tongue into four geometric portions seen in Figure 2.39. Each bone can be rotated to deform the geometric portion of the tongue skinned to the bone being rotated. Using only FK can somewhat limit the speed of control but it can increase precision since each bone can be directly manipulated.

Figure 2.39: Approach to rig a human tongue by [Bibliowicz 2005]; left: tongue bone hierarchy located along the center of the tongue; middle and right: respectively a tongue curl behavior and a tongue outwards behavior.

Holly [Holly 2006] also builds an FK bone hierarchy to control the tongue of a stylized character in Maya and then adds bones straying laterally to the tongue from the main FK bone hierarchy (see Figure 2.40). The main hierarchy curls the tongue upwards or downwards and the extra side-bones allow to curl the sides of the tongue so it forms a U-shape.

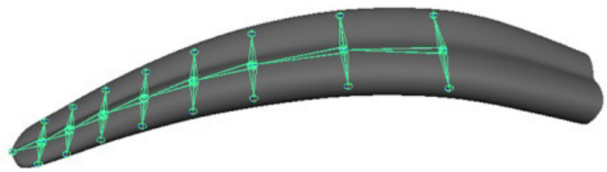


Figure 2.40: Approach by [Holly 2006] to rig a stylized tongue using a main bone hierarchy along the length of the model with lateral bones straying to the sides from the main bones.

Vasconcelos [Vasconcelos 2011] builds only two tongue bones for a realistic character in Blender. The root bone is closest to the character's throat and the tip bone nearer to the lips. The tip joint of the root bone is constrained to stretch to the root joint of the tip bone, allowing the tongue to be controlled only by translating and rotating the tip bone.

2.7 Facial Rigging for Key Frame Animation: Web Approaches

To further illustrate the diversity of approaches available in facial rigging for key frame animation, the following section presents a collection of approaches available in the web. An extensive search for this kind of approaches was conducted in the YouTube and Vimeo websites with relevant content identified from as early as 2007 until present time. These approaches are considered relevant because artists usually tend to show off their greatest quality portfolio in these websites to have a better chance to get a job in the industry, or to be recognized for their work and get feedback.

Tables 2.2 and 2.3, respectively available in Sections 2.7.1 and 2.7.2, list a selection of 48 facial rig approaches, each developed by a different artist identified in the websites YouTube and Vimeo. These approaches are focused on showing off the features of facial rig control interfaces rather than the features of facial rig mechanical setups. The reason is that the artists usually upload facial rig UIs instead of rig mechanics because the rig mechanics are a more time-consuming process to illustrate, as seen previously in Section 2.6.

The first 24 approaches include human-like characters, which more closely resemble humans. The other 24 approaches include anthropomorphic characters, which more closely resemble creature and cartoon-like character concepts.

2.7.1 Human

Table 2.2 lists the 24 human-like facial rigging approaches since 2007 consulted in YouTube and Vimeo. The column on the left identifies the year the approach was upload online, the next column provides an ID number to identify the approach, followed in the next column by the name of the digital artist who developed the approach. The remaining columns identify the interaction modes (window, camera or viewport) that the artist has provided in the user control interface of the character facial rig.

Year	ID	Artist	Interaction Modes Provided		
			Window	Camera	Viewport
2007	1	Denislav Georgiev	•	-	•
	2	Pumpkin Lim	-	•	-
2008	3	Ragh Avendra	-	•	-
2009	4	Alan Camilo	•	•	•
	5	Anson Au	•	•	-
	6	Chi Hung Cheung	-	•	-
	7	Matt Schiller	•	•	•
	8	Max Hammond	-	•	•
2010	9	Nathan Vegdahl	•	-	•
	10	Nox Labs	-	-	•
2011	11	Dan Neuffer	-	•	-
	12	David Marmor	-	•	-
	13	Derek Best	-	-	•
	14	Fangge Chen	-	•	•
	15	Gustavo Rosa	-	•	-
	16	Kenan Eren	-	-	•
	17	Kyle Hewitt	-	•	-

	18	Richard Vaucher	-	●	●
2012	19	Danilo Montesarchio	-	●	●
	20	Jaime Torres	-	●	-
2013	21	Bharani Kumar	-	●	-
	22	Colin Nebula	-	●	●
	23	John Lally	-	●	●
	24	Linil Komban	-	-	●

Table 2.2: Listing of human-like facial rigging approaches from 2007 until today as consulted in YouTube and Vimeo [Georgiev 2007, Lim 2007, Avendra 2008, Camilo 2009, Au 2009, Cheung 2009, Schiller 2009, Hammond 2009, Vegdahl 2010, Labs 2010, Neuffer 2011, Marmor 2011, Best 2011, Chen 2011, Rosa 2011, Eren 2011, Hewitt 2011, Vaucher 2011, Montesarchio 2012, Torres 2012, Kumar 2013, Nebula 2013, Lally 2013, Komban 2013].

From the 24 artists identified in Table 2.2, the artists Alan Camilo and Matt Schiller have implemented the facial rig controls in the three interaction modes. Figure 2.41 illustrates the 24 approaches built by these artists and brief descriptions are provided following on how the artists planned the facial rig controls based on the analysis of their upload videos.

The artist Denislav Georgiev [Georgiev 2007] provided access to window and viewport controls, respectively located in a window interface of the 3D tool and in the viewport on top and side-by-side to the facial model. Pumpkin Lim [Lim 2007] provided access to a single camera view which comprises the facial model with osipas and vertical sliders located over and side-by-side to the facial model. Ragh Avendra [Avendra 2008] used a fairly large number of horizontal sliders arranged in a camera view using the Facial Animation Toolset [Animation Institute 2014]. The controls provided by the artist Alan Camilo [Camilo 2009] are likely to be more complete since they are distributed in the three interaction modes, with the window mode including mainly buttons and the camera mode including mainly osipas. Anson Au [Au 2009] relied on window and camera controls with osipas and vertical sliders.

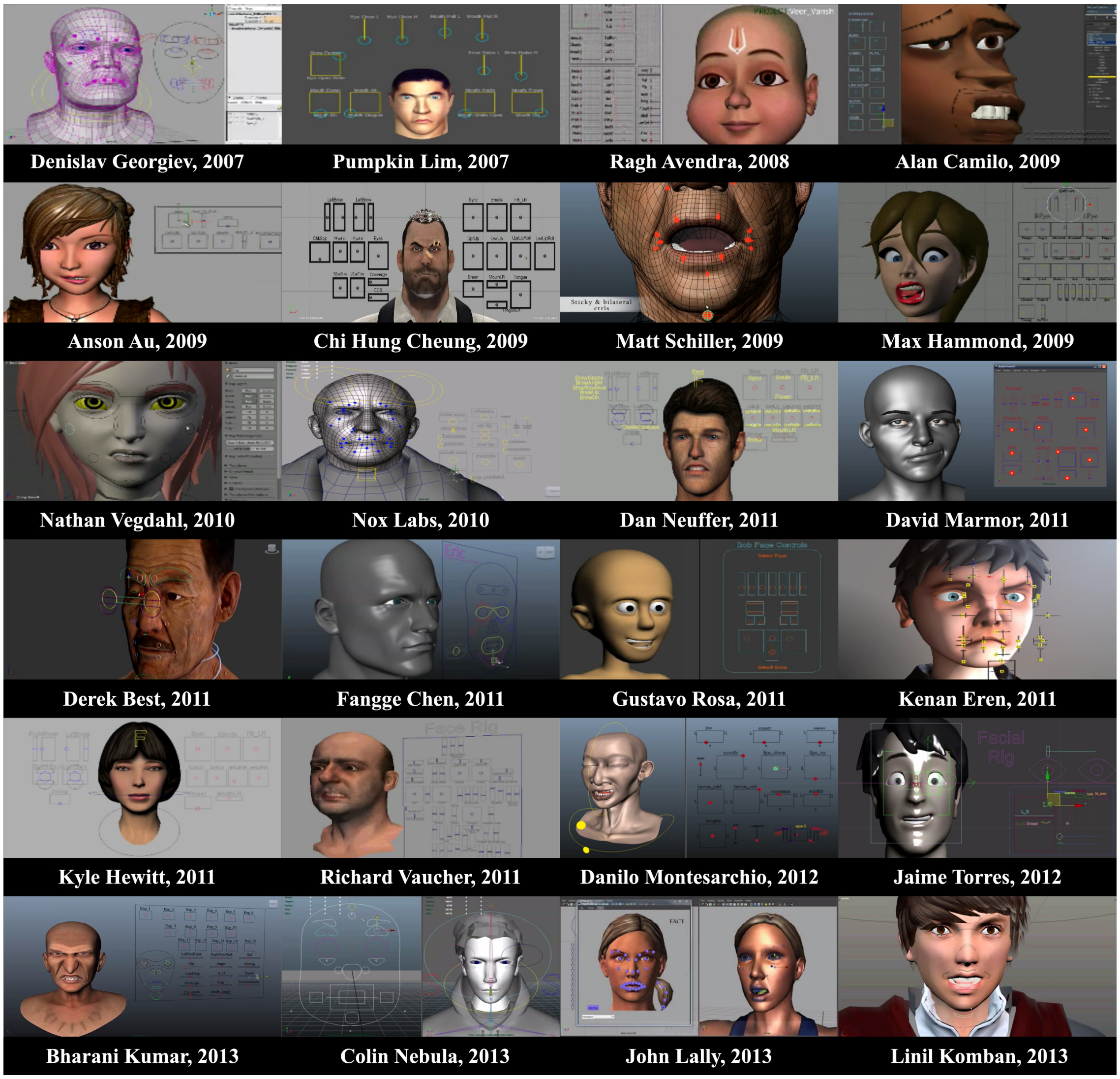


Figure 2.41: Screenshots of the 24 human-like approaches built in-between 2007 and today by [Georgiev 2007, Lim 2007, Avendra 2008, Camilo 2009, Au 2009, Cheung 2009, Schiller 2009, Hammond 2009, Vegdahl 2010, Labs 2010, Neuffer 2011, Marmor 2011, Best 2011, Chen 2011, Rosa 2011, Eren 2011, Hewitt 2011, Vaucher 2011, Montesarchio 2012, Torres 2012, Kumar 2013, Nebula 2013, Lally 2013, Komban 2013].

The artist Chi Hung Cheung [Cheung 2009] relied solely on camera controls distributed in the surrounding spatial area of the facial model. Matt Schiller [Schiller 2009] provided controls with a window, camera and viewport-based interface. In the screenshot emphasis is given to the lips controls located in the viewport over the surface of the facial model. Max Hammond

[Hammond 2009] provided numerous but also detailed camera-based controls accompanied with viewport controls. Nathan Vegdahl [Vegdahl 2010] used window and viewport controls, respectively based on buttons and sliders and also on splines distributed in the 3D space. Nox Labs [Labs 2010] provided an interface based on viewport controls and osipas displayed in the viewport rather than being separated or included in a camera view. The artist Dan Neuffer [Neuffer 2011] built camera controls distributed in the surrounding spatial area of the facial model. David Marmor [Marmor 2011] created a separated interface window with osipas and horizontal and vertical sliders as a camera-based interaction. Derek Best [Best 2011] configured viewport controls layered over the facial model and in its surroundings. Fangge Chen [Chen 2011] provided a bi-dimensional interface located side-by-side to the facial model to be used as a camera interface together with a viewport manipulation of the controls. Gustavo Rosa [Rosa 2011] made more use of a straightforward camera-based interaction. Kenan Eren [Eren 2011] provided viewport controls placed in front of the model following the location of the model's facial regions. Kyle Hewitt [Hewitt 2011] created camera controls distributed around the facial model in a camera view. Richard Vaucher [Vaucher 2011] provided bi-dimensional controls for detailed control which are also coupled with viewport controls located on the surface of the facial model. The artist Danilo Montesarchio [Montesarchio 2012] combined camera-based controls with viewport splines. Jaime Torres [Torres 2012] made use of large camera controls located side-by-side to the facial model for a more direct perception and manipulation. Bharani Kumar [Kumar 2013] focused in providing a schematic camera-based interface. Colin Nebula [Nebula 2013] provided access to camera and viewport controls, respectively located side-by-side and on top of the 3D model. John Lally [Lally 2013] also provided camera and viewport controls with the camera controls being available inside the viewport and in an exterior window. Lastly, Linil Komban [Komban 2013] focused on providing viewport gizmos.

2.7.2 Anthropomorphic

Table 2.3 lists the 24 anthropomorphic facial rigging approaches since 2007 consulted in YouTube and Vimeo that more closely resemble creature and cartoon-like character concepts.

Year	ID	Artist	Interaction Modes Provided		
			Window	Camera	Viewport
2007	1	Huseyin Ozturk	•	-	•
2009	2	Alessio Nitti	•	•	•
	3	Mario Aquaro	•	•	•
	4	Kevin Zheng	•	•	•
	5	David Marte	•	-	•
	6	Chris Pagoria	-	•	•
2010	7	Ahmad Sawalha	-	•	•
	8	Andrew Klein	-	•	•
	9	Felipe Nogueira	-	-	•
	10	Iker Santander	•	•	•
	11	Raul Abonitalla	-	•	•
	12	Thomas Mahler	-	•	-
	13	Tommaso Sanguigni	-	•	•
2011	14	Andy van Straten	-	-	•
	15	Josh Burton	-	-	•
2012	16	David Brooks	•	•	•
	17	Jonathan Soto	-	-	•
	18	Lab Trazos	-	•	-
	19	Pablo Sepulveda	-	-	•
	20	Sebastien Sterling	-	•	•
2013	21	Chris Dike	•	-	•
	22	Evelyn Seng	-	-	•
	23	Gary Suissa	-	•	•
	24	Paul Lombard	-	-	•

Table 2.3: Listing of anthropomorphic facial rigging approaches from 2007 until today as

consulted in YouTube and Vimeo [Ozturk 2007, Nitti 2009, Aquaro 2009, Zheng 2009, Marte 2009, Pagoria 2009, Sawalha 2010, Klein 2010, Nogueira 2010, Santander 2010, Abonitalla 2010, Mahler 2010, Sanguigni 2010, Straten 2011, Burton 2011, Brooks 2012, Soto 2012, Trazos 2012, Sepulveda 2012, Sterling 2012, Dike 2013, Seng 2013, Suissa 2013, Lombard 2013].

From the 24 artists identified in Table 2.3, the artists Alessio Nitti, Mario Aquaro, Kevin Zheng, Iker Santander and David Brooks have implemented the facial rig controls in the three interaction modes. Figure 2.42 illustrates the 24 approaches built by these artists. Brief descriptions are provided following on how the artists planned the facial rig controls based on the analysis of their upload videos.

The artist Huseyin Ozturk [Ozturk 2007] combined camera and window controls to manipulate a baby's face, also taking care to apply a particular schematic design to the camera interface. Alessio Nitti [Nitti 2009] combined viewport, camera and window controls, with the camera controls located in a separate window of the 3D tool interface. Mario Aquaro [Aquaro 2009] also made use of the three different interaction modes with a primary use of viewport controls coupled with camera controls layered in a separate window in the 3D tool interface. Kevin Zheng [Zheng 2009] also provided various controls distributed in the three different interaction modes, with a particular emphasis on the camera mode which can be moved around by the animator due to being located inside a window interface. David Marte [Marte 2009] built both window and viewport controls whereas Chris Pagoria [Pagoria 2009] focused more on building camera controls located side-by-side to the facial model with a viewport interaction. Ahmad Sawalha [Sawalha 2010] relied more on a limited number of fundamental camera controls coupled with viewport splines. Andrew Klein [Klein 2010] and Felipe Nogueira [Nogueira 2010] did a combination of camera and also of viewport controls, with [Klein 2010] having used more the camera controls and [Nogueira 2010] having used

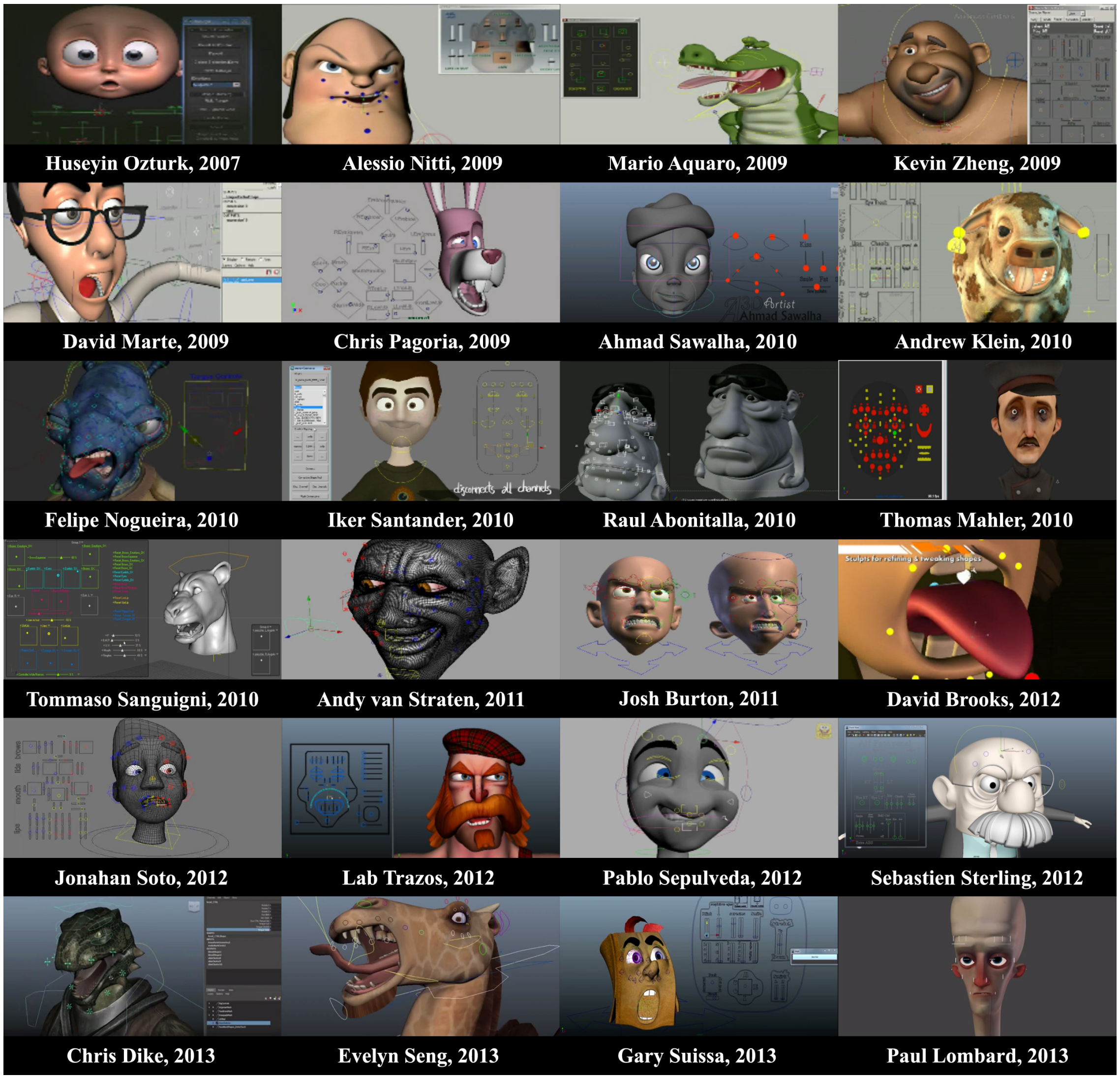


Figure 2.42: Screenshots of the 24 anthropomorphic approaches built in-between 2007 and 2013 by [Ozturk 2007, Nitti 2009, Aquaro 2009, Zheng 2009, Marte 2009, Pagoria 2009, Sawalha 2010, Klein 2010, Nogueira 2010, Santander 2010, Abonitalla 2010, Mahler 2010, Sanguigni 2010, Straten 2011, Burton 2011, Brooks 2012, Soto 2012, Trazos 2012, Sepulveda 2012, Sterling 2012, Dike 2013, Seng 2013, Suissa 2013, Lombard 2013].

more the viewport controls. Iker Santander [Santander 2010] provided window, camera and viewport controls, with the window controls being highly based on buttons. Raul Abonitalla [Abonitalla 2010] created a large number of controls in the surface of the facial model for accurate viewport interaction and then made use of a secondary viewport window to only display the effect of the controls on the facial model. Thomas Mahler [Mahler 2010] and Tommaso Sanguigni [Sanguigni 2010] resorted more to detailed camera-based interfaces, with the latter also using a reduced number of viewport controls. Andy van Straten [Straten 2011] made use of box-shaped gizmos as viewport controls distributed over the surface of the facial model. The artist Josh Burton [Burton 2011] relied more on viewport controls located on the surface and in the neighboring space of the facial model. David Brooks [Brooks 2012] provided a detailed control via window, camera and viewport controls which involved fully colored gizmos, likely to ease their selection and display in 3D space. Jonathan Soto [Soto 2012] provided a balanced mix between camera and viewport controls. Lab Trazos [Trazos 2012] made an extensive use of a camera-based interface coupled with another viewport window to display the effects on the facial model. Pablo Sepulveda [Sepulveda 2012] relied more on groups of circle-shaped viewport controls. Sebastien Sterling [Sterling 2012] relied heavily on a camera-based interface movable inside a window and coupled with viewport controls. Chris Dike [Dike 2013] resorted to using window sliders and viewport gizmos. Evelyn Seng [Seng 2013] built viewport controls based in gizmos distributed along the skin of the model. Gary Suissa [Suissa 2013] built both camera and viewport controls with a stronger use of the detailed camera-based interface. Lastly, Paul Lombard [Lombard 2013] focused on box-shaped viewport gizmos located on top of the facial model skin.

2.8 Discussion and Open Issues

As seen throughout the state of the art, facial rigging for key frame animation is a topic very much alive, which continues to be used extensively by 3D digital artists. But the process is also time-consuming and most artists tend to develop approaches which are feasible rather than ideal or desirably more generic and flexible. The main reason that explains this tendency is the fact that the rigger needs to overcome conceptual, technical and artistic challenges in a daily basis, as described previously in Section 2.3. Hence, independently of the type of rig being created for off-line or real-time systems, specific conditions apply which restrict the creation of the mechanical and control structure of the facial rig. These conditions are (i) the diversity of faces, (ii) the inconsistency of the facial movements, (iii) the lack of a rigging standard and (iv) the inherent complexity of the rig. These conditions were first identified in [Orvalho et al. 2012] (see Appendix A-2012-3), and are described in more detail following. As a result of the description of each of these conditions, the research direction of this thesis is then briefly addressed as a conclusion for this chapter and as an introduction to the results presented in the following chapters.

Diversity of Faces

Due to its large morphological and behavioral features, the illustration of a realistic human face in computer graphics is very complex. A human face is unique due to its variations in size and proportions of the bones and muscles [Ekman and Friesen 1975], which characterize each individual. In addition, the human face has also been extended in characters to include creature and cartoon-like faces. These extra faces do not necessarily follow the human structure and behavioral limits [Ritchie et al. 2005]. The ability to create either human, creature or cartoon characters expands the diversity of faces built and causes the rigging process to be difficult for artists to deal with.

Inconsistency of Facial Movements

Humans are naturally able to identify and classify unnatural behavior because of their familiarity and sensitivity to the face experienced on a daily basis [Mori 1970]. Moreover, the representation of facial behaviors can be more important than the display of the visual look in terms of human-like and empathy [Maddock et al. 2005]. The difficulty lies in the fact that the various facial styles generate a great diversity of behaviors which are not always produced exactly in the same way or as they would be expected. It is not just a matter of many behaviors being possible to simulate, it is also the question of how much their repetition can be divergent. A common solution is to combine FACS [Ekman and Friesen 1978] with the classic principles of animation when rigging a character [Harkins 2005, Osipa 2007] to produce realistic behaviors.

Lack of a Standard

As demonstrated in the state of the art, there is no character facial rigging standard. The reason is the fact that there are many different ways to build the mechanics of a facial rig as well as provide the controls to manipulate the rig, as seen in Sections 2.6 and 2.7 through the several approaches built by many artists worldwide. The great disparity and lack of consensus in facial rigging is because the artists seem to enjoy customizing their rigs to their liking instead of following predefined rules. In general, artists do not follow a formal criteria or method when creating a rig, that is an intuitive and creative process which depends on the preference of each artist, therefore all rigs tend to end up differently. A facial rigging standard becomes impractical also due to (i) the high number of software packages and in-house tools being used for character rigging which are based on their own specifications and (ii) the different digital media that the character is meant for, since there are a wide range of applications defined with different rig architectures (e.g. films, videogames, virtual reality, visual effects, etc.). As a result, it is often necessary during character production to create a variety of rig templates to support different behaviors. For instance, in the film The Incredibles (Pixar, 2004) characters with super powers had a rig that allowed squash-and-stretch movements and real world characters had rigs which were limited to normal movements that obeyed the laws of physics.

Inherent Complexity of the Rig

A rig can be more or less complex and accessible to setup depending on (i) the rigger's experience, (ii) the animator's creative feedback and (iii) the budget and time constraints in a film or videogame. In addition, the complexity of the facial rigging techniques is increasing [Maestri 1996, Osipa 2003, Orvalho 2007, O'Neill 2008, Dutreve et al. 2010, Miller et al. 2010] due to (i) the inherent complexity of the human facial behaviors [Ekman and Friesen 1975, Ekman and Friesen 1978, Faigin 1990, Fleming and Dobbs 1998], (ii) the ongoing evolution of computer graphics in terms of deformation accuracy and interactive animation control, (iii) the increasing need for characters to perform more complex facial animation for blockbuster feature films and videogames and (iv) the tendency that artists have to add a large number of controls to better handle all the subtleties of the different facial regions.

Research Direction

The conditions described previously, which were first identified in a generic survey on facial rigging [Orvalho et al. 2012] (see Appendix A-2012-3), together with the disparate literary and web approaches, which are described throughout this chapter as a survey focused on facial rigging for key frame animation (see Sections 2.6 and 2.7), have contributed in the last 15 to 20 years to restrict the creation of a facial rig and reveal that character facial rigging and animation is a fragmented field of expertise. Each 3D digital artist performs the setup of the rig mechanics and controls for a character's face mostly based on their own experience, without necessarily following any predefined method or guidelines to ease the process. As a result, there is no general consensus on how to build a facial rig, which makes it harder for both novice and experienced artists to rig their characters. Aware of the lack of a facial rigging consensus, the author presents in the next chapter a proof of concept that explains the optimized facial rig approaches devised in this research as a contribution to improve this field.

The facial rig approaches are then presented in Chapter 4 as (i) guideline proposals to place and configure the mechanical components in a facial rig and in Chapter 5 as (ii) a multimodal rig interface approach with multiple types of control to help provide freedom of manipulation according to the user preference.

Chapter 3

Proof of Concept

This chapter begins by stating the problem of the research and the solutions devised. It then presents two framework results that contributed to define the challenges and direction of the research. After that, the construction of a human subject facial model is briefly addressed, which served as the base model to implement the easy facial rig mechanical and control approaches described respectively in Chapters 4 and 5. In this chapter, the framework results presented in Section 3.3.1 are published in A Demo of a Dynamic Facial UI for Digital Artists [Bastos et al. 2011] (see Appendix A-2011) and the results presented in Section 3.3.2 are published in A Demo of a Facial UI Design Approach for Digital Artists [Bastos et al. 2012a] (see Appendix A-2012-4).

3.1 Problem Statement

The problem raised in this research is not only the fact that facial rigging is laborious and time-consuming [Orvalho et al. 2012, Bastos et al. 2012b], but that it is also a frustrating and startling process for digital artists. Facial rigging and animation is a frustrating job for artists because it involves (i) simulating the many different expressions in a character's face using rig mechanics that deform the facial model and (ii) deciding how to present and create a user control interface using rig controls that provide to the animator a precise manipulation of the different expressions independently of circumstances such as the angle of view of the face. Few artists become specialized in facial rigging and animation in the entertainment industry because this field is vast and disparate. Building a facial rig requires time and effort to acquire the required conceptual, technical as well as artistic skills, because it is an iterative process between the rigging and animation departments, in which the rigger needs to plan and build efficient technical solutions to better suite animation. As seen in Chapter 2, each artist does facial rigging differently, often more according to what is feasible to develop for the rig to work instead of more according to what would be ideal to develop in order to achieve an optimized rig with an enhanced user control interaction. For this reason, this field of expertise became fragmented, or deprived of consensual guidelines, making it harder to carry out efficiently. Therefore, riggers can benefit from improved solutions to ease their work and the work of the animators, who will use the rigs built by riggers.

Initially, the goal of this research was to define a standard for the facial rigging process. After studying the state of the art, it became evident that facial rigging is a complex problem for which devising a standard was impractical. A standard can be realized as a rule, or a set of rules, which are established as valid and necessary procedures to achieve a certain task. In order to set such rules it is necessary to conduct throughout research done in the long-term and supported by large data of information, to corroborate the rules in the standard. Searching for a standard in the field of facial rigging would mean conducting dawning efforts with professional artists worldwide involved in this field. But, many artists are prevented by copyright from exposing their work and other artists are not willing to respond to research requests. In addition, the impracticability of a facial rigging standard increases when dealing with anthropomorphic faces, due to (i) their diversity in shape, (ii) their inconsistency of movements, (iii) the disparity of rig approaches built by many digital artists [Orvalho et al. 2012] and (iv) the fact that the artists prefer to be able to customize the rig controls in order to make them more efficient and pleasant to manipulate, without being restricted to any predefined rules. Still, it is important to consider the aspect that the facial morphologies and behaviors of anthropomorphic characters need to remain desirably diverse and unpredictable in order to be appealing and emphatic for an audience. And for a character to be appealing for an audience, it is always expected that the character retains human features, so that the character can be recognized by the audience as an expressive agent, as is seen in the cartoon character created by [Harkins 2005].

3.2 Easy Facial Rigging and Animation Approaches

Given the impracticability of developing a facial rigging standard, the goal in this research became the study of easy facial rigging and animation approaches for a human subject facial model as a basis for other facial styles, to help bring a consensus to this field based on the fact that every face has a minimum set of human features. Two major challenges dependent of each other were considered to achieve the proposed goal. The first challenge is to ease the rigging of the underlying mechanics that deform a facial model, the second is to ease the rigging and animation of the user interface controls. A facial rig is not complete without including both mechanics and controls. The rig mechanics and controls are developed solely by the rigger, but to build the controls extra feedback from the animators is required in order to provide an interface for animation that facilitates the manipulation of the character's face. To provide the optimized facial rig approaches presented in this thesis, the starting point of research is the rig mechanical process, because the user interface controls need to trigger the already existent internal deformation mechanisms of a facial model.

In terms of the rig mechanics, a definition proposal is addressed on how to build the rig mechanical components for a human subject facial model. The rig mechanical definition is provided as suggested guidelines for artists to realize how to efficiently place and configure the mechanical components of a facial rig. The suggested guidelines are based on a set of rig construction and deformation conditions identified for the different human facial regions, in order to achieve a complete mechanical structure capable of deforming a human facial model accurately and as a basis for other facial styles. The facial rig mechanical optimization is presented in Chapter 4 as guidelines which were implemented in Blender for a human subject facial model, first in a per facial region basis in Section 4.1 and lastly for two distinct behaviors of the lips region, the zipper and sticky lips visual effects, in Section 4.2, to help provide a consensus to their rigging process, given that these are complex behaviors rarely simulated in the state of the art. In fact, the only widely known reference that describes these effects is Osipa [Osipa 2010], who provided objective procedures on how to deploy the sticky lips behavior. Still, there is no literary reference or even web reference, to the best of the author's knowledge, that describes how to setup a full rig for the lips region that supports both the zipper and sticky lips effects together and with control over the jaw and upper and lower lip regions. The optimization of the rig approaches for the lips effects was validated in an international user experiment conducted with five college professors who are also expert professional character animators.

In terms of the rig controls, a definition proposal is addressed based on a multimodal rig interface approach that provides multiple types of control, which are synchronized and scalable by the user: the window, camera and viewport interaction modes. The expectation of the multimodal interface approach relies on providing freedom of choice to the animator, based on the fact that different interaction modes are made available and that each mode can be more suitable in different circumstances during animation. An example is the control of the jaw thrust behavior, which is more adequate to manipulate from a profile view of the facial model rather than from a front view, because it involves a frontwards and backwards motion, described in FACS [Ekman and Friesen 1978] as the Action Descriptor 29, or AD 29. The availability of different interaction modes enables a more personalized control of the user over the different facial regions of the human subject facial model. The multimodal approach lets artists seamlessly choose which type of control to use for a given animation task by selecting amongst multiple control modes according to their preference and with real-time feedback of the transformation of the corresponding manipulators in the different interaction modes. Multiple and seamless controls are intended to help decrease user frustration based on the fact that more control options prevent the user from being restricted to a limited number of controls, which is usually the case in common rigs. The multimodal rig interface approach is presented in Chapter 5 as an interactive rig control system implemented in Blender for the human subject facial model. This system was first validated in an initial user experiment with 15 basic, skilled and expert digital artists, which is described in Section 5.1. The feedback obtained in the initial experiment motivated the author to build an improved version of the system, which was validated in Section 5.2 in a final international user experiment carried out online with 20 expert digital artists who are long-term professionals in the field.

The facial rig approaches are implemented in Blender for a human subject facial model, with the human subject being the author's face. The construction of the 3D facial model is described in detail in this chapter in Section 3.4. The implementation of the rig approaches for the human subject facial model is done in the popular open-source 3D tool Blender to (i) cause a significant impact in their growing community of artists, (ii) due to Blender users being highly predisposed to validation, as seen in a preliminary result [Bastos et al. 2012a] described ahead in this chapter in Section 3.3.2 and (iii) to leverage the use of Blender in the industry as a qualified platform with an extended workflow. The rig approaches are also oriented towards the key frame animation technique [Burtnik and Wein 1970] because (i) it is likely to be the technique that artists most use, as it continues to be popular in Computer Graphics (CG) animation [Orvalho et al. 2012] (ii) it is the most accessible and affordable technique since it does not require expensive equipment, (iii) results from other animation techniques need to be adjusted using key frame (e.g. after motion capture sessions), (iv) it allows digital artists to have a more precise manipulation of the animation using frame by frame control [Orvalho et al. 2012] and (v) the fact that key frame is currently at its peak of growth [Orvalho et al. 2012], thus more prone to welcoming rig optimization approaches.

The implementation being carried out in Blender does not mean that the rig mechanical and control approaches are enclosed to it. The approaches are likely to be applicable to other major 3D tools, because they are based in techniques which are found in the major 3D tools. As seen in the state of the art, artists use different tools to rig faces, with Maya leading the vigorous industry of key frame facial rigging and animation. But, in the last decade, Maya's leadership is followed by an exponential increase in the use of the open-source tool Blender, supported by a growing community of users worldwide. Blender being open-source “attracted hundreds of active developers and volunteers worldwide” [Mullen 2007], who propel the continuity of Blender as state of the art software in computer graphics, as is reported in the download statistics of Blender [Blender Nation 2009] and in the Blender website consultation statistics available in [Blender Foundation 2011]. In the last decade, Blender relied in the open-source community [Blender Foundation 2012], an idea which is recently more present in the industry, as Pixar senior scientist Tony DeRose said, “we had a competitive advantage for ten years, but now we get more value by letting everyone contribute” [DeRose 2013].

3.3 Framework Results

After the realization that a facial rigging standard was impractical, two preliminary results were devised and then implemented in Blender to assess the feasibility of the easy facial rigging and animation approaches. The preliminary results revealed fundamental to clear the pathway and helped define the framework of the research. The first preliminary result is a demo of a dynamic rig control user interface approach that extends the usability of the common facial rig UIs by providing a flexible configuration of the panels in a facial UI. The second preliminary result is a demo of a rig control interface design approach which is focused on the human face with the purpose to consolidate the first preliminary result, which was less consistent, due to the diversity of configurations that the user can achieve for the interface design. The first and second preliminary results are detailed following, respectively in Sections 3.3.1 and 3.3.2. For a throughout verification of the preliminary results, please refer to the Blender files entitled *section_3.3.1* and *section_3.3.2*, available in the folder *sections blend files* in the accompanying CD.

3.3.1 A Demo of a Dynamic Facial UI for Digital Artists

The first result is a demo of a dynamic rig control user interface that extends the usability of the common facial rig UIs. This dynamic interface allows an animator to re-dimension and re-position the regular panels found in a camera-based facial UI by (i) dragging the panels freely into new positions in space and (ii) dragging extra manipulators located in the extremities of the panels to re-dimension them. The goal of this preliminary result was to improve the common facial UIs found in the state of the art, which are usually cluttered with controls and have a limited user interaction. The feature of dragging the panels into new positions allows the user to reconfigure the layout design, which can be helpful in different circumstances, such as triggering a specific behavior with a closer zoom to the facial region to achieve more precision of deformation. The feature of re-dimensioning the panels allows the control manipulators to be moved further to achieve extra deformation, which can be relevant for cartoon characters, such as the character face seen in Figure 3.1, which was the subject facial model chosen to implement this preliminary result.

This first demo result allows to deform a facial model using a combination of bones and blendshapes, which were implemented in Blender as a fundamental set of rig mechanics for the cartoon facial model seen in Figure 3.1. The mechanics are triggered by a facial UI based on camera controls and composed of bones with custom shapes associated to them to give the bones their shape. Custom shapes are a technique in Blender that allows to visually represent a bone as any 3D object, allowing to change the appearance of the bone for that of the 3D object, in this case squares or rectangles, as is illustrated in Figure 3.1. Using panels based on bones instead of regular 3D objects keeps the rig clean, because it depends on the transforms of its own internal skeletal components instead of the transforms of exterior 3D objects. The triggering of the rig mechanics occurs via drivers that link each transformation channel of each manipulator seen in the camera controls to its corresponding deformation bone or blendshape. Figure 3.1 illustrates the dynamic facial UI applied to a cartoon character.

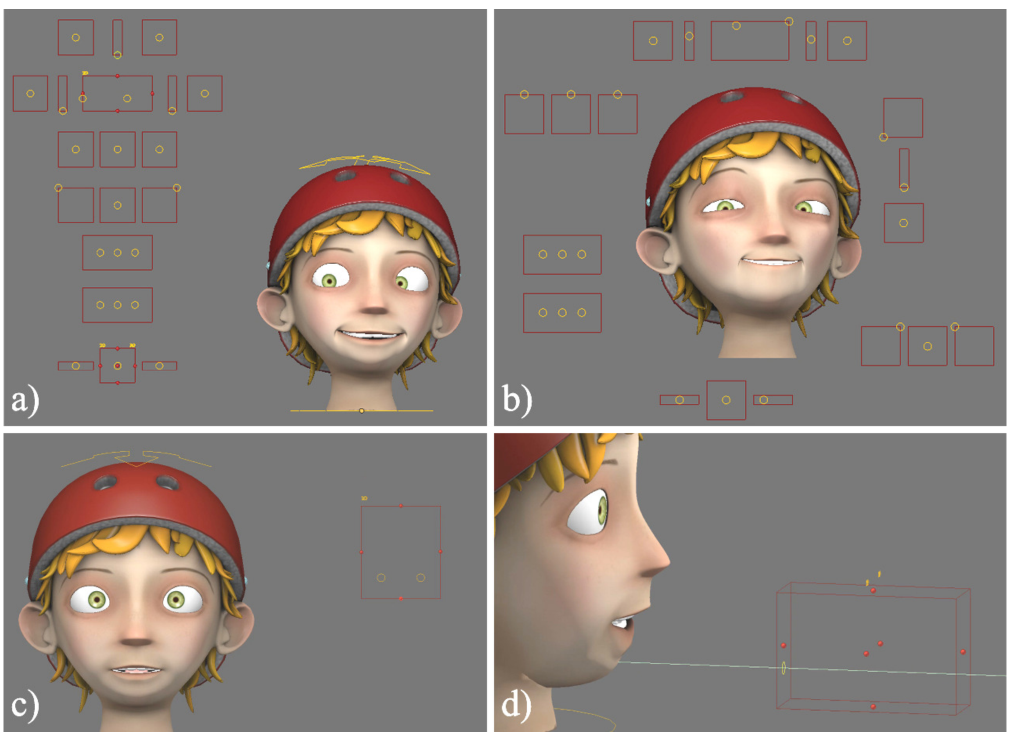


Figure 3.1: Screenshots of the dynamic facial interface; a) the initial static interface with a few manipulators moved; b) a new interface layout created by re-positioning the panels; c) and d) respectively the control panels of the eyeballs and jaw re-dimensioned and re-positioned (facial model is copyright of FaceInMotion).

The dynamic facial UI seen in Figure 3.1 provides a greater freedom of control because it combines a regular static UI seen in a) with a general re-dimensioning and re-positioning of the several control panels seen in b). Specific situations where this approach can be of benefit are seen in c) and d). In c), the control panel of the eyeballs is extended allowing further use of the manipulators. In d), the control panel of the jaw is transformed from a 2D into a 3D panel to facilitate specific jaw behaviors, namely the jaw thrust and sideways motions.

This first demo result improves on a considerable limitation of the common facial UIs, the fact that these UIs are usually setup to occupy fixed positions in 3D space. Also, the digital artist responsible for building the UI generally places it in a fixed position side-by-side to the face of the character. The dynamic facial UI presented provides the ability for the animator to easily relocate and also re-dimension the panels in a facial UI. The animator can customize the limits and rearrange the layout of the panels in the UI to gain more control over the animation of the several facial regions. It is possible to re-position the entire UI or a specific panel, transforming the panel as desired to fit the facial region that it controls.

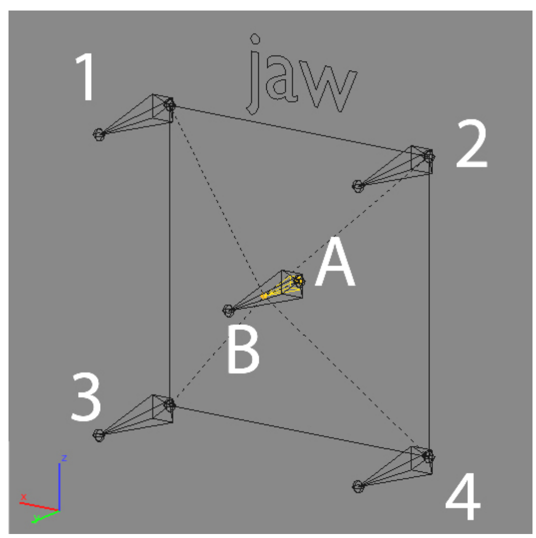


Figure 3.2: A six bone setup to dynamically control the jaw panel and manipulator handle.

The ability to relocate and rescale the panels in the interface is achieved using the bone rig mechanics illustrated in Figure 3.2. Bone A, the highlighted smaller bone in the center, is used to drag the panel, in this case relative to the jaw. Bone B is the controller of the jaw's animation, which is restricted to the borders of the jaw panel. The remaining four bones are placed in the corners of the panel so that each controls its corner. Bone A is set as the parent of bones

1 to 4 and also of bone B, keeping an offset between them as illustrated by the relationship lines in Figure 3.2. Bone A can be moved freely in space but the location of bone B is constrained to prevent it from moving to the front or backwards (Y axis in Blender). In order for the panel to follow along with the bones, either vertex groups or hooks can be used in Blender to associate each vertex of the panel to its corresponding bone. The setup illustrated in Figure 3.2 is finished by linking the translation of the jaw controller bone in the facial UI to the rotation of the deformation bone in the character's jaw, so that the jaw region of the facial model reacts to the motion of the control bone.

Although this was an appealing and promising preliminary result, it continued to clutter the viewport space, a limitation which becomes more evident if more controls are added to the interface. The ability to drag the several panels in space can also contribute for animators to build dispersive layouts, increasing the complexity of the visual space and increasing the chance to lose track of each control. Therefore, it would be hard to extend these features to the entire behavioral set of the face of a 3D character if only a camera-based interaction was used. The limitations found in this early prototype have led to the construction of the multimodal interface presented in Chapter 5, but first it was necessary to provide a more consensual design layout to the control interface, an issue which was addressed in the second preliminary result, described following.

3.3.2 A Demo of a Facial UI Design Approach for Digital Artists

This second demo result describes a rig control interface design focused on the human face as a basis of study to improve the issues found previously in the development of the dynamic facial rig UI. The goal was to develop a straightforward design prototype with a minimum number of control manipulators to prevent cluttering the visual space and still supporting a wide range of different facial behaviors. Also implemented in Blender, this demo result uses the same construction methods of the dynamic interface described in Section 3.3.1, with the exceptions that (i) the facial UI was redesigned as a fixed camera-based interaction applied to a human face and (ii) a pilot study was carried out with three digital artists with professional curriculum in 3D animation, who have validated this approach. Figure 3.3 illustrates the design approach of the facial user interface applied to a human facial model built using the open-source software MakeHuman[3], which was then imported to Blender to build the rig.

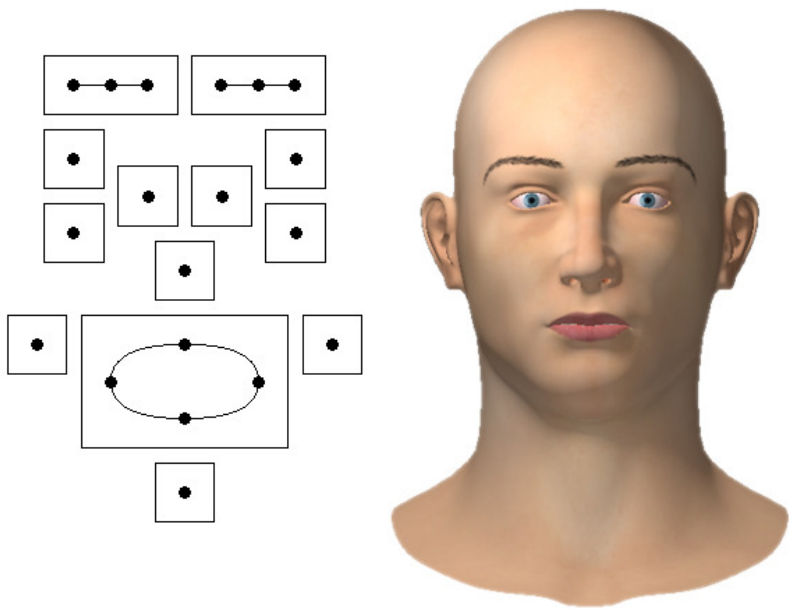


Figure 3.3: The straightforward design approach of the camera-based facial UI.

The user interface approach seen in Figure 3.3 was designed to support the six basic facial expressions of anthropomorphic faces identified by [Ekman and Friesen 1975]: surprise, fear, disgust, anger, happiness and sadness. An anthropomorphic face is realized as any face with the anatomical characteristics of a human face (e.g. number, symmetry and position of facial regions). The goal of the pilot study was to realize if the interface design could average the performance of the animators while posing the six expressions. The design is based in square-shaped widgets devised originally by [Osipa 2003], holding circular handles to produce the

[3] http://www.makehuman.org/

necessary muscular activity for the basic facial expressions. Each handle can move horizontally and vertically to change opposite parameters. The handles control the eyebrows, eyelids, nose, cheeks, lips and jaw, as the minimum facial regions necessary to be moved in order for the six basic facial expressions to be perceived accurately.

In the pilot study the animators were introduced to the UI controls and to the guidelines defined by [Ekman and Friesen 1975] to animate the expressions. Then they were asked to reproduce the six basic facial expressions using the interface without any further guidance. To evaluate the performance of the three animators when using the controls, two parameters were considered: the number of clicks and the time taken to animate the six expressions. The study reveals that there was little variation in the values each animator marked for both parameters, despite their natural difference in terms of creative performance. Table 3.1 shows the results of the animators' interaction. The optimal clicks seen in Table 3.1 were obtained based on the guidelines defined by [Ekman and Friesen 1975] for each of the six basic facial expressions.

		The Six Basic Facial Expressions						TOTAL
		Surp-rise	Fear	Dis-gust	Anger	Happ-iness	Sad-ness	
	Optimal Clicks	7	9	4	8	2	4	34
Animator 1	Clicks	9	9	6	8	2	4	38
	Time (secs.)	27	30	15	15	4	10	101
Animator 2	Clicks	7	13	5	8	2	4	39
	Time (secs.)	17	32	17	24	5	11	106
Animator 3	Clicks	8	9	5	12	2	6	42
	Time (secs.)	31	28	20	57	8	20	164

Table 3.1: Statistics of the animators' interaction with the facial UI seen in Figure 3.3.

Animator 1 did 38 clicks in 101 seconds. Animator 2 did 39 clicks in 106 seconds. Animator 3 did 42 clicks in 164 seconds. The number of clicks each animator did is very similar to each other and close to the optimal number of clicks (34). The time taken was similar for animators 1 and 2. Animator 3 required more time to place the controls thus achieving the most accurate expressions, as seen further ahead in Figure 3.4. Despite Animator 3 required more time, the number of clicks performed did not change significantly when compared with Animators 1 and 2.

Figure 3.4 illustrates, in the top row, the optimal poses built for each facial expression following the guidelines defined by [Ekman and Friesen 1975]. These were registered as the optimal number of clicks seen in Table 3.1. The three animators were given these guidelines and their results are shown in Figure 3.4.

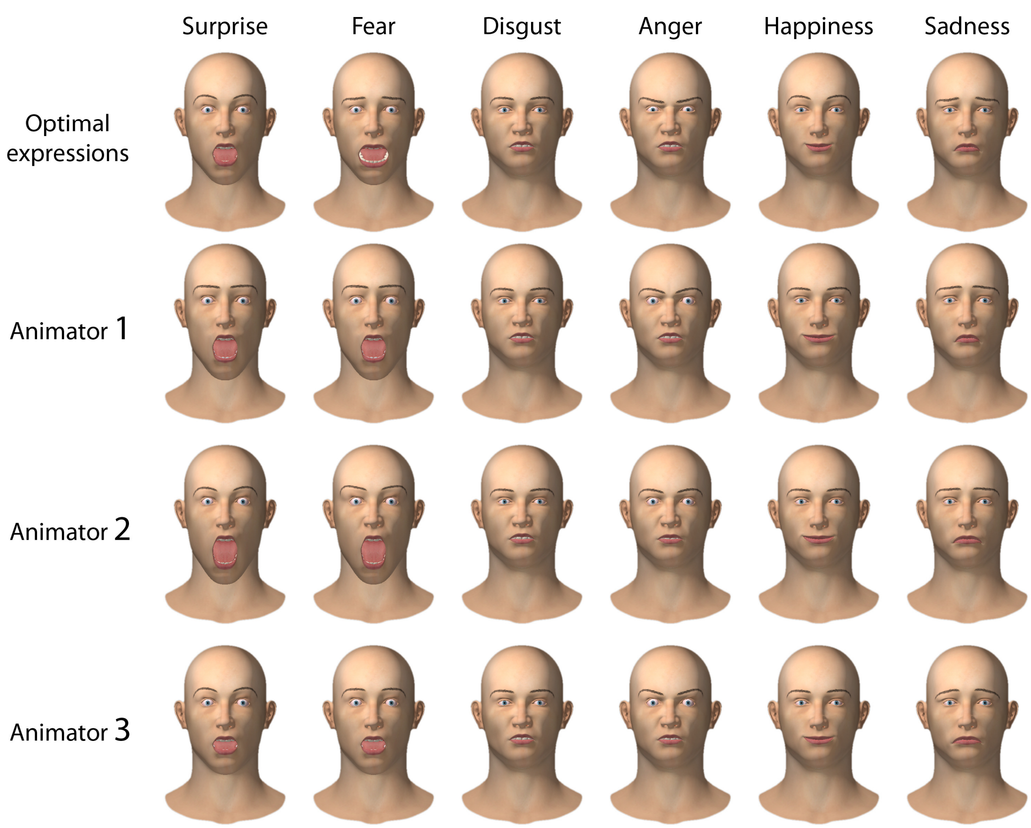


Figure 3.4: Comparison of the animators' performance results obtained using the facial UI design approach seen previously in Figure 3.3.

As seen in Figure 3.4, the animators were able to simulate the six basic facial expressions with increased accuracy. As a conclusion, the suggested facial rig UI design paradigm approximates the number of clicks and the time required to animate the six basic facial expressions defined by [Ekman and Friesen 1975]. Figure 3.4 shows that fear was the only expression that the animators could not simulate as well as its corresponding optimal expression. The reason is that the animators could have considered that the middle portion of the eyebrows can raise more in surprise and appear more straightened in fear [Ekman and Friesen 1975]. This occurrence is not considered a limitation, because the surprise and fear expressions share similarities [Jack et al. 2014] and therefore are prone to have different interpretations from the users.

3.4 Human Face Behavioral Study and Modeling of a Subject's Face

Both the preliminary results described in Sections 3.3.1 and 3.3.2 were important to realize that an accurate 3D facial model was required to serve as a basis to implement the facial rig approaches studied in this research. Even if the rigging process is carefully designed, only a carefully modeled human subject face would allow a correct implementation of the optimized facial rig approaches. An accurate model helps to design the rig because the rig components are built relative to the different target emotions that need to be expressed in each facial region. This section briefly describes the nuances of modeling the subject's face in Blender following the behavioral guidelines in FACS [Ekman and Friesen 1978] for each facial region. FACS was the reference chosen for the study of the facial behavior due to being the most complete and accurate, which includes photographs of individuals performing the different nuances of the muscular activity of the human face. For the reference model the author decided to reproduce his own face, for a number of reasons, namely (i) to avoid copyright issues, (ii) to speed up the process because the subject is always available and (iii) to have more freedom to explore the motion of the different facial regions in order to model them accurately.

A study of the subject's facial behaviors was performed following the guidelines in FACS [Ekman and Friesen 1978] to realize the limitations of the facial behaviors and consider those limitations for the rig setup. The study was carried out in December 3, 2012, in Faculdade de Engenharia da Universidade do Porto (FEUP), with the support of Centro de Eventos of FEUP. The setup is illustrated in Figure 3.5, it involved filming the subject using two cameras reference JVC GY-HM100E, two tripods, two lighting sets, a green background for clear perception of facial motion and a tape-measure to set the height of the tripods and the distances to the subject. One camera filmed the subject from the front and the other filmed the subject from the side to record the subject's profile.

Figure 3.5: Experiment setup in room I-104 in FEUP.

Several poses were simulated per each facial region (e.g. nose wrinkle, cheek raiser and lid compressor, jaw thrust, upper lip raiser, etc.) based on the muscular activity categorized in FACS by [Ekman and Friesen 1978] into Action Units, Action Descriptors and Gross Behavior Codes. This experiment allowed a better self-awareness of how the author's facial deformations occur via the analysis of the front and profile views of the subject in order to construct a high fidelity 3D model. An example of a facial expression simulated during the experiment which involved eyebrows movement is illustrated in Figure 3.6, the outer brow raiser (AU 2).

To build the facial model the facial feature points identified in the MPEG4 specification by [Pandzic and Forchheimer 2002] were used as reference together with high quality photographs of the front and side of the subject's face. These photographs were obtained using rigorous alignment and constant lighting to prevent shadow spots. The photographs were later used to texture the model to improve its realism (see Section 4.1.10).

Other concerns during modeling include the vertex count, edge flow, polygonal size, spacing and density. In overall, the model represents a balance between visual quality and workflow, to guarantee a believable deformation and a manageable smooth skinning of the base model to the facial rig. The edges of the model are built according to the overall shape of the face and to the motion direction of the facial muscles. Polygonal density is higher in areas that suffer more complex deformations (e.g. eyelids and lips) and the polygons in the facial skin are always four-sided (quads) to provide a more consistent and smoother deformation during animation.

Figure 3.6: Front and profile views of the subject performing a facial expression simulated at the time of the capture which involved simulating the outer brow raiser (AU 2).

Sections 3.4.1 to 3.4.9 provide a brief comprehensive insight on how each facial region of the 3D model was built based on the guidelines by [Ekman and Friesen 1978] to achieve a model ready for rigging.

3.4.1 Neck and Head

According to FACS [Ekman and Friesen 1978], the movements of the neck involve tightening it (AU 21 – Neck Tightener), as in wrinkling and bulging the skin of the neck, and also swallowing (GBC 80 – Swallow). The head includes gross behavior codes to shake it back and forth (GBC84 – Head Shake Back and Forth), nod the head up and downwards (GBC85 – Head Nod Up and Down), turning the head sideways on a vertical axis (AUs 51/52 – Head Turn Left or Right), pointing the head up or downwards (AUs 53/54 – Head Up or Down), tilting the head left or right, as in cocking the head to one side or to the other (AUs 55/56 – Head Tilt Left or Right), and moving the head forward or backwards (AUs 57/58 – Head Forward or Back). Based on the information described, an overview of the polygons in the neck and head is seen in Figure 3.7, which is combined with the several facial regions, with each facial region being described following.

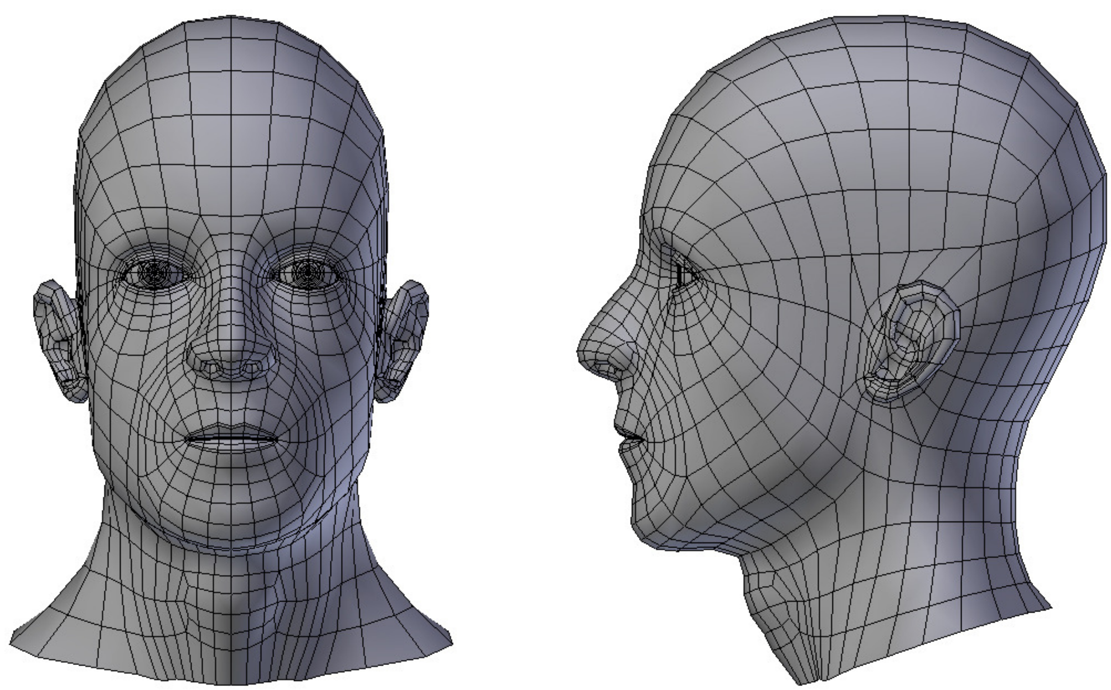


Figure 3.7: Front and profile views of the neck and head of the human subject facial model.

3.4.2 Eyebrows

According to FACS [Ekman and Friesen 1978], the movements of the eyebrows involve its inner, mid and outer regions. The inner region of the eyebrow has upward and horizontal movements that produce oblique shapes (AU 1 – Inner Brow Raiser). The mid eyebrow area moves slightly downwards. The outer portion of the eyebrow moves up to produce an arched shaping (AU 2 – Outer Brow Raiser) and it also lowers the eyebrow in that region (AU 4 – Brow Lowerer).

Based on the information described, the polygons of the eyebrows can be organized with an overall oval shape around the socket of the eyes, as seen in Figure 3.8. The arrows indicate the motion of the inner, mid and outer areas of the eyebrow. Two rows of polygons were built

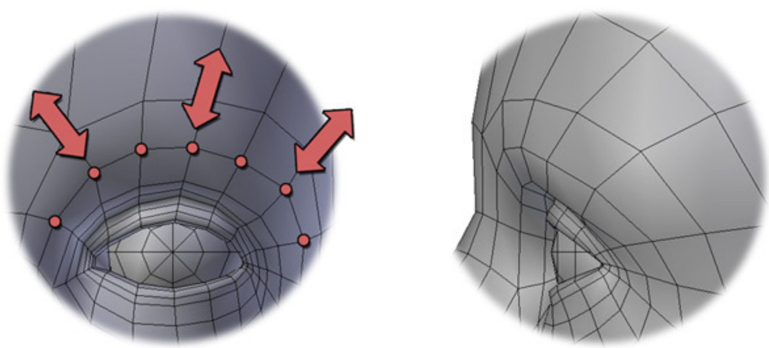


Figure 3.8: Front and profile views of the left eyebrow of the human subject facial model.

that extend to each area of the eyebrow to provide consistency. There is also a total of seven key rows of edges which are evenly spaced and with their respective vertices highlighted in Figure 3.8. The number of edges and vertices flows across the eyes into the eyelids.

3.4.3 Eyelids

FACS [Ekman and Friesen 1978] mentions that the movements of the eyelids involve raising and lowering their upper and lower regions (AU 5 – Upper Lid Raiser) to achieve poses such as closing (AU 43 – Eye Closure), tightening (AU 7 – Lid Teightener), blinking (AU 45 –

Blink) or eye winks (AU 46 – Wink). These require a vertical and concentric arcing motion towards the eyeball with occasional slight shape changes depending on the eye gaze. Hence, the eyelids edges should form concentric rings around the eye socket, as shown in Figure 3.9.

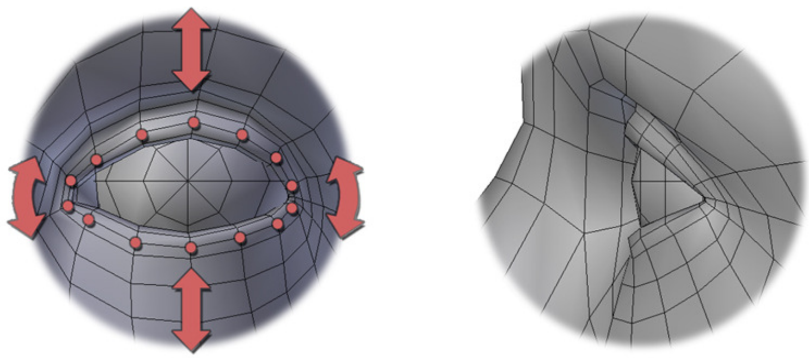


Figure 3.9: Front and profile views of the left eyelids of the human subject facial model.

The arrows in Figure 3.9 indicate the motions of the upper and lower regions of the eyelids. The rows of polygons coming from the eyebrow help define the seven key rows of edges in the top and bottom eyelids. The difference for the eyebrows is that the eyelids edges point even more towards the center of the eyeball. Therefore, shaping the eyelids rims as concentric edge loops ensures their physicality and behaviors and also makes them more comprehensive to smooth skin for animation.

3.4.4 Eyeballs

The eyeballs are spherical-shaped and have a rotation-based movement. In FACS [Ekman and Friesen 1978], the movements of the eyeballs involve turning the eyes left or right (AUs 61/62 – Eyes Turn Left or Right) and rolling up or down (AU 63/64 – Eyes Up or Down). The former also include the asymmetrical wall-eye (AU 65 – Wall eye), which occurs when one eyeball alone is turned, and the cross-eye phenomena (AU 66 – Cross-eye), which occurs when the eyeballs cross. Figure 3.10 shows the eyeball model built.

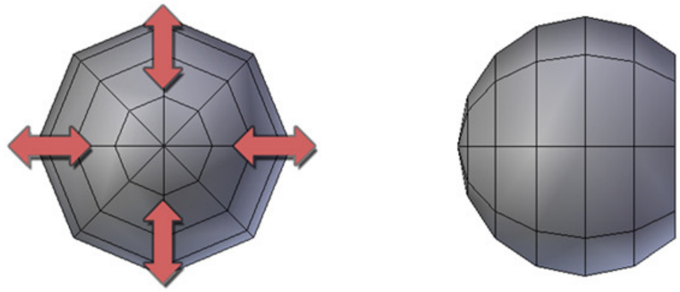


Figure 3.10: Front and profile views of one of the eyeballs of the human subject facial model.

The physicality of the eyeballs and their behaviors suggest the use of a primitive sphere. Eight rows of meridians are used to keep the model manageable and perceptible, since Blender's subdivision surface will add on top of the base quality of the model. Two rows of polygons are deleted in the back of the eyeball, as these will not be seen, even in extreme human-like eyeball rotations. The pupil dilation, which can be realized as the outwards and inwards scale of vertices found in the center of the pupil, also suggests the use of circular loops in this region. If necessary the sclera can also be simulated using overlapping spherical-shaped meshes.

3.4.5 Nose

According to FACS [Ekman and Friesen 1978], the movements of the nose include sniffing (GBC 40 – Sniff), snarling (AU 9 – Nose Wrinkler) and dilating (AU 38 – Nostril Dilator) or compressing the nostrils (AU 39 – Nostril Compressor). Figure 3.11 shows the edge layout and motion directions of the nose.

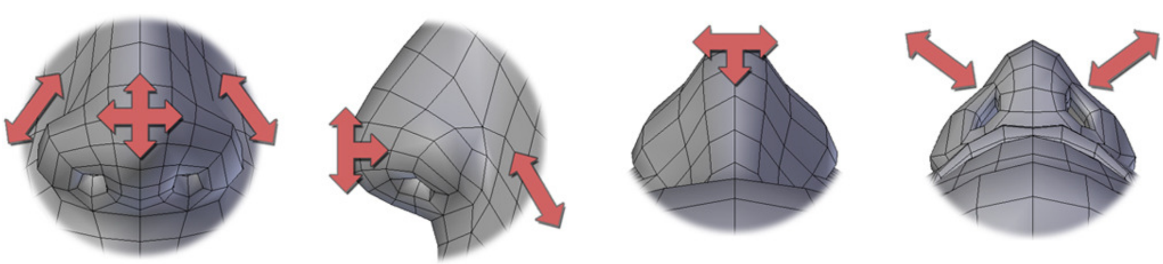


Figure 3.11: Front, side, top and bottom views of the nose of the human subject facial model.

The edges in the nose itself were modeled more according to the physicality of the nose than according to its behaviors because there is a more rigid morphology in this region. But the edges in the sides of the nose and nostrils follow specific behavioral conditions. The sides are subjected to the snarl behavior, which involves wrinkling that area, therefore these edges flow downwards into the cheeks, in accordance to the edges coming from the bottom eyelids. The dilation and compression of the nostrils suggest having circular loops in this region.

3.4.6 Cheeks

FACS [Ekman and Friesen 1978] mentions that the cheeks can be raised towards the bottom eyelid (AU 6 – Cheek Raiser and Lid Compressor) and sucked into the mouth (AD 35 – Suck) or blown outwards (AD 33 – Blow), either with the lips closed to keep the air inside the mouth or with the lips slightly opened to puff the air to the outside (AD 34 – Puff). Figure 3.12 illustrates the edge layout in the cheeks and their main motion directions.

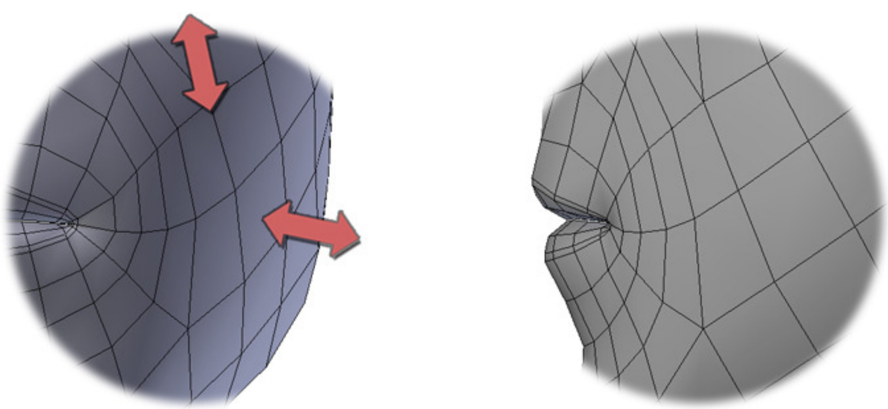


Figure 3.12: Front and profile views of the left cheek of the human subject facial model.

The movements of the cheeks suggest a large flexibility of their oval-shaped area when a gradual expansion or contraction of the cheeks occurs. For this reason, the polygons in the cheeks are organized as a uniform grid. This grid flows in accordance to the edges in the eyelids, nose and lips, resulting in a fair compromise in terms of edge integration.

3.4.7 Jaw

FACS [Ekman and Friesen 1978] mentions three consecutive actions for the jaw behavior (AUs 25, 26 and 27 – Lips Part, Jaw Drop and Mouth Stretch). These begin with exposing the inside of the mouth upon gradually lowering the mandible and go eventually into flattening and stretching the cheeks if opening the mouth too wide. The jaw can also be pushed forward in a thrust action to make the chin stand outwards (AD 29 – Jaw Thrust). It moves sideways to displace the chin from the midline of the face (AD 30 – Jaw Sideways) and also does the clench, a behavior that involves tightening the lower with the upper jaw region, producing bulges along the jaw bone closer to the ears region (AU 31 – Jaw Clencher). Figure 3.13 illustrates the edge flow in the jaw region.

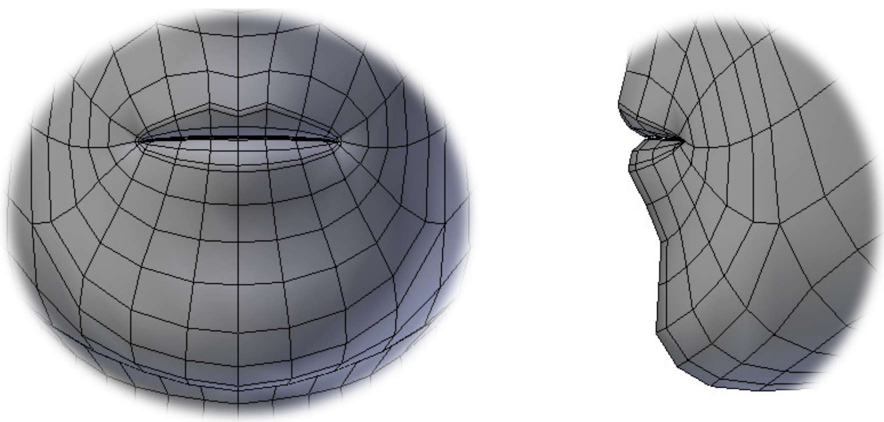


Figure 3.13: Front and profile views of the jaw of the human subject facial model.

As seen in Figure 3.13, the edges of the jaw follow from the cheeks downwards into the chin and inwards into the corners of the lips. The movements of the jaw do not suggest a large flexibility above the lips line, because those areas are more rigid. Instead, these areas follow along the jaw line to connect with the chin, lips and cheeks.

3.4.8 Lips

The lips are likely the most complex region in the face to model and then rig, as they are able to produce many different movements, especially during speech. According to FACS [Ekman and Friesen 1978], the upper and lower lips have vertical, horizontal and oblique motions that range in distance to allow for lip stretching, narrowing and also folding. There are a number of Action Units and Action Descriptors allowing this: pulling downwards the upper lip (AU 8 + 25 – Lips Toward Each Other), raising the center of the upper lip upwards (AU 10 – Upper Lip Raiser), pulling the upper lip upward (AU 11 – Nasolabial Furrow Deepener), pulling the corners of the lips obliquely (AU 12 – Lip Corner Puller), causing the cheeks to puff out (AU 13 – Sharp Lip Puller), tightening the corners of the mouth (AU 14 – Dimpler), pulling the corners of the lips downwards (AU 15 – Lip Corner Depressor), pulling the lower lip downwards (AU 16 and 16 + 25 – Lower Lip Depressor), pushing the lips forward (AU 18 – Lip Pucker), pulling the lips backwards (AU 20 – Lip Stretcher), funneling the lips outwards (AU 22 and 22 + 25 – Lip Funneler), tightening the lips making their red parts appear more narrow (AU 23 – Lip Teightener), pressing the lips together (AU 24 – Lip Presser), sucking the lips inwards (AU 28 and 26 + 28 – Lips Suck) and biting the lips with the teeth (AD 32 – Bite). Figure 3.14 illustrates the edge flow in the lips and their main vertices.

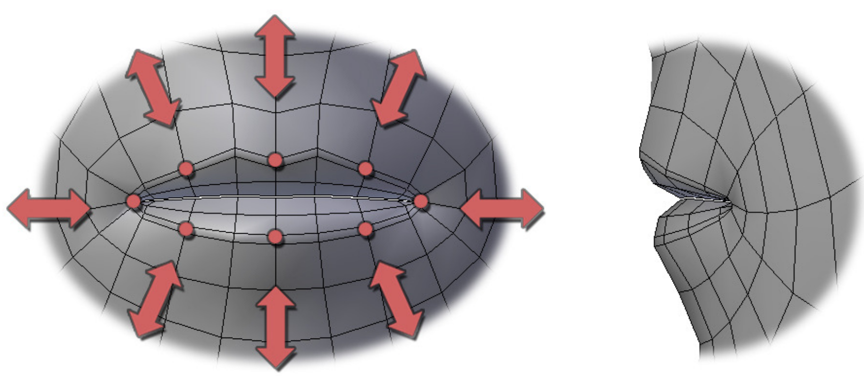


Figure 3.14: Front and profile views of the lips region of the human subject facial model.

The eight key feature points [Pandzic and Forchheimer 2002] seen in Figure 3.14 define the forces responsible for the main motions of the lips. In the human subject facial model they are coupled with eight more vertices which were added in-between. The concentric edges flow from the points, assuring a vertical, horizontal and oblique deformation for correct raise, lower, widen, narrow, frown and puckered actions. There is also an even spacing of the polygons along the lips and up until the lip corners and two extra edges on the top and bottom regions of each lip corner to maintain the detail when the mouth opens. The even spacing assures that the sides of the lips can extend keeping a proportional distance between the polygons as well as adapting to the edges in the sides of the nose and in the cheeks.

3.4.9 Tongue

FACS [Ekman and Friesen 1978] mentions that the tongue can protrude (AD 19 – Tongue Show), be pushed against the cheek to produce a bulge (AD 36 – Bulge) and it can wipe the lips (AD 37 – Lip Wipe). These behaviors suggest a high flexibility and therefore a polygonal arrangement with a higher vertex count and a carefully evened distribution. This part of the model is based in a royalty free model[4] (copyright TurboSquid) and in Sintel open-source model[5] (copyright Blender Foundation). Figure 3.15 shows the details of the tongue model as well as the inner cavities of the mouth and the teeth.

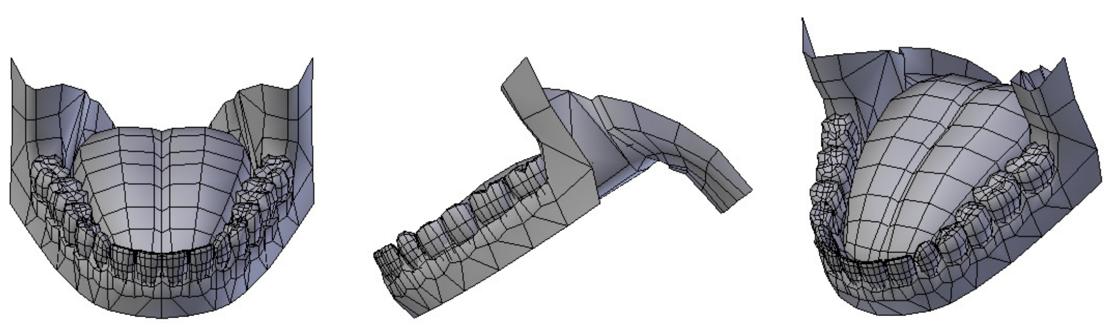


Figure 3.15: Front, profile and perspective views of the tongue of the human facial model.

[4] http://www.turbosquid.com/3d-models/old-man-head-max/245360
[5] https://durian.blender.org/sharing/

In this chapter, Sections 3.4.1 to 3.4.9 have provided a brief and comprehensive insight on how each facial region of the 3D facial model was built based on the guidelines by [Ekman and Friesen 1978] to achieve a model ready for rigging. The facial model obtained represents a balance between visual quality and workflow, to provide an accurate deformation and a manageable smooth skinning of the base model to the facial rig. The facial model was obtained using the modeling approaches described throughout this chapter and is now prepared to be rigged, because only a carefully modeled human subject face can allow a correct implementation of the optimized facial rig approaches presented in the following chapters. An accurate model helps the rigger to design the facial rig because the rig components are built relative to the different target emotions that need to be expressed in each facial region.

Chapter 4

Easy Facial Rig Mechanics

Setting the facial rig mechanics for a facial model is a key task that digital artists perform to assure subtle character facial skin deformations that result from the manipulation of the facial rig controls. But, the setup of the facial rig mechanics for a facial model is a laborious, time-consuming and disparate process. This chapter presents generic and validated guideline proposals to build facial rig mechanics as a complete hybrid bone-blendshape rig approach to pose a human subject facial model accurately and as a basis for other facial styles. These guidelines describe how to setup the fundamental facial rig mechanics, the bones and blendshapes (or BTMs), for the human subject facial model built previously in Section 3.4 of Chapter 3. A user experiment was also carried out with expert character animators, who have evaluated specific guidelines to setup a rig mechanical approach that sheds light on the configuration of the zipper and sticky lips effects, which are fundamental lip deformations for realistic character facial animation. The results reveal that these approaches help to ease the rigging process of the fundamental facial rig mechanics and of the zipper and sticky lips effects. In this chapter, the results presented in Section 4.1 are published in Generic and Certified Facial Rig Mechanical Approaches for Key Frame Character Animation [Bastos 2012] (see Appendix A-2012-1) and the results presented in Section 4.2 are published in Easy Character Facial Rig Mechanical Approaches to Animate Zipper and Sticky Lips Effects [Bastos 2013] (see Appendix A-2013-1).

4.1 Generic and Validated Facial Rig Mechanical Approaches for Key Frame Character Animation

The rig mechanical approaches presented in this section are focused in the human face as a basis for other facial styles, because any face is expected to have human features in order to be recognized by an audience as an expressive character, as seen in the cartoon character created by [Harkins 2005]. The implementation of the approaches is done in the popular open-source 3D tool Blender for the reasons highlighted previously in Section 3.2 of Chapter 3: (i) cause a significant impact in the growing community of Blender artists, (ii) due to Blender users being highly predisposed to validation, as seen in the preliminary result [Bastos et al. 2012a] described in Section 3.3.2 of Chapter 3 and (iii) to leverage the use of Blender in the industry as a qualified platform with an extended workflow.

The validation of the optimized facial rig mechanical design approaches involved a two stage process: (i) a realization of the behaviors of the human face by identifying its voluntary and involuntary actions – voluntary as deliberate and involuntary as spontaneous [Scherer and Ekman 1984], and (ii) the construction of the corresponding most adequate facial rig mechanical approaches for the facial behaviors, which means building a hybrid facial rig composed of bones and blendshapes. The goal is to provide guideline proposals for the rigger to build efficient facial rig mechanics using these techniques, which are provided in the major 3D tools. These approaches are not meant as a role model, their goal is to help bring a consensus into this vast and fragmented field of expertise, given the inherent difficulty in building an efficient facial rig.

The facial behaviors studied are those considered previously to build the human subject facial model, the muscular activity in FACS, categorized by [Ekman and Friesen 1978] into Action Units (AUs), Action Descriptors (ADs) and Gross Behavior Codes (GBCs). These are then realized as spatial transformations using the flight dynamics orientation, where “rotations are defined as roll, pitch, and yaw, mapped to X, Y and Z in a Z-up coordinate system" [O'Neill 2008]. The Z-up coordinate system is the default in Blender, where positive or negative values of X, Y and Z respectively correspond to the right or left, frontwards or backwards and upwards or downwards motion directions.

The construction of the facial rig mechanical approaches lies in determining which of the two main facial rigging techniques to use per facial region: bones, blendshapes (BTMs) or both. As described in the state of the art, bones can be realized as articulated hierarchical structures connected to each other to work in a parent-child relationship. Bone-driven deformations use either a smooth [Yang and Zhang 2006] or a rigid [Labourlette et al. 2005] skinning [Lewis et al. 2000], which is realized as the weighting of the influence that each bone has in each vertex of the facial model. Bones allow a rotation-based motion of the facial skin [Parke and Waters 2008] with hierarchically averaged weight values, being more suitable for arcing deformations [Ritchie et al. 2005] that occur in larger or denser facial regions. BTMs are differently sculpted duplicates of a facial model that blend from one to the other via a weight parameter. They allow a per-vertex tweak of the facial geometry [Parke and Waters 2008] with more precision, being more suited for predefined linear skin deformations [Ritchie et al. 2005] in lesser dense facial areas.

Other rigging techniques such as clusters, lattices and wires are not considered for the facial rig mechanical optimization approach to prevent inconsistency, as these techniques are (i) less often used by digital artists [Bastos et al. 2012b] and (ii) they are not always available in a similar way in the major 3D tools. In contrast, bones and BTMs are (i) more consistently found in the main 3D tools (e.g. respectively called skelegons and endomorphs in LightWave) and (ii) they are the most efficient, as digital artists use them the most [Bastos et al. 2012b].

To define whether to use bones or BTMs and how many per facial region, a set of four construction and deformation conditions were devised. Construction is relative to the work of the rigger to build the rig structures and deformation is the resulting visual quality of the facial movement in the character model. The four conditions are (i) hierarchy relation, (ii) predefined motion, (iii) motion dynamics and (iv) deformation area size. Hierarchy relation is the interaction of the rig mechanical components with each other via their skinning values. Predefined motion is the chance for behavior prediction and resulting mechanical degree of freedom and creativity. Motion dynamics is the evaluation of whether the motion is linear or non-linear. Deformation area size is the deformation dimension of the behavior, considered either small, medium or large.

The following sections present the generic and validated guideline proposals to place and configure the rig mechanical components for the different facial regions of the human subject facial model. The mechanical techniques are described in terms of whether to use bones, blendshapes or both, with the bone skinning paints illustrated. For a throughout verification of the facial rig mechanics developed in this section, please refer to the Blender file entitled *section_4.1*, available in the folder *sections blend files* in the accompanying CD.

4.1.1 Neck and Head

Table 4.1 and Figure 4.1 respectively list and illustrate the deformation areas of the neck and head, respectively the *tightener* and *swallow* (AU 21 and GBC 80) and the *shake back and forth*, *nod up and down*, *turn left or right*, *up or down*, *tilt left or right* and *forward or back* (GBCs 84, 85 and AUs 51, 52, 53, 54, 55, 56, 57, 58). These are, respectively, the upwards or downwards, the downwards, the backwards and frontwards, the upwards and downwards, the turning left or right, the upwards or downwards, the tilting left or right and the frontwards or backwards motion directions.

Anatomic Area	Motion Types		Corresponding AUs & GBCs	Motion Directions
Neck	Involuntary	Tightener	4	↓
	Voluntary	Swallow	80	↑↓
Head		Shake Back and Forth (SBF)	84	←→
		Nod Up and Down (NUD)	85	↑↓
		Turn Left or Right (LR)	51, 52	←→
		Up or Down (UD)	53, 54	↑↓
		Tilt Left or Right (TLR)	55, 56	←→
		Forward or Back (FB)	57, 58	←→

Table 4.1: Identification of the neck and the head behaviors from FACS [Ekman and Friesen 1978].

Table 4.2 and Figure 4.1 respectively list and illustrate the rig mechanics for the neck and head regions. The neck and the head hold a number of different behaviors for which a combination of bones and BTMs is required. Two BTMs are used, one per each linear motion, the swallowing and tightening of the neck. Two bones are also used to control the rotations of the neck and the head, respectively illustrated as A and B in Figure 4.1. Bone A deforms the

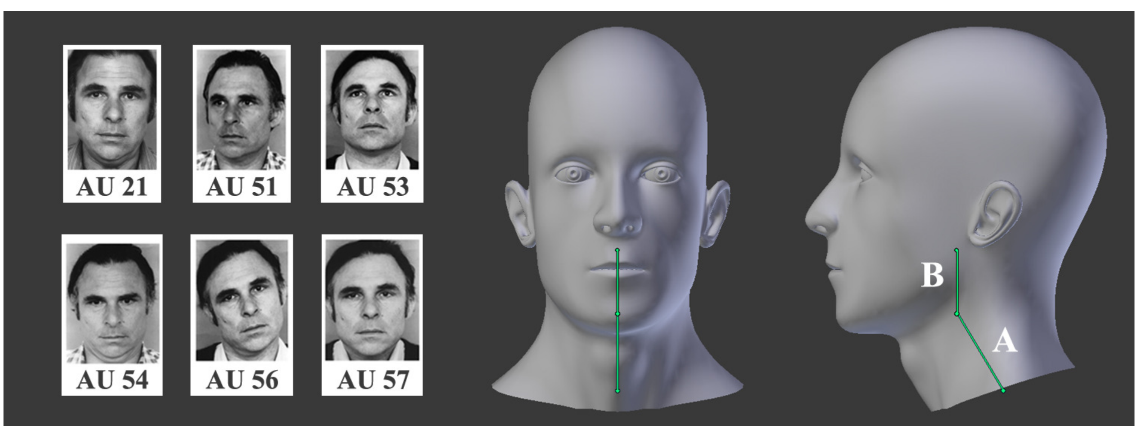


Figure 4.1: Behaviors of the neck and head and corresponding rig approach; left: muscular activity examples of the neck and head by [Ekman and Friesen 1978]; right: front and profile views of the neck and head bones, respectively A and B.

neck region and bone B serves as a hierarchy manager to which the upper and lower jaw bones (described ahead in Section 4.1.7) are linked to. Bone B does not need to deform the head region because that task is accomplished by the upper jaw bone.

Motion Types	Construction and Deformation Conditions				Resulting Mechanical Approach	Amount
	Hierarchy Relation	Predefined Motion	Motion Dynamics	Deformation Area Size		
Tightener	-	Yes	Linear	Medium	BTMs	2
Swallow						
SBF	Yes	-	Non-Linear	Large	Bones	2
NUD						
LR						
UD						
TLR						
FB						

Table 4.2: The rig mechanics validation approach for the neck and head regions.

In terms of *hierarchy relation*, using two bones assures a proper interaction between the neck and the head. The head tip will also link to the upper and lower jaw bones, which will average their skinning values with the neck region. The rotation based motion of the neck and head is a non-linear *motion dynamics*, not holding specific *predefined motion* and their *deformation area size* is large, when compared to the other facial regions, which suggests using bones for their rotation-based behaviors. The behaviors of the neck, the tightener and swallow, are rigged using BTMs because they do not have a direct *hierarchy relation* to each other or to other rig components, their linear motion is predefined. Furthermore, despite occupying a medium-sized *deformation area*, the skin deformations in the neck are more wrinkled and can be pre-built using the vertex-based sculpting benefit of BTMs, instead of using more bones. Figure 4.2 illustrates the bone skinning paints for the neck region.

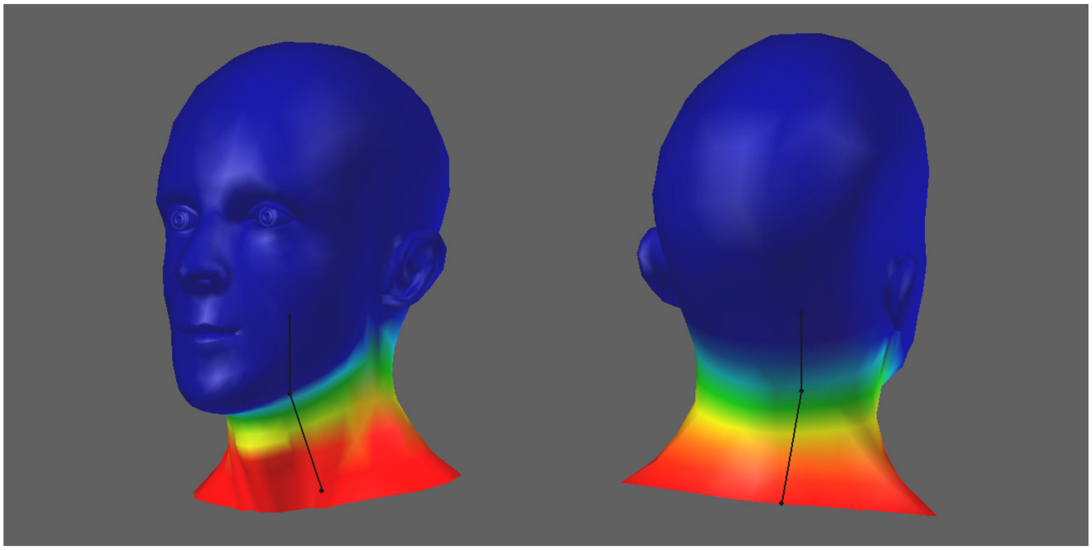


Figure 4.2: Skinning paints for the neck; left and right: frontwards and backwards views.

4.1.2 Eyebrows

Table 4.3 and Figure 4.2 respectively list and illustrate the deformation areas of the eyebrows: *inner*, *mid* and *outer* (AUs 1, 4, 2). These are, respectively, the upwards or downwards and/or inwards, the upwards or downwards and the upwards motion directions.

Anatomic Area	Motion Types		Corresponding AUs	Motion Directions
Eyebrows	Voluntary	Inner Brow Raiser (IBR)	1	↑↓←
		Mid Brow Raise/Lower (MBRL)	4	↑↓
		Outer Brow Raiser (OBR)	2	↑

Table 4.3: Identification of the eyebrows behaviors from FACS [Ekman and Friesen 1978].

Table 4.4 and Figure 4.3 respectively list and illustrate the eyebrows rig mechanics. Six bones are used, one per each eyebrow region (inner, mid, outer) and in each side of the face. The bones colored in green deform their respective areas, the bones colored in black serve as hierarchy managers that link the former to the tip of the head bone, which is located in the center of the facial model, as seen previously in Figure 4.1.

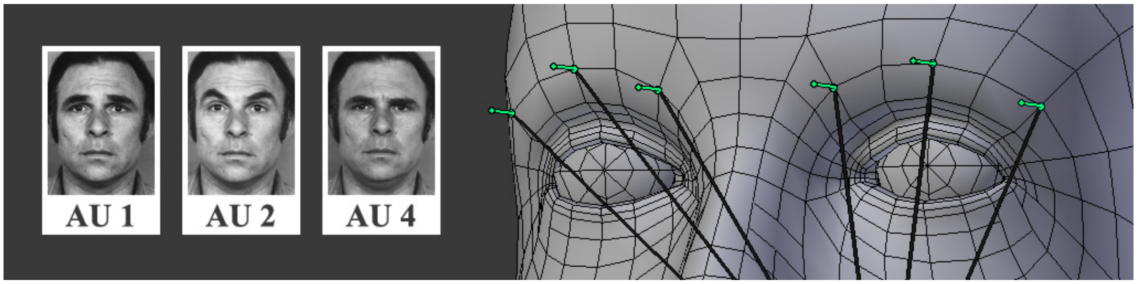


Figure 4.3: Behaviors of the eyebrows and corresponding rig approach; left: eyebrows muscular activity examples by [Ekman and Friesen 1978]; right: eyebrows bone mechanics.

Motion Types	Construction and Deformation Conditions				Resulting Mechanical Approach	Amount
	Hierarchy Relation	Predefined Motion	Motion Dynamics	Deformation Area Size		
IBR						
MBRL	Yes	Yes	Linear	Medium	Bones	6
OBR						

Table 4.4: The rig mechanics validation approach for the eyebrows regions.

In terms of *hierarchy relation*, using bones allows averaged skinning values between each bone located in each of the different areas of the left and right eyebrows. For *predefined motion* and *motion dynamics*, despite the eyebrows motions are quite predictable and linear, there are still eight directions per eyebrow to consider, meaning that a total of eight BTMs would be necessary instead of using only three bones. Therefore, resorting to BTMs would likely increase the setup work for the rigger and the overall complexity in this region. In terms of *deformation area size*, the behaviors of the eyebrows occupy sequential medium sized areas, which can be better encompassed and distributed using a reduced number of bones. Figure 4.4 illustrates the skinning paints for the bones responsible for deforming the inner, mid and outer regions of the eyebrows, which are colored in green in Figure 4.3.

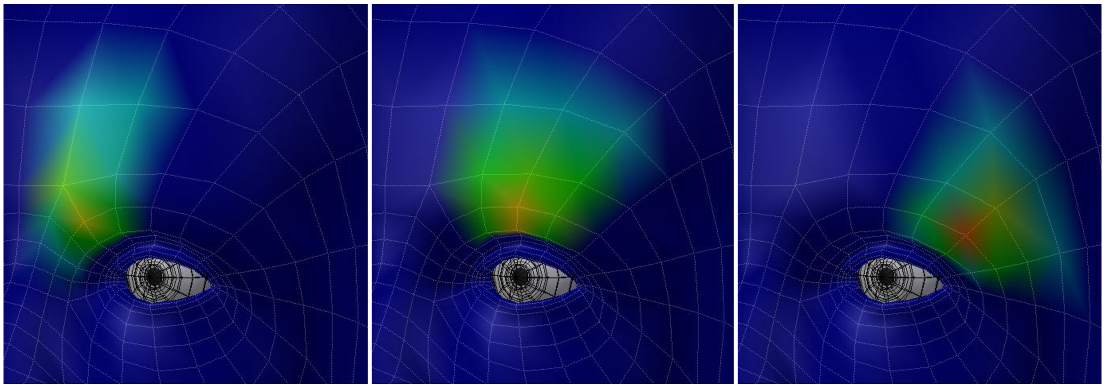


Figure 4.4: Left eyebrow skinning paints; left to right: inner, mid and outer eyebrow regions.

4.1.3 Eyelids

Table 4.5 and Figure 4.5 respectively list and illustrate the eyelids behaviors: *roll* and *pitch* (AUs 5, 7, 43, 45 and 46), respectively the voluntary upwards or downwards (X rotation in side view) and the involuntary left or right (Y rotation in front view) motion directions.

Anatomic Area	Motion Types & Axis		Corresponding AUs	Motion Directions
Eyelids	Voluntary	Roll (Rotation X)	5, 7, 43, 45, 46	↑↓
	Involuntary	Pitch (Rotation Y)		→←

Table 4.5: Identification of the eyelids behaviors from FACS [Ekman and Friesen 1978].

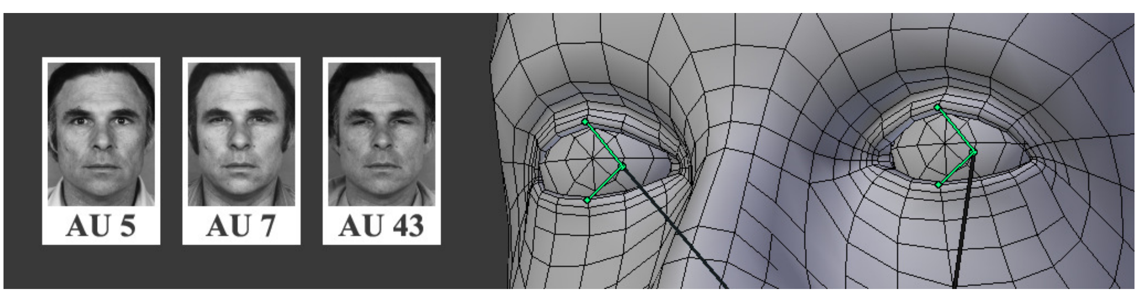


Figure 4.5: Behaviors of the eyelids and corresponding rig approach; left: eyelids muscular activity examples by [Ekman and Friesen 1978]; right: eyelids bone mechanics.

Table 4.6 and Figure 4.5 respectively list and illustrate the resulting rig mechanics used for the eyelids. Four bones are used, one per each eyelid (the upper and lower left and upper and lower right eyelids), based in [Maraffi 2003]. The green bones deform their respective areas, the bones colored in black are hierarchy managers linking the former to the tip of the head bone, which is located in the center of the facial model, as seen previously in Figure 4.1.

Motion Types	**Construction and Deformation Conditions**				**Resulting Mechanical Approach**	**Amount**
	Hierarchy Relation	**Predefined Motion**	**Motion Dynamics**	**Deformation Area Size**		
Roll	Yes	-	Non-linear	Small	Bones	4
Pitch						

Table 4.6: The rig mechanics validation approach for the eyelids regions.

Bones are used to rig the eyelids because they provide a *hierarchy relation* that eases the interaction of the skinning values between the eyelids, eyebrows and the upper jaw. The rotation-based motion of the top and the bottom eyelids around the eyeballs does not have a well established *predefined motion*, since the angle of rotation varies, suggesting bones to be used. The eyelids *motion dynamics* is non-linear, due to their rotation around the eyeballs, also indicating the use of bones. Not resorting to bones would require a large number of BTMs, likely three per each eyelid instead of a single bone, which would result in more work for the rigger. The *deformation area size* in the eyelids is small and there is a high vertex-count in the eyelids, which delays their sculpting process, denoting a more adequate use of bones in these regions. Figure 4.6 illustrates the skinning paints for the upper left and lower left eyelids.

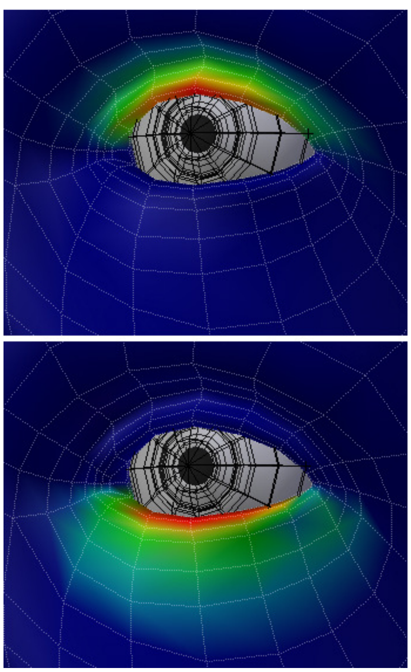


Figure 4.6: Skinning paints for the left eyelids; top and bottom: the upper and lower eyelids regions.

4.1.4 Eyeballs

Table 4.7 and Figure 4.7 respectively list and illustrate the eyeballs behaviors: *roll* (AUs 63 and 64) and *yaw* (AUs 61, 62, 65 and 66), respectively the upwards or downwards (X rotation in side view) and the left or right (Z rotation in top view) motion directions.

Anatomic Area	Motion Types & Axis		Corresponding AUs	Motion Directions
Eyeballs	Voluntary	Roll (Rotation X)	63, 64	↑↓
		Yaw (Rotation Z)	61, 62, 65, 66	→←

Table 4.7: Identification of the eyeballs behaviors from FACS [Ekman and Friesen 1978].

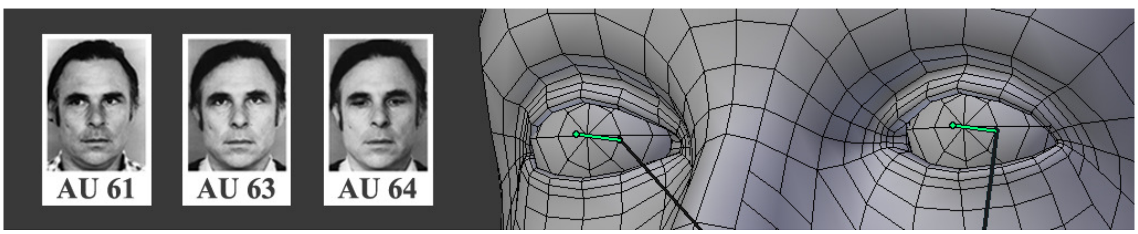


Figure 4.7: Behaviors of the eyeballs and corresponding rig approach; left: eyeballs muscular activity examples by [Ekman and Friesen 1978]; right: eyeballs bone mechanics.

Table 4.8 and Figure 4.7 respectively list and illustrate the eyeballs rig mechanics. One bone is used per eyeball, based in [Harkins 2005]. The green bones deform their respective areas, the bones colored in black are hierarchy managers that link the former to the tip of the head bone, which is located in the center of the facial model, as seen previously in Figure 4.1.

Motion Types	Construction and Deformation Conditions				Resulting Mechanical Approach	Amount
	Hierarchy Relation	Predefined Motion	Motion Dynamics	Deformation Area Size		
Roll	-	-	Non-linear	Medium	Bones	2
Yaw						

Table 4.8: The rig mechanics validation approach for the eyeballs.

The eyeballs were modeled as separate meshes from the facial skin, therefore they are not directly subjected to a *hierarchy relation* that involves elaborate skinning weights. They also do not have a *predefined motion* because their rotation angle varies due to having a non-linear *motion dynamics*. The former conditions suggest the use of bones to support their variety of rotations instead of using a large number of BTMs. Their *deformation area size* is medium and also contained, thus better encompassed using bones. The skinning paints for the eyeballs only require a full weight paint of all the vertices of each eyeball to each bone, as is illustrated in Figure 4.8 for the left eyeball as an example.

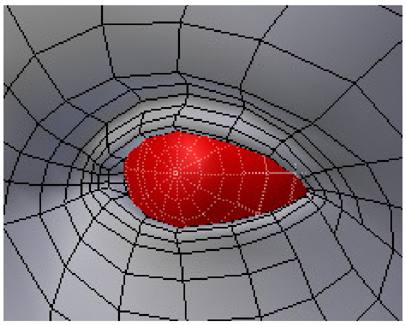


Figure 4.8: Skinning paints for the left eyeball.

4.1.5 Nose

Table 4.9 and Figure 4.9 respectively list and illustrate the nose behaviors: *sniff* (GBC 40), *snarl* (AU 9) and the nostrils *compress or dilate* (AUs 38, 39), respectively the upwards or downwards and/or left or right, the upwards or downwards and the inwards or outwards motion directions.

Anatomic Area	Motion Types		Corresponding AUs & GBCs	Motion Directions
Nose	Voluntary	Sniff	40	↑↓→←
	Involuntary	Snarl	9	↑↓
		Compress or Dilate (NCD)	38, 39	←→

Table 4.9: Identification of the nose behaviors from FACS [Ekman and Friesen 1978].

Table 4.10 and Figure 4.9 respectively list and illustrate the nose rig mechanics. Eight BTMs are used, four for the *sniff*, two for the *snarl* (one per each side of the face) and two for the *compress* and *dilate* behaviors.

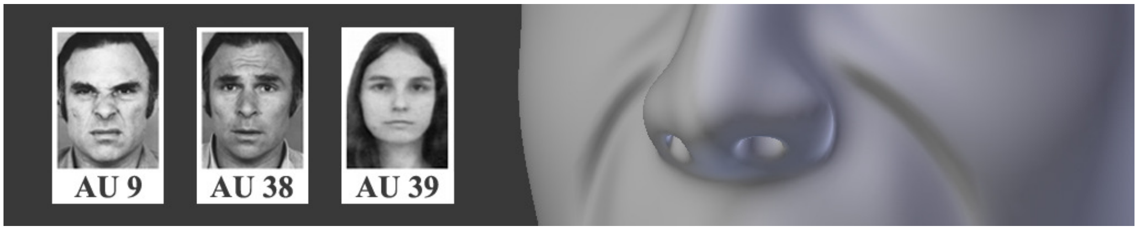


Figure 4.9: Behaviors of the nose and corresponding rig approach; left: nose muscular activity examples by [Ekman and Friesen 1978]; right: BTMs for the nose snarl behavior.

Motion Types	**Construction and Deformation Conditions**				**Resulting Mechanical Approach**	**Amount**
	Hierarchy Relation	**Predefined Motion**	**Motion Dynamics**	**Deformation Area Size**		
Sniff						4
Snarl	-	Yes	Linear	Medium	BTMs	2
NCD						2

Table 4.10: The rig mechanics validation approach for the nose.

As seen in Table 4.10, the motions of the nose can be narrowed to a number of *predefined motions* in areas with medium *deformation area sizes* responsible for specific wrinkles in those regions. In fact, “the inner, nasal branch of the sneering muscle gives the face a crinkled-up look when it contracts” [Faigin 1990]. Therefore, the use of blendshapes is an ideal approach, because they provide a more complete control over the vertices to build more defined wrinkles. BTMs also provide an independent sculpting of the linear *motion dynamics* of the nose, which are not dependent of *hierarchy relations* relative to other components of the rig since the use of BTMs to rig these behaviors will not produce excessive bone skinning weights in this region. Any excessive bone weights would be hard to combine with existing bone weights in other regions such as the eyelids and the lips, thus saving work to the rigger.

4.1.6 Cheeks

Table 4.11 and Figure 4.10 respectively list and illustrate the cheeks behaviors: *cheek* *raiser/lower* (AU 6) and *suck or blow/puff* (ADs 33, 34 and 35), respectively the upwards or downwards and the inwards or outwards motion directions.

Anatomic Area	**Motion Types**		**Corresponding AUs & ADs**	**Motion Directions**
Cheeks	Voluntary	Cheek Raiser/Lower (CRL)	6	↑↓
		Suck or Blow/Puff (SBP)	33, 34, 35	←→

Table 4.11: Identification of the cheeks behaviors from FACS [Ekman and Friesen 1978].

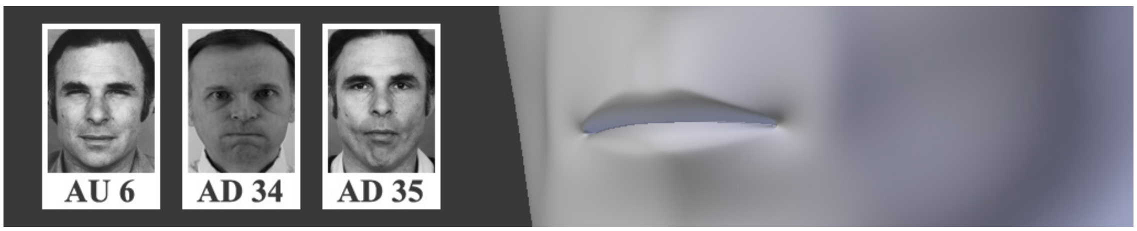


Figure 4.10: Behaviors of the cheeks and corresponding rig approach; left: cheeks muscular activity examples by [Ekman and Friesen 1978]; right: BTMs for the cheeks suck behavior.

Table 4.12 and Figure 4.10 respectively list and illustrate the cheeks rig mechanics. Four BTMs are used for the *CRL*, two per each cheek, and another four BTMs for the *SBP*, also two per each cheek.

As with the nose, the cheeks behaviors can also be narrowed down to a number of *predefined motions* responsible for deformations with a linear *motion dynamics* in areas with medium *deformation area sizes*. In addition, the behaviors of the cheeks are not directly dependent of *hierarchy relations* relative to the skinning of other components of the facial rig

Motion Types	**Construction and Deformation Conditions**				**Resulting Mechanical Approach**	**Amount**
	Hierarchy Relation	**Predefined Motion**	**Motion Dynamics**	**Deformation Area Size**		
CRL	-	Yes	Linear	Medium	BTMs	4
SBP						4

Table 4.12: The rig mechanics validation approach for the cheeks regions.

mechanics. Furthermore, the sculpting ability provided by the use of blendshapes will allow the rigger to simulate with more precision the cheeks behaviors, which include a manageable vertex-count to ease sculpting. In this way, there is also no cluttering of the already existing skinning weights since bones are not used in the cheeks.

4.1.7 Jaw

Table 4.13 and Figure 4.11 respectively list and illustrate the jaw behaviors: *roll* (AUs 25, 26, 27), *thrust* (AD 29), *yaw* (AD 30) and *flex* (AD 31), respectively the upwards or downwards (X rotation in side view), frontwards or backwards (Y translation in side view), left or right (Z rotation in top view) and the inwards or outwards motion directions, or bulging in the sides of the face below the ear, which results from the jaw being clenched. Therefore, the clenching motion of the jaw is better perceived when the face is viewed from the front.

Anatomic Area	Motion Types & Axis		Corresponding AUs & ADs	Motion Directions
Jaw	Voluntary	Roll (Rotation X)	25, 26, 27	↑↓
		Thrust (Translation Y)	29	↙↗
		Yaw (Rotation Z)	30	→←
	Involuntary	Flex	31	←→

Table 4.13: Identification of the jaw behaviors from FACS [Ekman and Friesen 1978].

Table 4.14 and Figure 4.11 respectively list and illustrate the jaw rig mechanics. Here two bones are used to support the *roll*, *thrust* and *yaw* behaviors. One bone for the upper part of the face and the other for the lower part, based in the upper and lower jaw by [Miller 2004], but here with identical idle positions. Two BTMs are also used, one per each side of the face, to deal with the *flex* behavior. The upper and lower jaw bones, which are colored in green in Figure 4.11, deform their respective areas and they are directly linked to the tip of the head bone, which is located in the center of the facial model, as seen previously in Figure 4.1.

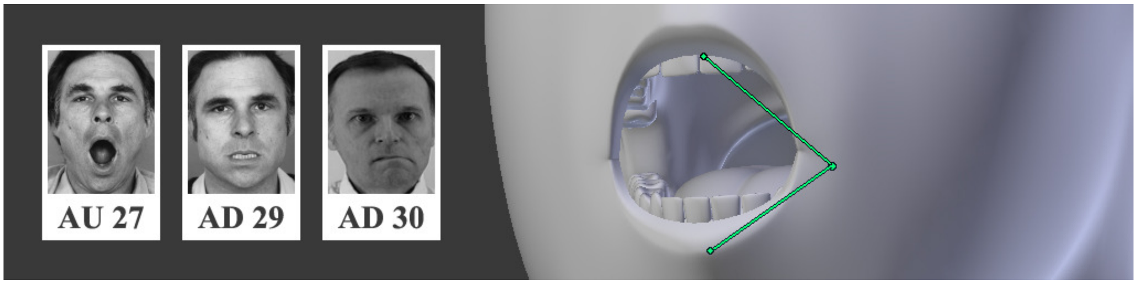


Figure 4.11: Behaviors of the jaw and corresponding rig approach; left: jaw muscular activity examples by [Ekman and Friesen 1978]; right: a deformation caused by the upper and lower jaw bones.

Motion Types	**Construction and Deformation Conditions**				**Resulting Mechanical Approach**	**Amount**
	Hierarchy Relation	**Predefined Motion**	**Motion Dynamics**	**Deformation Area Size**		
Roll	Yes	-	Non-linear	Large	Bones	2
Thrust			Linear			
Yaw			Non-linear			
Flex	-	Yes	Linear	Small	BTMs	2

Table 4.14: The rig mechanics validation approach for the jaw region.

The jaw region uses a rig based on bones and BTMs. In terms of *hierarchy relation*, two bones for the *roll*, *thrust* and *yaw* (upper and lower jaw) prevail as many BTMs would be required instead. These bones allow an averaged skinning between each other and also other bones in the rig (e.g. neck and lips). In terms of *predefined motion*, bones suit more likely the jaw since their translation and rotation values cannot be predefined. For the *flex* behavior, pre-built BTMs are more indicated due to this behavior being a subtle predefined involuntary motion driven by the clenching of the jaw. In terms of *motion dynamics*, bones for the *roll* and *yaw* behaviors fit the non-linear nature of these motions. The *thrust* behavior, despite having a linear motion, can be dealt with by sympathy using the forward and backwards translation of the existent lower jaw bone, which is already skinned to this region via the skinning values for the *roll* and *yaw* behaviors. The *flex* behavior is also a linear motion, in this case a lateral bulge of the facial skin that occurs on both sides of the face that are closer to the ears. But each bulge in each side of the face has its own direction, being more adequate to rig using the linear-based motion of BTMs, which also contributes to balance the use of bones. In terms of the *deformation area size*, the use of bones for the *roll*, *thrust* and *yaw* behaviors encompass the large areas influenced by these behaviors, allowing a faster skinning in these areas. In opposite, the *flex* behavior occurs in two small *deformation area sizes*, with a more reduced vertex-count, which are likely not to be tedious for the rigger to sculpt using BTMs. Figure 4.12 illustrates the bone skinning paints configured for the upper and lower jaw regions.

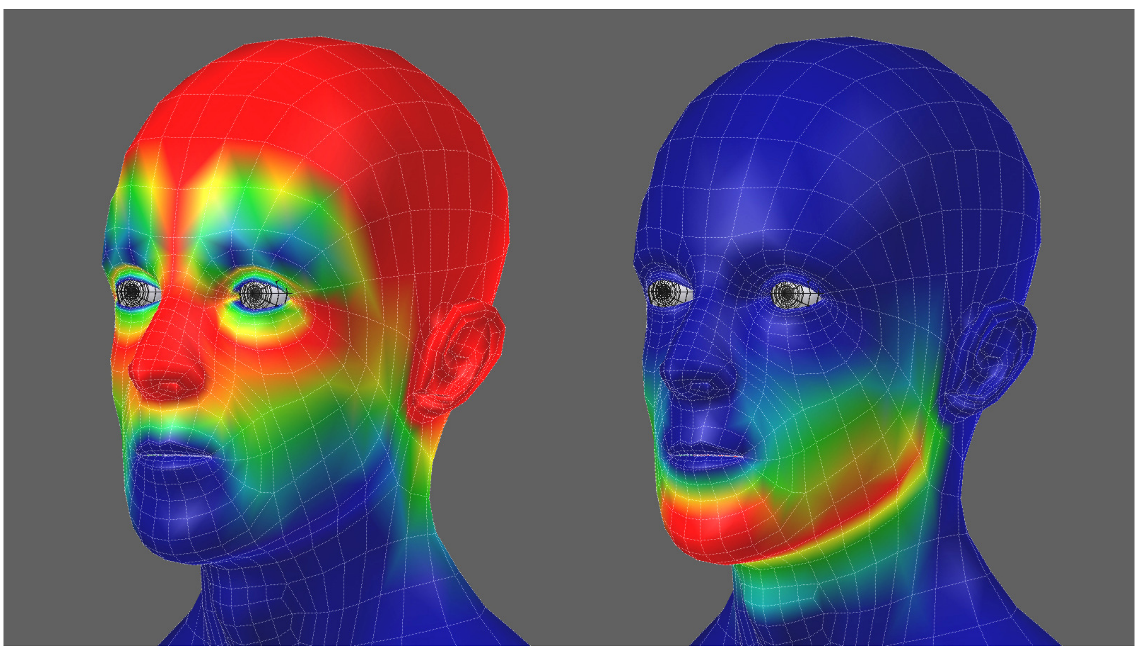


Figure 4.12: Skinning paints for the jaw; left: upper jaw; right: lower jaw.

4.1.8 Lips

Table 4.15 and Figure 4.13 respectively list and illustrate the lips behaviors: *overall shaping* (AUs 11, 13, 8+25, ADs 32, 36, 37, GBC 50), *upper lip raiser* (AU 10), *lower lip depressor*

(AU 16+25), *narrow or widen* (AUs 20 and 23), *frown or smile* (AUs 12, 14 and 15) and *pucker, suck, presser or funneler* (AUs 18, 22, 24 and 28), respectively the upwards or downwards and/or left or right and/or frontwards or backwards, the upwards, the downwards, the left or right, the down and left or right and/or the upwards and left or right, and the upwards or downwards and left or right and frontwards or backwards motion directions.

Anatomic Area	**Motion Types & Axis**		**Corresponding AUs, ADs & GBCs**	**Motion Directions**
Lips	Voluntary	Overall Shaping (OS) (Translation/Rotation XYZ)	11, 13, 8+25, 32, 36, 37, 50	↑↓→←↙↗
		Upper Lip Raiser (ULR)	10	↑
		Lower Lip Depressor (LLD)	16+25	↓
		Narrow or Widen (NW)	20, 23	→←
		Frown or Smile (FS)	12, 14, 15	↓→←↑→←
		Pucker, Suck, Presser or Funneler (PSPF)	18, 22, 24, 28	↑↓→←↙↗

Table 4.15: Identification of the lips behaviors from FACS [Ekman and Friesen 1978].

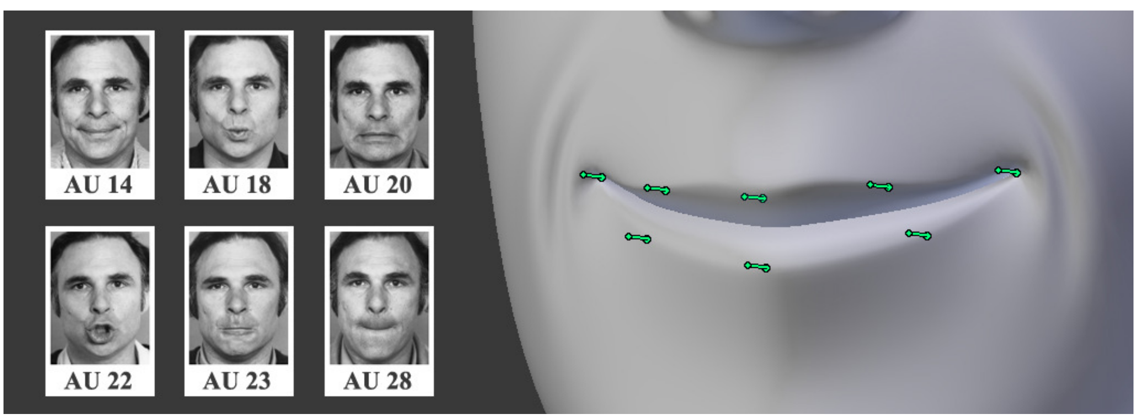


Figure 4.13: Behaviors of the lips and corresponding rig approach; left: lips muscular activity examples by [Ekman and Friesen 1978]; right: lips bones and BTMs.

Table 4.16 and Figure 4.13 respectively list and illustrate the fundamental rig mechanics of the lips, which involve using eight bones for their *overall shaping*, which is equivalent to the number of feature points used by [Pandzic and Forchheimer 2002]. Another fourteen BTMs are used for the remaining behaviors of the lips. To achieve more complex expressions such as visemes, another eight bones can be added for instance, although by default any extra bones may be hidden and only accessible to the animator as an extra option to achieve further manipulation precision. The bones colored in green deform their respective areas and other secondary linkage bones, which are hidden in Figure 4.13 to ease the visualization, connect the main lips deformation bones to the tip of the head bone, which is located in the center of the facial model, as seen previously in Figure 4.1.

Motion Types	Construction and Deformation Conditions				Resulting Mechanical Approach	Amount
	Hierarchy Relation	Predefined Motion	Motion Dynamics	Deformation Area Size		
Overall Shaping	Yes	-	Linear and Non-linear	Medium	Bones	8
ULR	-	Yes			BTMs	1
LLD						1
NW						4
FS						4
PSPF						4

Table 4.16: The rig mechanics validation approach for the lips.

The lips are the most complex region to rig in a facial model because a more complex combination of bones and BTMs is required to support their vast range of behaviors. In terms of *hierarchy relation*, a distribution of eight bones along the upper and lower lips, as seen in Figure 4.13, assures a proper interaction with other rig components via the averaging of the skinning values (e.g. with the upper and lower jaw bones). The distribution of eight bones along the lips also allows for an accurate *overall shaping* of the lips by manipulating each lip bone to adjust deformation details in the lips which do not have an exact *predefined motion*. If necessary, the addition of eight extra bones in-between the suggested base eight bones, can assure a higher precision of deformation. The *motion dynamics* in this case is both linear and non-linear as the lips can be moved and twisted in each of their individual areas, hence the use of bones to better support both translation and rotation transformations. The lips medium *deformation area size* suggests the use of bones for the purpose of pose adjustment and the use of BTMs to achieve general predefined poses. Besides the *overall shaping*, the remaining behaviors of the lips, listed previously in Table 4.15, are more common lip expressions, sometimes extended to include well defined poses such as visemes. Thus, the former can be considered *predefined motions* with no direct *hierarchy relation* to other components in the facial rig. Although occupying a medium-sized *deformation area*, these deformations are more wrinkled and can be pre-built using the vertex-based sculpting benefit of BTMs instead of increasing the existent number of bones. The bone skinning paints for the left corner area of the lips is illustrated in Figure 4.14 as an example for the remaining bones that are surrounding the lips region as seen in Figure 4.13. The lips region also includes the zipper and sticky lips effects, for which an in-depth rig mechanical optimization is available ahead in Section 4.2.

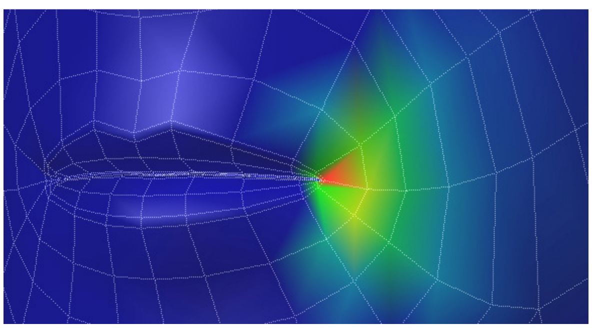


Figure 4.14: Skinning paints for the left corner area of the lips region.

4.1.9 Tongue

The tongue is only briefly mentioned in FACS [Ekman and Friesen 1978], with the indication of the tongue show, bulge and lip wipe, respectively ADs 19, 36 and 37. The former, coupled with the generic tongue positions presented by [Fleming and Dobbs 1998], seen previously in Figure 2.8 in Section 2.3.2 in Chapter 2, and based on the approaches by [Bibliowicz 2005] and [Holly 2006], have led to the development of a rig mechanical approach with a flexible hierarchy of bones for the tongue, which is illustrated in Figure 4.15.

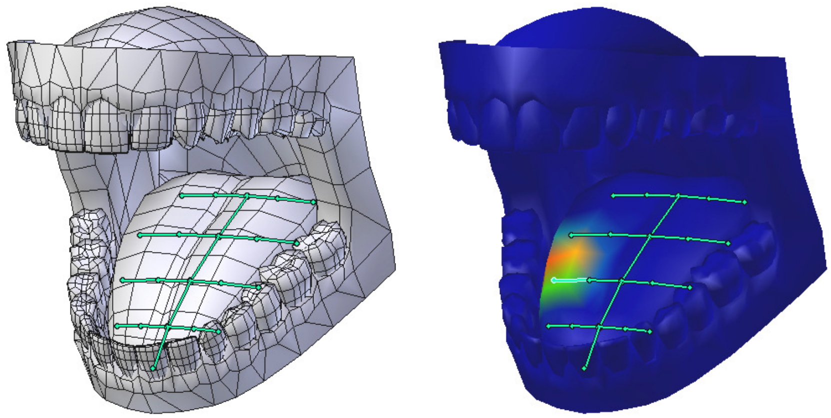


Figure 4.15: Rig approach for the tongue; left: tongue skeleton with a main hierarchy placed along the center of the tongue model and with lateral bones emerging from the main bone hierarchy; right: skinning paints seen in the model that correspond to the highlighted lateral bone in the right side of the tongue skeleton.

The skeletal approach seen in Figure 4.15 takes the shape of a hierarchical tree, with main bone branches arranged in sequence and distributed along the center of the tongue and then two secondary lateral bone branches spreading from the sequential nodes of the main branch. This skeletal structure works with constraints assigned to the second, third and forth bones in the main hierarchy that copy the rotation of the first bone, which is the root bone of the main hierarchy. The constraints allow the tongue to roll, turn, pitch and curl by only rotating the first bone. The lateral bones use the same constraint approach to achieve a U-shape of the tongue. In this case, the lateral bones located in the outside copy the rotations of the lateral bones located in the inside, allowing the rotation of the inside bones to affect the outside bones. The skinning paints for one of the lateral sections of the tongue is illustrated in Figure 4.15 as an example.

4.1.10 Facial Deformation Overview

After the optimization process of the fundamental facial rig mechanics for the facial model, a texture was built based on the photographs obtained of the human subject and it was applied to the facial model in Blender. Figure 4.16 illustrates the general deformation abilities of the facial model by demonstrating an expressive facial pose compared with the original neutral pose, with the facial texture applied to improve the realism of the model.

Figure 4.16: Final look of the human subject facial model in a neutral pose (top row) and in an expressive pose (bottom row); left: shading with wireframe; middle left: shading with wireframe and textures; middle right: shading with textures; right: shading with textures and subdivision surface to smooth the model.

Although presented in a white background with no other environment elements, the final look of the facial model seen in Figure 4.16 demonstrates a subtle level of realism in terms of the character's skin deformation, as a result of the optimized facial rig mechanical approaches presented throughout this section. The approaches are considered generic due to their focus in the human face as a basis for other facial styles. The approaches act as validated guideline proposals for human facial rig mechanical construction because they are based in a set of construction and deformation conditions described and demonstrated previously per each facial region. The approaches were defined based on the behaviors found in FACS [Ekman and Friesen 1978], because this is likely the most detailed reference in facial behavior.

The fundamental facial rig mechanics presented throughout this section allow the facial model to deform with a high level of realism and credibility. But the lips region contains two other behaviors which are more complex to simulate. These behaviors are not described in FACS [Ekman and Friesen 1978] but they are relevant to further increase the realism of a facial model, especially in close-up shots of the face, in which the details of deformation in the lips become more relevant. These behaviors are known in the industry as the zipper and sticky lips effects and they were also considered for the facial rig mechanical optimization. Section 4.2 describes these behaviors and presents the facial rig mechanical approaches devised to support and shed light on their preparation for animation.

To conclude this section, Tables 4.17 and 4.18 respectively resume the human facial behaviors identified throughout this section by studying FACS [Ekman and Friesen 1978] and the corresponding fundamental rig mechanics developed for each facial region of the human subject facial model.

Anatomic Area	Motion Types & Axis		Corresponding AUs, ADs & GBCs	Motion Directions
Neck and Head	Involuntary	Tightener	4	↓
	Voluntary	Swallow	80	↑↓
		Shake Back and Forth (SBF)	84	←→
		Nod Up and Down (NUD)	85	↑↓
		Turn Left or Right (LR)	51, 52	←→
		Up or Down (UD)	53, 54	↑↓
		Tilt Left or Right (TLR)	55, 56	←→
		Forward or Back (FB)	57, 58	←→
Eyebrows	Voluntary	Inner Brow Raiser (IBR)	1	↑↓←
		Mid Brow Raise/Lower (MBRL)	4	↑↓
		Outer Brow Raiser (OBR)	2	↑
Eyelids	Voluntary	Roll (Rotation X)	5, 7, 43, 45, 46	↑↓
	Involuntary	Pitch (Rotation Y)		→←
Eyeballs	Voluntary	Roll (Rotation X)	63, 64	↑↓
		Yaw (Rotation Z)	61, 62, 65, 66	→←
Nose	Voluntary	Sniff	40	↑↓→←
	Involuntary	Snarl	9	↑↓
		Compress or Dilate (NCD)	38, 39	←→
Cheeks	Voluntary	Cheek Raiser/Lower (CRL)	6	↑↓
		Suck or Blow/Puff (SBP)	33, 34, 35	←→
Jaw	Voluntary	Roll (Rotation X)	25, 26, 27	↑↓
		Thrust (Translation Y)	29	↙↗
		Yaw (Rotation Z)	30	→←
	Involuntary	Flex	31	←→
Lips	Voluntary	Overall Shaping (Translation/Rotation XYZ)	11, 13, 8+25, 32, 36, 37, 50	↑↓→←↙↗
		Upper Lip Raiser (ULR)	10	↑
		Lower Lip Depressor (LLD)	16+25	↓
		Narrow or Widen (NW)	20, 23	→←
		Frown or Smile (FS)	12, 14, 15	↓→←↑→←
		Pucker, Suck, Presser or Funneler (PSPF)	18, 22, 24, 28	↑↓→←↙↗

Table 4.17: Resume of the behaviors of the human face per each facial region based on FACS [Ekman and Friesen 1978].

Anatomic Area	Motion Types	Construction and Deformation Conditions				Resulting Mechanical Approach	Amount
		Hierarchy Relation	Predefined Motion	Motion Dynamics	Deformation Area Size		
Neck and Head	Tightener	-	Yes	Linear	Medium	BTMs	2
	Swallow						
	SBF	Yes	-	Non-Linear	Large	Bones	2
	NUD						
	LR						
	UD						
	TLR						
	FB						
Eyebrows	IBR	Yes	Yes	Linear	Medium	Bones	6
	MBRL						
	OBR						
Eyelids	Roll	Yes	-	Non-linear	Small	Bones	4
	Pitch						
Eyeballs	Roll	-	-	Non-linear	Medium	Bones	2
	Yaw						
Nose	Sniff	-	Yes	Linear	Medium	BTMs	4
	Snarl						2
	NCD						2
Cheeks	CRL	-	Yes	Linear	Medium	BTMs	4
	SBP						4
Jaw	Roll	Yes	-	Non-linear	Large	Bones	2
	Thrust			Linear			
	Yaw			Non-linear			
	Flex	-	Yes	Linear	Small	BTMs	2
Lips	Overall Shaping	Yes	-	Linear and Non-linear	Medium	Bones	8
	ULR	-	Yes			BTMs	1
	LLD						1
	NW						4
	FS						4
	PSPF						4

Table 4.18: Resume of the fundamental rig mechanics definition per each facial region of the human subject facial model.

4.2 Easy Character Facial Rig Mechanical Approaches to Animate Zipper and Sticky Lips Effects

This section presents optimized rig setups in Blender for the zipper and sticky lips effects as design approaches to facilitate their laborious and time-consuming development. The effect known as zipper lips is a voluntary behavior realized as the action in effort of closing the lips together while keeping the jaw open, as in zipping the lips [Osipa 2007]. The effect known as sticky lips is the natural tendency that real-world lips have to adhere to each other toward their corners while opening the jaw [Osipa 2010]. Sticky lips is a more complex behavior than zipper lips because it involves deforming the inner portions of the lips. Sticky lips occurs mostly during speech when the lips get dry. It can also occur when a person is not speaking for some time and then opens the mouth slowly, causing the upper and lower lips to initially appear glued and then slowly begin separating from each other. These effects are known to require complex rig setups which are not easy to learn and to carry out by digital artists. The approaches presented in this section are straightforward rig mechanical setup guidelines that support the zipper and sticky lips behaviors for the human subject facial model and provide a closure to the rig mechanical optimization process described throughout this chapter.

The state of the art in character facial rigging [Orvalho et al. 2012, Bastos et al. 2012b] includes limited literature on the zipper and sticky lips effects, despite they contribute highly to enhance character realism in close-up shots of a face. The reasons for the lack of literature regarding these effects are (i) the inherent complexity that a facial rig requires to support them, (ii) only few artists show off their work in these topics (e.g. Chad Vernon [Vernon 2010]) and (iii) that even fewer artists describe their rig setups. A single literary reference, by Jason Osipa [Osipa 2010], presents a detailed description of a sticky lips rig. In this section a clear distinction and throughout description of the zipper and sticky lips is presented as well as their interaction with the rig mechanical components of the neck, head, jaw and lips. In light of the former, the approaches presented in this section are intended as a straightforward and complete learning reference for riggers and artists in general to reproduce these effects accurately. These approaches are not meant as an imperative model, their goal is to help bring a consensus to the rig setup of the complex zipper and sticky lips behaviors, which are not easily found and reproduced in the state of the art. For a throughout verification of the facial rig mechanics developed in this section, please refer to the Blender file entitled *section_4.2*, available in the folder *sections blend files* in the accompanying CD.

4.2.1 Validation of the Easy Zipper and Sticky Lips Rig Approaches

The rig approaches for the zipper and sticky lips act as guideline proposals validated via these conditions: (i) the rig approaches integrate with existing lips controls (zipper and sticky lips couple with the regular lip controls built previously in Section 4.1.8), (ii) they are extensible to different numbers of lip controls (e.g. 8 or 16), (iii) they involve only bones, constraints and drivers, which are universal rigging techniques available in the major 3D tools [Bastos et al. 2012b], (iv) they are bone-based, therefore they have a higher compatibility with game development since most game engines do not support blendshapes [Skonicki 2008], (v) they do not require programming skills, being more accessible to less experienced artists, (vi) they are open-source, therefore available for all digital artists, (vii) they are cross operating system as Blender is available for different platforms and (viii) they were validated in a qualitative user experiment with expert artists (see ahead in Sections 4.2.4 and 4.2.5) and considered by the users as a relevant contribution for a topic little documented in the state of the art.

The rig setup for the zipper and sticky lips effects was carried out on top of the base rig structures for the neck and head, jaw and lips, which were described previously, respectively in Sections 4.1.1, 4.1.7 and 4.1.8. Figure 4.17 illustrates the rig setups for the former facial regions combined and further described to act as a basis for the zipper and sticky lips rig setups, which are respectively described ahead in Sections 4.2.2 and 4.2.3.

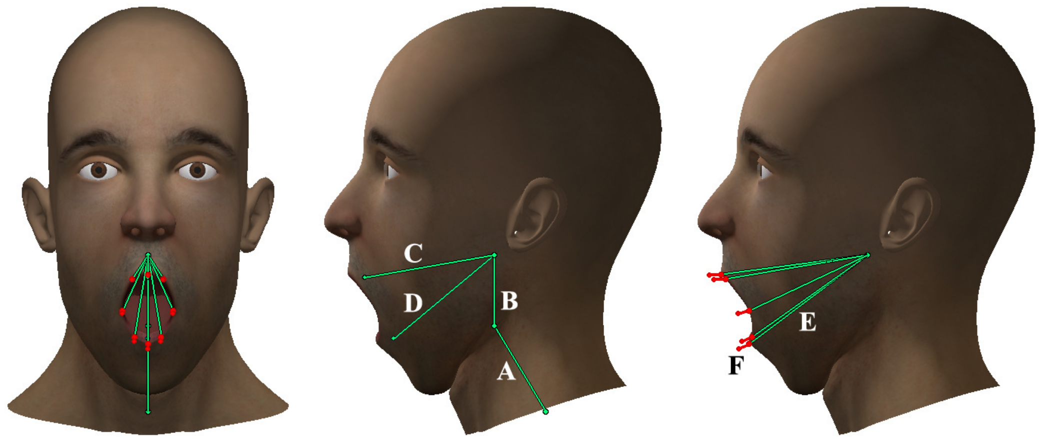


Figure 4.17: Views of the facial model and base rig mechanical structure for the lips; left: front view with eight control points for the lips; middle and right: profile views showing the two layers of the base rig mechanical structure.

The bones seen in Figure 4.17 are organized in two layers. The first layer contains the neck, head and the upper and lower jaw bones (respectively A, B, C and D). The neck region is deformed by bone A. Bone B acts as a hierarchy manager, it links the jaw bones to the neck, it does not deform the head because this task is dealt with by the jaw bones, C and D, which respectively deform the upper jaw area (the cranium) and the lower jaw area.

In the second layer a group of bones is distributed along the lips to deform them. Bones E and F are replicated in the eight lip locations seen in Figure 4.17. Bone E connects bone F to the tip of the head bone and is constrained to copy the rotation of the upper and lower jaw bones. The constraints applied to the top half lip connection bones, located above the center line of the lips, are set so that their owners are more influenced by the rotation of the upper jaw bone. The constraints applied to the lip connection bones in the bottom half portion, located below the center line of the lips, cause their owners to be more influenced by the rotation of the lower jaw bone. Lastly, Bone F is parented to bone E and it deforms the geometry of the lips which is located closest to it.

4.2.2 The Zipper Lips Rig Mechanical Approach

The zipper lips effect is a voluntary behavior realized as the action in effort of closing the lips together while keeping the jaw open, as in zipping the lips [Osipa 2007]. While the jaw is in a downwards pose, the bones deforming the lips return to their original positions in order to close the mouth while the jaw is maintained in its downwards position. Figure 4.18 illustrates the rig mechanical setup developed to manage the zipper lips effect.

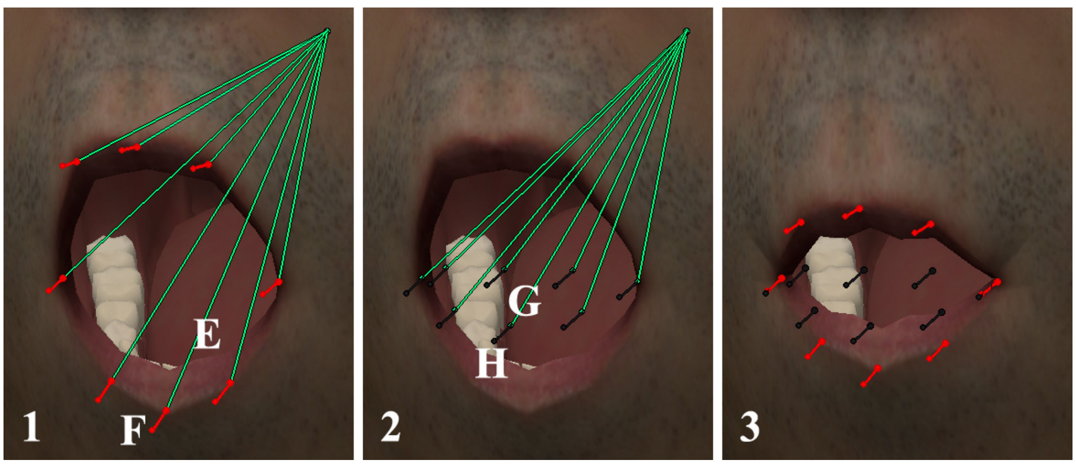


Figure 4.18: Rig mechanical setup for the zipper lips effect; left: layer with base rig mechanical structure (the lips deformation bones); middle: layer with the zipper bones; right: the zipper effect applied to half strength.

To generate the zipper lips effect a new extra layer of bones seen in part 2 of Figure 4.18 is added to the existing skeletal structure described previously in Section 4.2.1. The extra layer has the same rig structure of the layer seen in part 1 of Figure 4.18 but it serves as a reference for the original positions of the bones deforming the lips. Therefore, the bones in this extra layer are always posed to keep the mouth closed and they are also hidden from the animator to prevent them to be selected accidentally.

To achieve the zipper effect, a constraint is applied to each bone deforming the lips (e.g. bone F) to copy the transformations of its corresponding zipper bone (e.g. bone H). To make sure that the zipper effect works independently of mouth shapes that can eventually be posed by the rotation of the upper or lower jaw bones, the linkage bones of the zipper bones layer (e.g. bone G) are constrained to copy the rotation of the upper and lower jaw bones with half influence in each in order to average their movements. The process described for bones G and H is then repeated for the remaining bones of the zipper bones layer which is seen in part 2 of Figure 4.18.

Lastly, to allow the user to control the amount that the zipper effect is applied, a driver is assigned to the influence of each of the copy transforms constraints which were associated to each bone responsible for deforming the lips (e.g. bone F). The influence value can then be driven using a rig control manipulator (e.g. a window slider or a viewport gizmo).

4.2.3 The Sticky Lips Rig Mechanical Approach

The sticky lips effect is the natural tendency that real-world lips have to adhere to each other toward their corners while opening the mouth [Osipa 2010]. This effect occurs while the jaw is moving in a downwards direction. In the initial moments of the downwards motion of the jaw, the upper and the lower lip regions stick to each other, not separating immediately, but then the lips slowly begin to release from each other as the downwards motion of the jaw increases and the mouth is more and more opened.

Figure 4.19 illustrates the rig process developed to handle the sticky lips effect. To generate the sticky lips effect a new extra layer of bones seen in parts 2 and 3 of Figure 4.19 is added to the existing skeletal structure seen in part 1 of Figure 4.19. This new extra layer involves a more complex bone setup because it deals with the deformation of the inner geometric portions of the lips. Initially, these regions need to adhere to each other and then gradually cease to adhere to each other, an effect which is achieved using three different bone groups described in detail following.

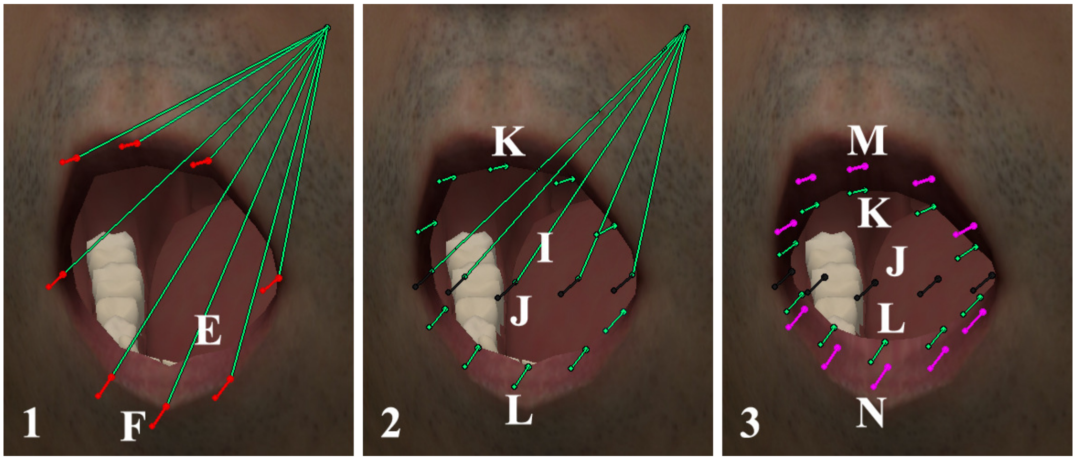


Figure 4.19: Rig mechanical setup for the sticky lips effect; left: bones layer for the lips; middle: layer with bones to control the most inner edge of the lips; right: layer with static bones that act as references for the inner edge deformation bones (effect applied to 2/3).

The bones in the first group are located along the center lip line (e.g. bones I and J). Bone J is parented to bone I, which is constrained to copy the rotations of the upper and lower jaw bones with half influence in each in order to average their movements and adapt to eventual mouth shapes being posed by the rotation of the upper or lower jaw bones. By mimicking the rotations of the jaw bones, the bones in the first group are kept centered in the mouth and in-between the top and bottom lips.

The second group of bones is located along the most inner edge of the lips (e.g. bones K and L). These bones are parented to the main deformation bones of the lips (e.g. bone L is parented to bone F) and they are responsible for the deformation of the inner portion of the lips. These bones are able to move but the user is not given direct control of them, since they operate automatically and do not need to be visible.

The third group of bones (e.g. bones M and N) share the same position and parenting of the bones in the second group, except they do not deform the lips geometry, instead they remain static in that area, being parented to their corresponding lip deformation bones, in order to only move along with the lip deformation bones.

The purpose of the first and last groups of bones is to act as references for the second group of bones. When the upper or lower jaw bones are moved to open the mouth, bones K and L gradually stop copying the transformations of bone J, located in the center lip line, and gradually begin copying the transformations of bones M and N, respectively. As a result, the upper and lower inner portions of the lips break apart from each other in respect to the mouth opening. In part 3 of Figure 4.19 the sticky lips effect is executed by two thirds of its total, with bones K and L getting closer to bones M and N, respectively.

Figure 4.20 illustrates the constraints and drivers assigned to bone K as a rig setup example which is repeated for the bones in the second group. The bone nomenclature seen in the constraints stack and drivers illustrated in Figure 4.20 coincides with the nomenclature used in the previous figures to facilitate the rig setup description. In the facial rig available in the Blender file relative to this section, the nomenclature used is more detailed, to more easily distinguish between the several components that compose the entire rig setup of the lips.

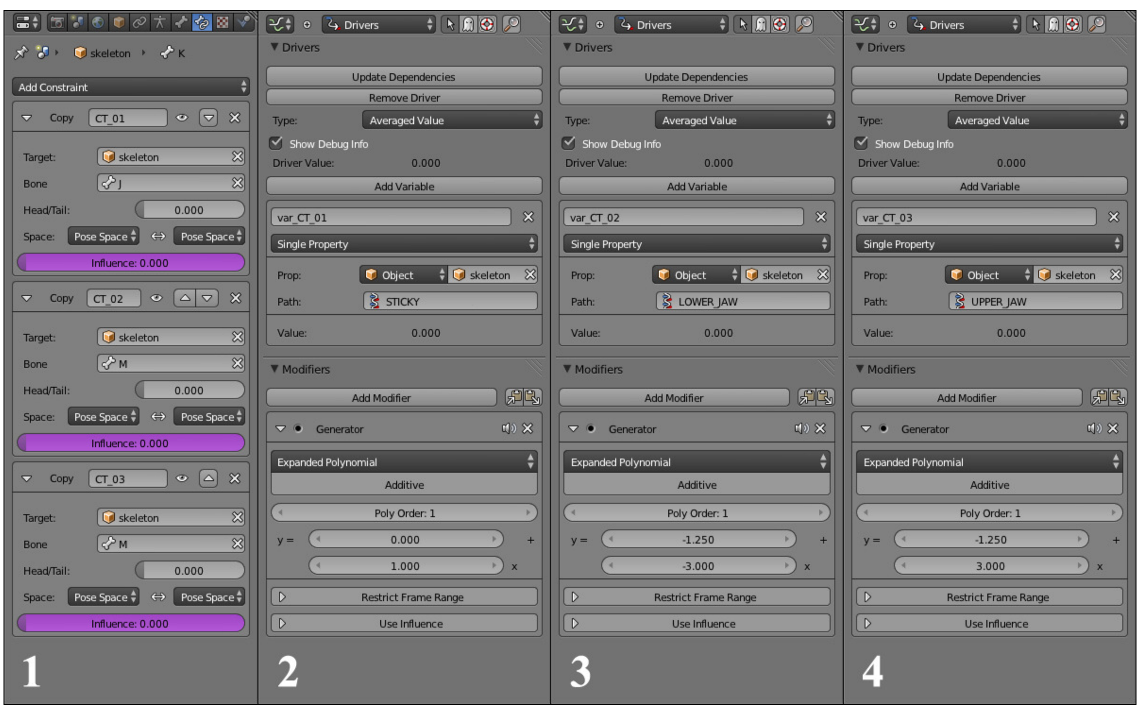


Figure 4.20: The setup of the constraints and drivers for the deformation bone K.

Three constraints assigned to bone K are seen in part 1 of Figure 4.20. Parts 2, 3 and 4 of Figure 4.20 respectively show the drivers assigned to the influence of each of the constraints in the constraint stack. The influence of a constraint is the intensity of its use, measured 0 to 1. The first constraint, in the top of the stack, copies the transformations of bone J (seen in part 2 of Figure 4.19). The influence of this constraint is driven by a control slider ranging from 0 to 1, here called STICKY (part 2 of Figure 4.20) and located in the Blender interface.

This slider allows the animator to control the amount that the sticky lips effect is applied. The coefficients of the expanded polynomial of this driver are set to 0.000 and 1.000 in order for the control slider range to correspond to the influence range of the constraint.

The second constraint, in the middle of the stack, copies the transformations of bone M (seen in part 3 of Figure 4.19). The influence of this constraint is driven by a control slider for the lower jaw, here called LOWER_JAW (seen in part 3 of Figure 4.20), ranging from 0 to -1 as the lower jaw moves downwards to open the mouth.

The third constraint, in the bottom of the stack, also copies the transformations of bone M and its influence is driven by a control slider for the upper jaw, here called UPPER_JAW (see part 4 of Figure 4.20). This slider ranges from 0 to 1 because the upper jaw moves upwards to open the mouth. Hence, the sticky lips effect is triggered by the user manipulating either the upper or the lower jaw controls. The coefficients of the expanded polynomials of the second and third drivers are set to -1.250 with -3.000 and -1.250 with 3.000 in order to provide more strength at first and less strength in the end. The reason for the inversion of -3.000 to 3.000 is the negative downwards motion of the control slider of the lower jaw, from 0 to -1, whereas the control slider of the upper jaw moves upwards positively, from 0 to 1. In order to achieve a more realistic sticky lips effect, the values of the coefficients of the drivers managing the constraints of the deformation bones of the inner portions of the lips are fine tuned to allow the deformation bones in the center region of the lips to cease copying bone J faster.

The setup described allows bone K to cease copying bone J and begin copying bone M as either the upper or the lower jaw is opened. Because Blender evaluates the constraints stack from top to bottom, the effects of the second and third constraints are only visible if the influence of the first constraint is different than zero. Hence, the sticky lips effect is only considered if the animator defines a sticky value, otherwise bone K continues to produce a regular inner lip motion by only copying the transformations of bone M. The setup described is then repeated for each bone responsible for the deformation of the inner portion of the lips.

4.2.4 User Evaluation: A Qualitative Experiment

To evaluate the zipper and sticky lips rig approaches developed, a qualitative international user experiment was conducted online with five expert users in character animation who are also professors of 3D in higher education institutions in Portugal and in Brazil.

Experiment Design

The users were asked to score the rig approaches of the zipper and sticky lips in a number of parameters according to their satisfaction with the optimization of the facial rig construction processes and given their professional and lecturing experience. The users are Gustavo Rosa of the Veiga de Almeida University of Rio de Janeiro (UVA/RJ), Nelson Gonçalves of the Superior School of Education of Viseu (ESEV/IPV), Nuno Estanqueiro of the Superior School of Social Communication of Lisbon (ESCS/IPL), Pedro Teixeira of the Superior School of Technology of Barcelos (EST/IPCA) and Ricardo Megre of the Arts School of the Catholic University of Porto (EA/UCP).

Experiment Stages

The evaluation was carried out individually in the following two stages, which are detailed ahead: (i) description and user hands-on and (ii) an evaluation questionnaire via an interview.

- *Stage 1: description and user hands-on*

In the first stage the users were initially presented with a video showing a series of facial rig control approaches collected from the state of the art in character facial rigging and animation. The purpose of this video was to introduce the topic and generate a compelling environment for the experiment. Following, a description of the sticky lips approach created by [Osipa 2010] was provided as a means of comparison for a final follow-up description of the author's approaches presented in Sections 4.2.1 to 4.2.3. The users were then invited to explore the facial rig directly in Blender in the author's machine via a remote control session using the software TeamViewer [TeamViewer 2014]. The purpose of this session was for the users to test the rig features and the rig construction workflow (user hands-on). This stage was carried out in a period of 60 minutes, of which 40 minutes were used for the author to describe the rig and 20 minutes for user hands-on.

- *Stage 2: evaluation questionnaire via interview*

In the second stage an evaluation questionnaire was conducted via an interview with no limit of time. Each user was asked to score the zipper and sticky lips rig setup process based on (i) their experience as professionals and professors, (ii) on the description provided by the author and (iii) on the user hands-on period that they have carried out. Users were told to feel free to provide any comments during their scorings, which were provided in a 0 to 10 scale according to a set of parameters presented following in Section 4.2.5.

4.2.5 Evaluation Parameters and Experiment Results

The parameters to conduct the user evaluation were devised based on usability engineering criteria by [Nielsen 1992], on the usability metrics by [Dix et al. 2004] and on the inherent characteristics of the facial rigging process. The parameters are *(i) easy construction and maintenance* (if the steps required for the rig setup process are easy to carry out, to keep track of and to fix or change if necessary), *(ii) integrability of features* (if the features of the rig are well integrated with each other and are suitable for being used with rigs relative to other facial regions, given that these approaches are focused on the region of the lips), *(iii) overall complexity of learning* (if the rig setup process is easy to learn [Nielsen 1992]), *(iv) easy to reproduce without errors* (if the rig setup process is easy to remember and repeat successfully even after a period of not using it [Nielsen 1992]), *(v) easy to extend to other facial styles* (potential of the rig to be applicable to other character faces like cartoons), *(vi) balance work/efficiency* (if the amount of work to build the rig justifies the visual deformation efficiency obtained), *(vii) compatibility off-line/real-time* (if the rig can be used for characters in both films and videogames, given that most game engines are prepared to support bone-driven deformations), *(viii) go cross-platform* (potential of the rig to be recreated in other 3D tools considering that it is based on rigging techniques available in the major 3D tools), *(ix) applicability to lecturing* (if the rig setup has potential to be lectured to experienced students in character animation as an advanced workshop topic in a classroom environment [Dix et al. 2004]) and *(x) applicability to the industry* (if the rig setup process has potential to be used as a facial rigging approach in production pipelines of companies working in the entertainment industry [Dix et al. 2004]).

The scores given by each user are presented in Table 4.19. The individual and averaged scores given by the users reveal the relevance of the rig approaches to help ease the facial rigging and animation of the zipper and sticky lips effects. An average score of 8,8 is given by the five users in a scale 0 to 10, with 43 in 50 scoring slots (86%) having scores equal or higher than 8. Only 7 slots (14%) have scores lower than 8, which are justified following.

Users 1 and 5 scored 7 in the parameter *(i) easy construction and maintenance*, based on the fact that the maintenance of a facial rig is a hard task with which expert digital artists are more capable to deal with. This score was justified by the users more as safety conduct than

Evaluation parameters	Expert Users Scores					Average per parameter
	U1	U2	U3	U4	U5	
(i) easy construction and maintenance	7	8	8	9	7	7,8
(ii) integrability of features	9	10	10	10	10	9,8
(iii) overall complexity of learning	9	10	9	10	8	8,8
(iv) easy to reproduce without errors	8	10	8	10	8	8,8
(v) easy to extend to other facial styles	10	7	8	9	10	8,8
(vi) balance work/efficiency	5	9	10	10	5	7,8
(vii) compatibility off-line/real-time	10	10	9	9	10	9,6
(viii) go cross-platform	9	10	9	10	10	9,6
(ix) applicability to lecturing	8	10	9	10	7	8,8
(x) applicability to the industry	7	9	9	8	8	8,2
Average per user	8,2	9,3	8,9	9,5	8,3	8,84\8,8

Table 4.19: Users self assessment criteria of the zipper and sticky lips rig approaches in a scale 0 to 10; left to right: column with evaluation parameters; columns with scores by each user; column with averages of the scores per each parameter; bottom row: average scores per user.

the cause of any sort of limitation that was found in the facial rig approaches.

User 2 scored 7 to the parameter *(v) easy to extend to other facial styles*, based on the fact that a larger number of lip controls might be required to allow an accurate transition from the human subject facial model to other facial anatomies. This score is also described as a safety conduct pointing to an eventual accessible adaptation of the facial rig approaches, given that the lips can adopt different shapes in more stylized characters (e.g. cartoon lips). As seen previously, the number of controls in the lips can be increased if necessary.

Users 1 and 5 scored 5 in the parameter *(vi) balance work/efficiency*, based on (i) the fact that the zipper and sticky lips effects are used in specific situations and (ii) that a faster solution (e.g. an automatic facial motion capture system) is sometimes a strategy that digital artists might prefer to reduce the amount of rigging work. However, motion capture solutions to track the lips remain prone to errors [Anderson et al. 2013]. Also, mocap equipment (i) is usually not cost free, unlike the approaches presented, which rely solely in the artist and in using the open-source tool Blender, (ii) the end results produced by motion captured lips can require extra fine tuning by the animators via key frame and (iii) an underlying optimized facial rig mechanical structure is required to guarantee a customized control of the lips and to provide extra lip deformation for cartoon lips, which are not feasible to be motion captured, "you could not mocap Shrek for example" [Orvalho 2007].

User 5 scored 7 in the parameter *(ix) applicability to lecturing*, based on the fact that there are a number of topics which can be lectured in advanced rigging. But the user also mentions that the implementation of the presented rig approaches as a specific workshop on facial rigging would definitely be more feasible and interesting to lecture.

User 1 scored 7 in the parameter *(x) applicability to the industry*, based on the fact that the limited production time in a company can cause these approaches to be considered only in case they are used more than once. In this regard one needs to consider that a faster solution can tend to provide acceptable rather than desirable end visual results.

As final remarks, the zipper and sticky lips approaches presented are considered easy because (i) they act as logical guideline proposals that help place and configure the facial rig mechanics for the zipper and sticky lips effects, a topic little documented in the state of the art, and (ii) they are validated according to the conditions described in Section 4.2.1 and in a qualitative international user validation that resulted in positive scores awarded by expert professional users who are also higher education professors in the field.

Despite the implementation of the lips rig approaches was carried out in Blender, the approaches are highly prone to be implemented in other 3D tools, since they are based in techniques available in the major 3D production tools (e.g. 3dsMax, Maya and Softimage).

Chapter 5

Easy Facial Rig Controls

The setup of the facial rig controls for a facial model is a laborious, time-consuming and disparate process. This chapter presents a facial rig control approach that lets animators personalize their control of a character's face. This approach is implemented for the human facial model built in Section 3.4 of Chapter 3. Aware of the fact that a control definition approach could not rely in setting rules, the opposite research direction became the goal: build a multimodal facial rig with a greater freedom of control. This hypothesis is based on providing multiple control inputs that breakthrough with the contingency of the rig control interaction paradigms found in the state of the art. The common rig control paradigms constrain animators due to being based on specific isolated designs, with benefits and counter backs perceived only by the artist who developed them and limited to their current design. A multimodal facial rig can be a relevant contribution for the awaited balance of the workflow between rigger and animator because it helps diminish the likeliness of unnecessary rig setups and unexpected manipulations by providing a more free manipulation. The validation of the multimodal approach was performed in an initial user experiment with 15 basic, skilled and expert digital artists and then in a final international user experiment online with 20 expert artists. In this chapter, the results presented in Section 5.1 are published in FAC^{2}RIG: A Complete and Customizable Rig Interface System Approach for Facial Animation [Bastos and Fernandes 2013] (see Appendix A-2013-2) and the results in Section 5.2 are published in Multimodal Facial Rigging: Towards a User-Adapted Animation Interface [Bastos et al. 2014a] (see Appendix A-2014).

5.1 FAC²RIG: A Complete and Customizable Rig Interface System Approach for Facial Animation

This section presents the development and the initial validation of a complete and user customizable rig interface system approach for facial animation. This approach is based on providing multiple control design inputs to help decrease the disparity of the facial rig control approaches found in the state of the art. The state of the art approaches restrain animators because they are built on the fly and are generally limited to one interaction mode (window, camera or viewport). This limitation is frustrating for animators because they are confined to a predefined interaction during their laborious job, in which usually the production of one second of animation requires about one hour of labor [Lewis and Anjyo 2010].

Because “there are no standards for the creation and definition of a facial rig” [Orvalho et al. 2012], then the use of rules in this field may not be prudent. Animation is a creative job [Thomas and Johnston 1981] in which, ideally, there should be little interaction constraints. Therefore, a question is conceived: would animators enjoy having a customized control of a character's face and would that ability improve their job? This hypothesis led to developing a system baptized FAC²RIG, a complete and user customizable rig interface design approach for character facial animation. The anticipated advantages of this system are (i) the decrease in the user frustration by accessing different control options and (ii) the increase in the precision of manipulation of a facial expression by doing an initial blocking out [Williams 2001] of the required expression using one input mode and refine it via another input mode. This interaction model aims towards a consensus for the facial rigging and animation process by attempting to optimize the diverse data common to a facial rig. This model lets animators choose the interaction mode that may be more appropriate to control each behavior in each facial region, as each mode is advantageous in different circumstances. The benefits of each interaction mode are described in detail ahead in Section 5.1.2.

The complete and user customizable interface approach was implemented in the popular open-source 3D tool Blender in order to (i) cause a significant impact in their growing community of artists, to (ii) leverage the use of Blender in the industry, (iii) due to the author's realization of the higher predisposal of the artists for validation (as seen in the preliminary result [Bastos et al. 2012a], described earlier in Section 3.3.2) and (iv) the fact that Blender is a qualified platform with an extended workflow. Still, FAC^{2}RIG's method can be applicable to other 3D tools, since its construction is based in the entire spectrum of the state of the art in this field [Orvalho et al. 2012, Bastos et al. 2012b]. It can also be used by animators with experience in other 3D tools (e.g. Maya), because Blender supports different input conventions. The system approach was tested by 15 digital artists (6 basic, 6 skilled and 3 experts) who evaluated it in terms of facilitating the manipulation of a character's face. An overview of the system is presented following. For a throughout verification of the facial rig controls developed in this section, please refer to the Blender file entitled *section_5.1*, available in the folder *sections blend files* in the accompanying CD.

5.1.1 System Overview and User Interaction Framework

FAC^{2}RIG is planned to benefit riggers and animators. For riggers it acts as a set of guideline proposals to create complete and accurate facial rig controls, to help the rigger build a more user friendly control rig. For animators it acts as a multimodal user interface customizable according to user preference to create complete and appealing character facial expressions. The anticipated advantages of this system – the decrease in user frustration and the increase in control precision – can help reduce the need for animators to request improvements to the rigging department, because there are more options of control available. Figure 5.1 illustrates the user interaction framework of FAC^{2}RIG, which merges the different interaction modes described in the state of the art to benefit from the advantages inherent to each mode.

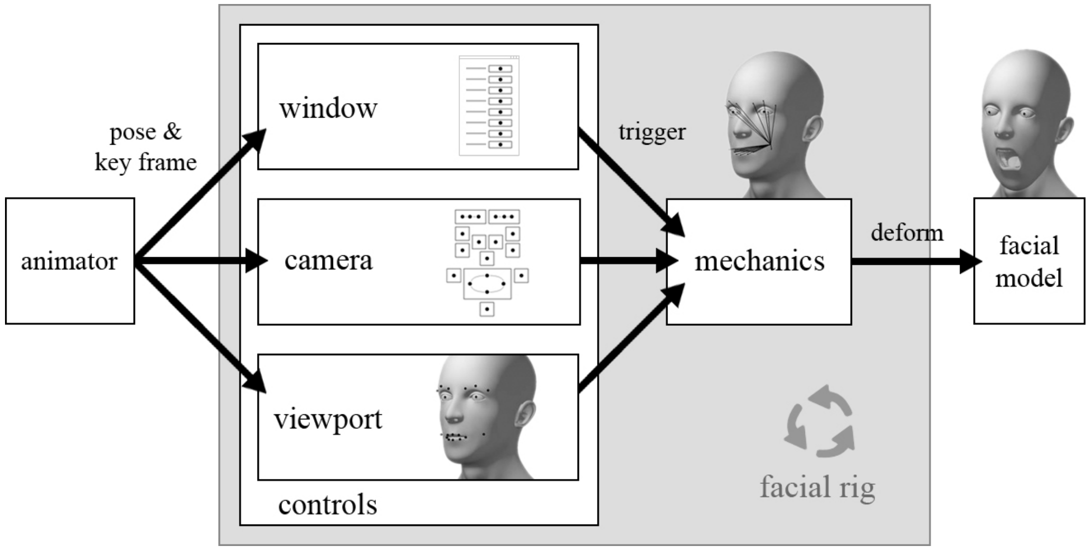


Figure 5.1: Overview of the FAC²RIG multimodal user interaction framework including the window, camera and viewport-based modes.

The user interaction framework seen in Figure 5.1 was created based on the hypothesis that multiple control input modes can decrease frustration in the manipulation and provide a more enjoyable user interaction. The illustrated model allows animators to pose and to key frame a character's face using the window (WIN), camera (CAM) and/or the viewport (VIP) interaction modes. Figure 5.2 illustrates how this framework is presented to the animator.

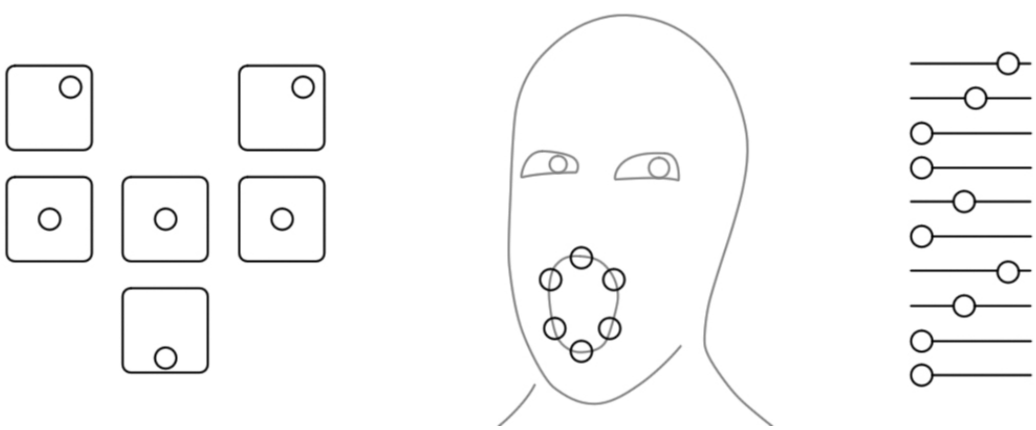


Figure 5.2: Visual representation of the user interaction framework of FAC²RIG; left, middle and right: respectively the positions on the screen of examples of the camera, viewport and window-based control handles.

The interface layout seen in Figure 5.2 is designed to prevent the cluttering of the visual space. The viewport interaction mode is displayed in the center of the screen to allow a straightforward perception of the facial model. The window mode is displayed to the right in the predefined properties window of the Blender interface, commonly used by the Blender artists to hold sliders (e.g. [Vegdahl 2010]). The camera mode is displayed to the left.

5.1.2 System Definition

The system is applied to the realistic human subject facial model built previously in Section 3.4 of Chapter 3. This facial model balances visual appeal with a fast deformation via the rig mechanics which were developed and described previously in Chapter 4. The facial behaviors are confined to the human face as a basis for other facial styles, because independently of the character face used, that face will always have anthropomorphic characteristics (e.g. number, symmetry and position of facial regions) in order to be recognizable by an audience as an expressive character. Even a cartoon character (e.g. with only one eyeball and a mouth) will display variations of the behaviors originally seen in humans, but with different proportions, often exaggerated to different degrees of freedom.

The different rig control interaction modes of the system are available per each facial region. Each control manipulator located in each interaction mode triggers the underlying optimized facial rig mechanics described in Chapter 4 that deform their corresponding facial regions in the 3D model. The different rig control interaction modes of the system are the window, camera and viewport-based modes. The user can manipulate either the window mode (1D), the camera mode (2D) or the viewport mode (3D) and receive visual feedback of the deformation in the 3D viewport where the facial model is displayed. The window mode is considered to provide a more exact interaction through a more precise input. The benefits of the window mode are (i) the ability for the user to set numerical values in sliders, (ii) it does not occupy the 3D space or the skin surface of the 3D facial model, (iii) it contains textual information and (iv) due to being located in the predefined windows of the 3D tool interface, it can be relocated to fit different types of windows. The camera mode is considered more intuitive as it (i) presents a schematic view of the controls, (ii) each osipa control component can accommodate more than one manipulator if necessary, (iii) it allows deformations independent of the angle of view of the 3D model and (iv) it eases the perception of control limits due to the manipulators being constrained to visual delimiters. The viewport mode is considered more straightforward, as it (i) presents more transformation freedom because a single gizmo can hold more behaviors, (ii) the deformation caused by the manipulation of a gizmo is directly acknowledged due to the gizmo being stationary on the surface of the facial model or nearby the facial region, (iii) the manipulation can feel more natural for the user due to the gizmo behavior being unconstrained, as objects tend to be in the real world, and (iv) in the 3D viewport the gizmos manipulation can be coupled with zooming in/out the model to perceive the purpose of a control and the resulting deformation.

Each interaction mode has advantages and disadvantages, for instance, it might be less efficient to clutter the camera and viewport modes with textual information, but the window mode may not be as direct and straightforward. It may also be more difficult to locate a given desired slider and memorize its function in the window mode, but sliders provide the option to adjust deformation by typing in values which can be set to range in different measures.

The system controls are implemented in Blender using bones, which are connected to the deformation bones via a network of drivers. The drivers average the modifications done by the user in each control knob and provide that information to the rig mechanics. An optimized naming convention was defined to easily distinguish between the deformation and the control bones (e.g. the left inner eyebrow deformation bone is *B_eyebrow_LI_DEF*, the respective camera and viewport control bones are *B_eyebrow_LI_CAM* and *B_eyebrow_LI_VIP*).

5.1.3 System Interface Design

The interface design of the system is seen in Figure 5.3. It is based on sliders, osipas and gizmos, respectively distributed in the different types of interaction modes: window (WIN), camera (CAM) and viewport (VIP). The sliders, osipas and gizmos control the different behaviors required for animators to accurately manipulate the human subject facial model. The controls were designed using Python modules *bpy.types.Panel*, *bpy.types.Operator* and *bpy.app.handlers* [Blender Foundation 2014b], respectively used for panel/sliders design and function, buttons design and function and interactive update functions upon user input.

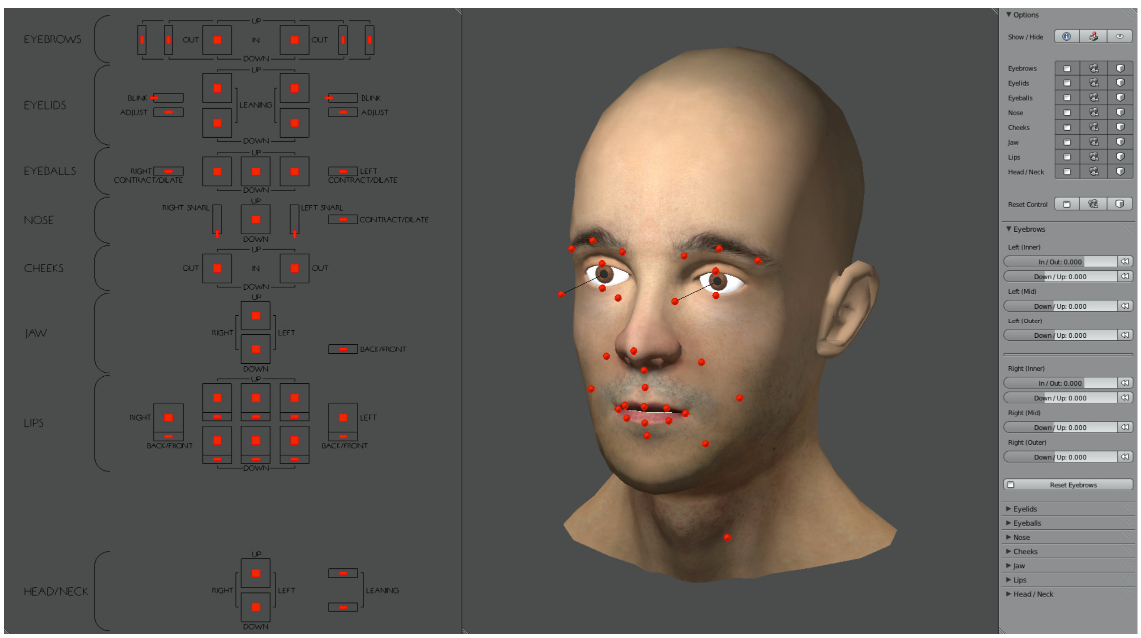


Figure 5.3: The FAC^{2}RIG control system approach implemented in Blender; left: camera mode; middle: viewport mode; right: window mode with a customization panel in the top.

The controls seen in Figure 5.3 in the camera and viewport modes are bones with custom shape designs assigned. The custom shapes are respectively squares and rectangles for the osipas and spheres for the gizmos. These custom shapes provide a visual simplification and help prevent an intentional preference of the user for a given control mode, as a result of the appeal of the 3D look of the manipulator [O'Neill 2008] instead of its function.

The different user interface interaction paradigms of the system seen in Figure 5.3 were designed per each facial region based on a combination of the feature points identified in the MPEG-4 standard [Pandzic and Forchheimer 2002] with the collection of the facial behaviors identified in FACS [Ekman and Friesen 1978]. Two other conditions were also considered: association and simplification. Association is relative to the user efficiently associating a given control manipulator with its corresponding behavior, via a pre-planned location of each osipa, gizmo and panel of sliders based on a top to bottom layout that corresponds to the top to bottom layout of the human subject facial regions. Simplification is designing the function of each control so it produces the deformation expected by the user (e.g. the translation of a control knob to the right results in a deformation motion of the facial skin to the right).

Figure 5.4 shows a specific example of the setup of the different interaction modes for the area of the left inner eyebrow region considering: (i) inclusion of the least possible number of controls keeping their necessary functionality and (ii) an organization layout based on the muscular activity of the human face, because the position and motion of each of the different control manipulators is mapped to correspond to the deformation of the target emotion to be expressed, therefore according to the position and motion direction of its corresponding behavior. As a result, when the user manipulates a given control in a direction, generally the corresponding facial skin deformation will also occur in that direction, independently of the interaction mode being used.

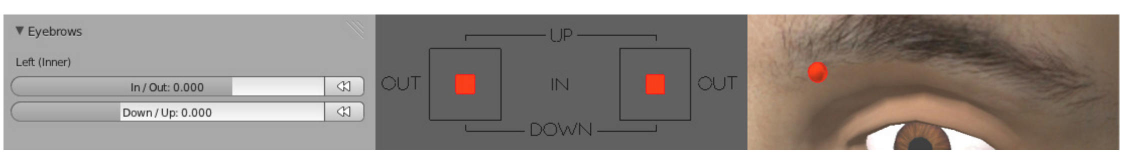


Figure 5.4: The design of the inner eyebrow control modes of the FAC^{2}RIG system approach in Blender; left: window mode based on sliders and buttons; middle: camera mode based on manipulators limited to square-shaped delimiters; right: gizmo-based viewport mode.

The goal of the approach seen in Figure 5.4 was to find a balance between the necessary facial behaviors and the user interaction design of the rig interface. Given this goal, each interaction mode was configured for each facial region as described following.

In the window-based control mode, each slider is built via property handlers in Python with textual information corresponding to its motion direction. To design the sliders in the window mode, ID properties were built inside Blender and then added to the existing script, instead of also being scripted. The reasons are that (i) ID properties are stored in the Blender file when the file is saved and (ii) they are easier to keep track of because they can be created for the different bones in the skeleton, therefore being visible on the Blender interface inside the custom properties panel of each control bone they belong to, instead of cluttering the Python code of the system. Buttons to reset sliders' values were built via operator classes and are located to the right of each slider, because generally the first intention of the user is to perform a deformation using a slider, followed by a reset event, only if necessary.

In the camera mode osipas are used. As mentioned in FACS [Ekman and Friesen 1978], the inner eyebrow can raise, lower and move to the inside or outside. In the FAC^{2}RIG design, the horizontal and vertical motion of the osipa manipulator (seen in Figure 5.4 in the center of each delimiter) has a correspondence with the horizontal and vertical motion of the inner eyebrow. This correspondence provides an optimized interaction because these motions are opposite and it is not correct for them to occur at the same time as they would null each other.

In the viewport mode, a single gizmo is located in the skin of the inner eyebrow region (see Figure 5.4). This gizmo can be moved only in the horizontal and vertical axis, since no other movements are necessary in this area of the eyebrow. A likely option could be to use this gizmo to also control the middle or the outer area of the eyebrow, but this option would not correspond to the separation by areas of the motion control within each facial region, as is outlined in FACS [Ekman and Friesen 1978].

5.1.4 System Interface Customization

A panel with options for customization of the interaction is provided in the top of the window mode, seen previously in Figure 5.3 under the name *Options* and illustrated in detail in Figure 5.5. The customization tools assure that the animator has a free control workflow and that the system does not get too overwhelming to deal with. These tools include, from left to right and top to bottom: (i) information visibility, to show or hide window and camera-based text, (ii) visibility of the manipulators of viewport only, for fast previewing of a deformation during manipulation to prevent visual cluttering of the 3D model, (iii) visibility of the entire set of rig control modes, useful for a rapid displaying or hiding in order to focus in a particular facial region with no need to turn off each manipulator one by one, making it easier to integrate with other controls that may be necessary to rig into other parts of the character's body, (iv) support for displaying multiple interaction modes together via toggle operator buttons, which help ease the visual recognition of an active interaction mode for each facial region and allow a rapid switch between the different modes if necessary, (v) reset buttons for each mode with the extended benefit of doing so in a per region basis for each facial region and in each interaction mode and (vi) a highly perceptible user interface design with eye-catching icons: an info question mark for text information display, a joystick for manipulator display in the viewport space, an eye to toggle the display of the combination of the different interaction modes per facial region, a panel for the window mode, a camera for the camera mode and a cube for the viewport mode. The interface also displays tooltips that describe the buttons' purposes when the mouse is paused over them.

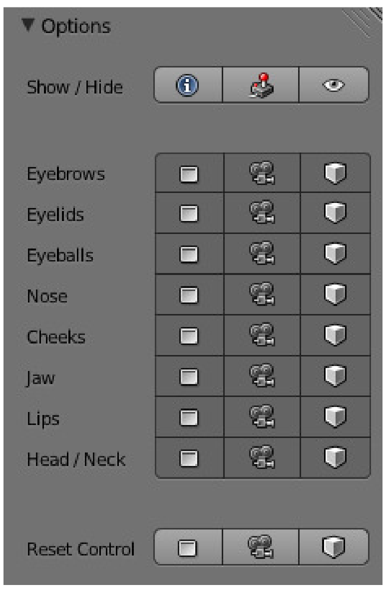


Figure 5.5: Options panel in the window mode for customization of the control interaction.

5.1.5 Facial Control Customization: A User Experiment

Is there any consensus in facial rigging and animation? The different interaction modes of the FAC^{2}RIG system approach are helpful in different circumstances? To answer these questions a user experiment with 15 digital artists was conducted in the 1st Blender PT Conference (*http://problender.pt/en/conf2013*) [Bastos et al. 2014b], held in the School of Arts of the Catholic University of Porto (*http://artes.ucp.pt*) in April 6-7, 2013. The goal of this experiment was to formally determine if there is an interaction mode that is most preferred by the users and how the system approach can be improved considering the user preferences.

Experiment Design

The experiment was performed with 15 users who are 3D digital artists, 6 with basic skills, 6 skilled and 3 experts. The users skills were perceived via a questionnaire (explained ahead in the experiment stages). Basic users have little or no experience in Blender and in facial rigging and animation. Skilled users are comfortable with Blender and somewhat experienced in facial rigging and animation. Expert users master Blender and reveal a solid experience in facial rigging and animation. The following three hypothesis have been formulated for the user experiment, which are evaluated in the experiment results shown ahead in this section.

1. FAC^{2}RIG lets users customize their interaction according to their preference and independently of their experience and also drives users to produce more appealing results with a heightened performance;
2. The different interaction modes in FAC^{2}RIG have benefits and counter backs, thus the user is able to produce better results when coupling the different modes rather than focusing in a single mode;
3. Users may reveal a higher preference for the viewport mode, be highly discriminative of the window mode and use the camera mode only in particular occasions.

Experiment Stages

The experiment occurred in two stages: (i) a group training stage to get users familiarized with the system followed by (ii) an individual tasking stage to produce facial poses using the system. Each stage is detailed following.

- *Stage 1: Training Stage*

In the training stage the users were gathered in a room where each user had a computer to practice the FAC²RIG system (see Figure 5.6). This stage had a total duration of 30 minutes and was divided in the following moments: (i) system explanation and users hands-on (20 minutes) and (ii) consent sheet and skills questionnaire (10 minutes). In the first moment the system was projected in a large screen and briefly explained, after which each user trained the system freely. The users were encouraged to practice with the different interaction modes and they were clarified of any questions relative to handling the system. In the second moment each user signed an experiment participation consent sheet (see Appendix B) and completed a questionnaire to assess their level of expertise (see Appendix C).The users level of expertise was also perceived during their performance in the tasking stage.

Figure 5.6: Training stage of the FAC²RIG system approach; left: digital artists training the system in the training room; right: 3D digital artist José Ricardo Silva performing tasks with the system.

- *Stage 2: Tasking Stage*

In the tasking stage each user participated in an individual session to simulate a number of facial poses (see Figure 5.6). This stage occurred in a separate room that was pre-arranged with a workstation running the system. A laptop computer displayed the facial poses which were requested for the user to simulate within a limited period of time, of which the user was aware. This stage had a total duration of 12 minutes per each user and was divided into the following moments: (i) choose your interaction mode per each facial region (4 minutes), (ii) customize your overall facial interaction (4 minutes) and (iii) participate in an evaluation questionnaire via an interview (4 minutes). In the first moment, the user simulated 1 pose in 30 seconds per each facial region (eyebrows, eyelids, eyeballs, nose, cheeks, jaw, lips and head), totalizing 8 poses in 4 minutes. For each pose the user had a pre-arranged Blender file with the required facial region slots pre-activated (the slots are seen previously in Figure 5.5). The pre-activation of the facial region slots allowed the user to focus in each facial region at a time, leading the user to decide within 30 seconds which interaction mode to use amongst the three interaction modes available. Users were only told that they were free to use the system as they prefer, as long as they achieved the desired expressions in 30 seconds per each facial region. In the second moment, the user simulated one final global pose which involved controlling each of the 8 facial regions in a total time of 4 minutes, equal to the overall time used previously for the set of 8 isolated poses. A Blender file was pre-arranged with all the regions slots deactivated in order to provide a restart of the interaction. The final global pose requested to the users is the visual result of merging each of the previous individual poses for each facial region. This aspect was not mentioned to the users to avoid leading them in reusing any control options they might have used in the first moment. Lastly, in the third moment, each user was asked to score the system according to a number of topics using a 0 to 10 scale and also provide comments. These topics are presented further ahead in Table 5.1.

Experiment Results

To answer the formulated hypothesis and evaluate the users performance and enjoyment, the collected data was studied considering: (i) the visual accuracy of the poses, (ii) the number of clicks performed in each interaction mode and (iii) the overall satisfaction of the users.

In terms of visual accuracy, it was verified that the users in general were able to pose the character's face according to the original pose model. A selection of the most precise poses is illustrated in Figure 5.7 for one basic, one skilled and one expert user per each facial region and also for the final global pose. Despite each user is different in skill, when the selected poses are compared with the original poses seen in the first row of Figure 5.7, a significant

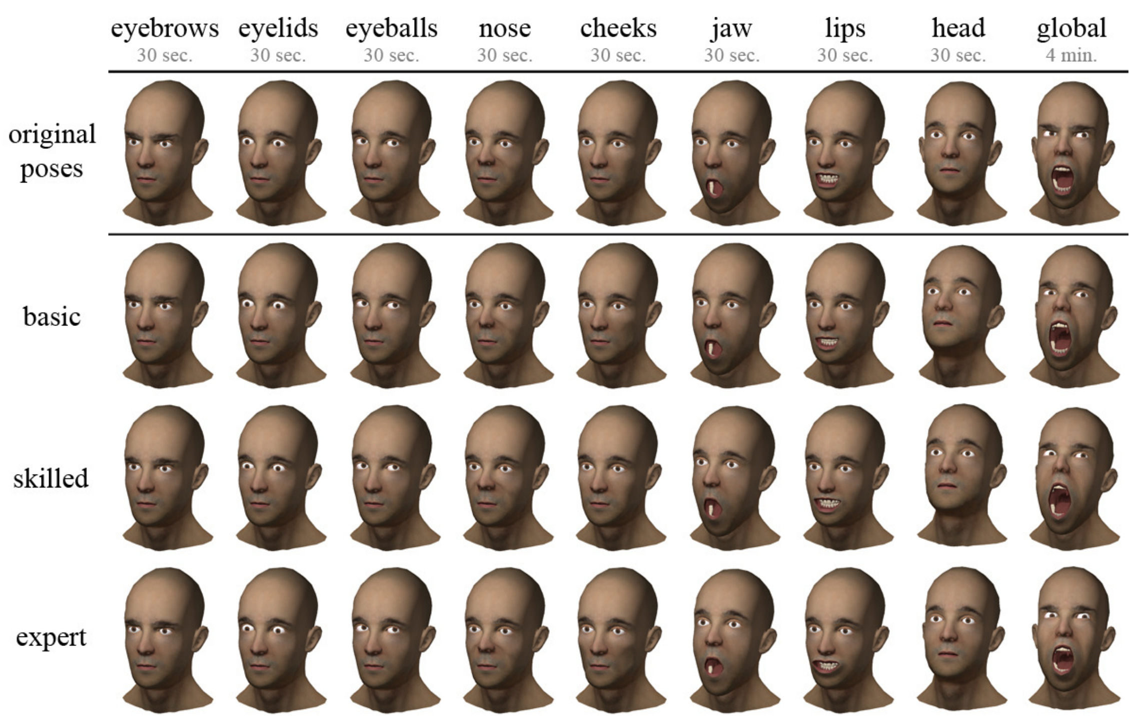


Figure 5.7: Visual accuracy of the facial poses produced by a basic, a skilled and an expert user selected amongst the group of 15 users and in comparison to the original poses required for users to simulate; top to bottom rows: respectively the original poses and the poses obtained by the basic, the skilled and the expert user per each facial region.

visual similarity is perceptible. This visual similarity demonstrates that the system contributes to average and improve the users performance because the users were able to pose accurate expressions for each facial region and also for the global pose, which combines the motions in each facial region. In general the 15 users managed to produce appealing results with a heightened performance, with the non-expert users producing poses similar quality-wise to those produced by the experts.

The evaluation of the number of clicks is also revealing. The first moment of the task stage allowed to acknowledge if users prefer a given mode over another based on their first decision. The second moment allowed to realize if in fact there were any variations from the control mode used in the first moment, and if users actually prefer a single mode or they perform sequential tweaks using other modes. A graphic showing the number of clicks made by the 15 users is presented following in Figure 5.8. This graphic also presents a comparison of the users preferences in each interaction mode and in both moments of the task stage.

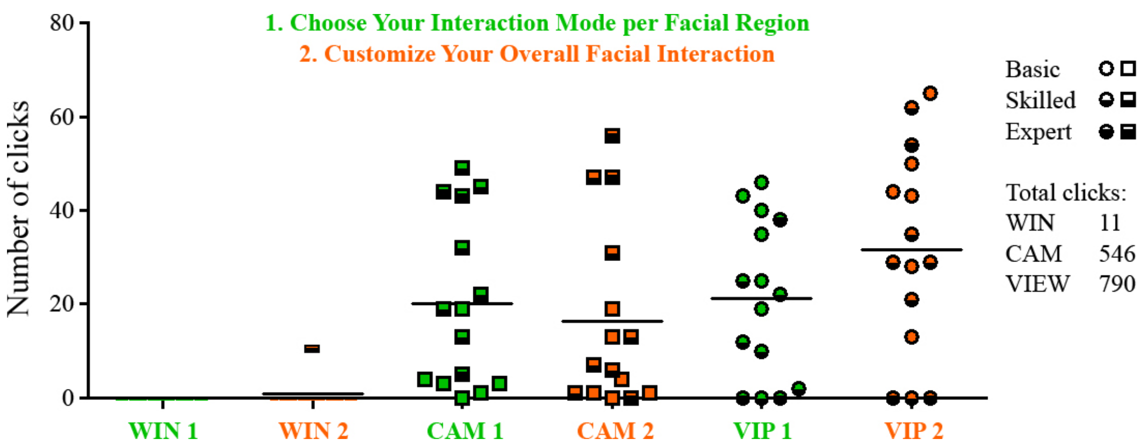


Figure 5.8: Graphical representation of the number of clicks performed by the group of 15 digital artists in each interaction mode and in each moment of the task stage.

Figure 5.8 provides a number of conclusions for the hypotheses highlighted previously in the experiment design. Within hypothesis (i), in terms of the interaction mode preferred, there is a distinction from the 1st to the 2nd moment of the task stage, in which users tended to use more either the camera (CAM) or viewport (VIP) in the 1st moment and in the 2nd moment that tendency was inverted. This inversion of tendency in the users interaction reveals the relevance of providing access to the different interaction modes, to let users customize their interaction according to their preference during the posing of the facial expressions.

Within hypothesis (ii), two of the three expert users have relied exclusively on the camera mode, with one of these users even opting to hide the other interaction modes in order to focus in the camera mode alone. The third expert user was more keen of the viewport mode and only used a limited number of times the window and camera modes, but did not exclude these. The basic and skilled users coupled the viewport and the camera modes more often and obtained likable visual results, which reveals that the access to the different interaction modes is beneficial in a per user basis as well as for different users.

Within hypothesis (iii), users in general had a somewhat stronger preference for the viewport mode, but closely followed by the camera mode and with little use of the window mode. According to the users feedback, the camera mode was also highly considered due to its schematic design. Also relevant is the fact that there was a reduced discrepancy of use from the 1st to the 2nd moment of the task stage in the camera mode when compared to the discrepancy of use between the two moments in the interaction with the viewport mode.

Table 5.1 shows the classifications given to the system in a scale 0 to 10 by the selected basic, skilled and expert users, it also shows their average scores per parameter, the average scores per parameter of the 15 users and the total average scores (in the last row of the table).

The criteria used for the classifications is based in [Dix et al. 2004] and is focused on the (i) overall interaction of the different modes, (ii) the users adaptability to the system, (iii) the availability of advanced options, (iv) the easiness of operation (learnability), (v) the system reliability (error tolerance), (vi) the option to access different controls, (vii) the ability to customize the controls and (viii) the pertinence of the system going cross-platform.

Evaluation criteria		**Basic**	**Skilled**	**Expert**	**Users average scores per parameter**	
					3 users	**15 users**
Overall Interaction	WIN	3	6	10	6,3	4,2
	CAM	8	9	10	9	7,8
	VIP	10	10	10	10	8,9
Adaptability		9	7	10	8,6	8,3
Advanced options		10	10	10	10	8,3
Learnability		10	9	10	9,6	8,8
Error tolerance		8	10	10	9,3	8
Different controls		9	10	10	9,6	8,4
Controls customization		10	10	10	10	9,4
Go cross platform		8	10	10	9,3	9,1
Users total average scores		**8,5**	**9,1**	**10**	**9,75**	**8**

Table 5.1: Users self assessment criteria of the FAC^{2}RIG system approach in a scale 0 to 10; left to right: column with the evaluation criteria; columns with scores by a basic, a skilled and an expert user selected amongst the group of 15 users; column with the average of the scores by the former 3 users; column with the average of the scores by the group of 15 users.

Despite it is verified in Figure 5.8 that the users neglected the use of the window mode, they classify it as significant, in particular to insert numerical values, with a resulting score average of 6,3 by the 3 users and 4,2 by the 15 users, as seen in Table 5.1. In the same way, the expert users who resorted exclusively to the camera mode have also provided high classifications to the viewport and window modes. They mentioned that the camera mode feels more rapid to them, which in their opinion does not mean that the other modes are not necessary. In fact, it has been realized that the users generally started the tests with the intention to use a given control mode and eventually realized the advantages of coupling the different modes. All users classified positively the access to the different control modes and the ability to configure them. The users described the system as “very interesting”, “very easy to use" and "able to achieve fast results", with one user mentioning: "the option of having a more intuitive interface (camera), a more easy interface (viewport) and a more rigorous interface (window) allows to please various needs".

As a final remark, it was also verified that the skilled users are more prone to couple the different modes. The other users tend to focus more in one of the control modes for a specific task and then call upon another interaction mode to achieve a different task (e.g. a per facial region workflow versus an overall facial interaction). The fact that the users are able to focus their interaction in a single mode and then also couple the different modes means that they are able to complete sequential tasks one after another. It also means that when necessary, the users are also able to mix the different modes to better control the nuances of the facial skin deformation in the different facial regions. Neither basic, skilled or expert users experienced problems either in a more focused interaction or when coupling the different modes.

5.1.6 FAC^{2}RIG Benefits and Limitations

FAC^{2}RIG advances the state of the art in character facial rigging and animation based on the fact that riggers can benefit from an optimized methodology to build the controls in a facial rig, which results in providing customizable rig controls that adapt to the preferences of the animators. The user experiment results revealed that the system advances the state of the art in facial rigging and animation by merging the WIN, CAM and VIP interaction modes to provide an improved user performance. However, two limitations were identified: (i) that the system averaged the inputs in the different interaction modes and instead should synchronize them to provide an updated visual feedback of the controls and (ii) that the system could also provide a parameter to exaggerate the facial behavior, to achieve more flexibility of control. These limitations were corrected and tested in a follow-up validation. The upgraded version of the system together with a final user experiment is presented following in Section 5.2.

5.2 Towards a User-Adapted Animation Interface via Multimodal Facial Rigging

The initial version of the FAC^{2}RIG system, presented in Section 5.1, was developed with the purpose to help provide a uniform integrative of control interaction to facial animation. The user experiment with 15 digital artists revealed necessary improvements to the system. This section describes the development and validation of an upgraded version of the FAC^{2}RIG interface system which is oriented towards a more user-adapted interaction. The user-adapted interaction is more efficient because it resolves the limitations found previously based on new features: (i) a synchronized input between its different interaction modes and (ii) a scalable control of the facial skin deformation. The approach continues to be a hybrid user interface that incorporates different types of rig controls within a single facial animation system, but it now allows the users to seamlessly switch between the controls in the different user interface modes with extra control provided over the elasticity of the facial skin deformation.

The following sections present the algorithm methods that were developed to support the user-adapted features of the upgraded version of the FAC^{2}RIG system and also a final user experiment with twenty expert digital artists who are long-term professionals in character rigging and animation. The evaluation parameters of the experiment focused on the added value of the user-adapted system approach in comparison to conventional interaction modes, based on the time required to pose facial expressions, the precision of the interaction and the users' satisfaction. The results of the time and precision tests, together with the users' scores, help clarify the relevance of the user-adapted multimodal rig interface system approach as an interaction model that decreases the effort required to rig and animate a character's face, making this process more efficient and enjoyable for the user. For a throughout verification of the facial rig controls developed in this section, please refer to the Blender file entitled *section_5.2*, available in the folder *sections blend files* in the accompanying CD.

5.2.1 A User-Adapted FAC^{2}RIG System

This section presents the design, workflow and implementation of the algorithms that correct the limitations identified in the FAC^{2}RIG system. The user-adapted multimodal approach includes (i) a synchronized control of the different interaction modes to let animators personalize their interaction avoiding extrapolations of the user manipulation and (ii) the exaggeration of the facial deformation to (a) overcome the difficulty of the rigger in building a facial rig that completely satisfies the animator in terms of the facial deformation length and (b) overcome the difficulty of the animator in achieving a greater flexibility of control. The algorithms also provide (i) a correct deformation independently of the control knob being manipulated, (ii) a range from -1 to 1 for the window (WIN) controls adapted to the camera (CAM) and viewport (VIP) control ranges, (iii) the ability to drag and scale the CAM controls to provide an enhanced user perception, (iv) the ability to display the WIN controls in a per facial region basis when their corresponding CAM or VIP controls are selected, to prevent cluttering the space in the WIN mode and (v) the ability of the algorithms to deal with the upper and lower parts of the face, with the user experiment including the eyebrows, the upper and lower jaw, the lips and the head as an example for the remaining facial regions.

Algorithms Design

The flowchart design seen ahead in Figure 5.9 overviews the cyclic process of the algorithm methods. The algorithms run continuously to accept user input events in one of the three interaction modes. The camera and viewport controls (respectively indicated as CAM and VIP) can be selected by the user and if so they can be used via a translation in space and their corresponding WIN controls are displayed. The window controls (indicated as WIN) cannot be selected since they are sliders. They are either displayed or hidden and, if displayed, they can be manipulated by dragging the slider knob. The *select WIN* box seen in Figure 5.9 is used to represent the clearing of the selection of the CAM or VIP controls in 3D space by selecting secondary 3D space controls. The secondary controls are then checked for being selected and, in case they are, the WIN slider is made available for manipulation, while also keeping the CAM and VIP controls unselected and able to be manipulated again if necessary.

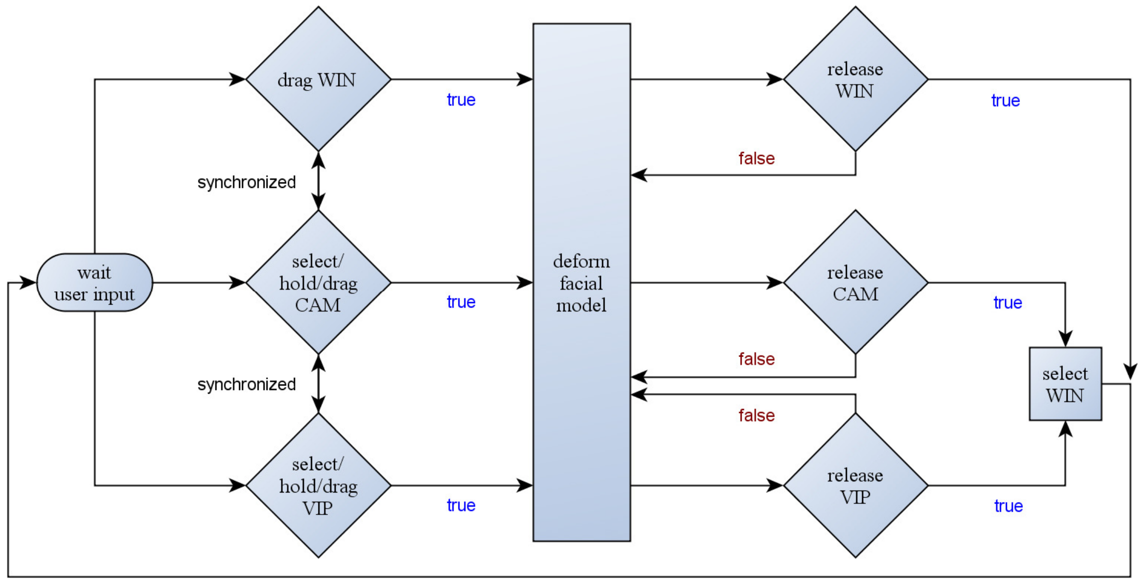


Figure 5.9: Flowchart of the user-adapted interaction for the FAC²RIG system approach; left to right: user inputs are permanently accepted in one of the 3 interaction modes to deform the facial model, in case a control is manipulated the other controls mimic the one being handled, upon release of the current control the system restarts the acceptance of input.

Algorithms Workflow

The algorithm methods deal with a number of components of the facial rig (e.g. skeletons, blendshapes and constraints). To assure their proper debugging, a naming convention was defined. Figure 5.10 highlights a single facial behavior in the human subject facial model as an example: the inwards and outwards motion of the left inner eyebrow region [Ekman and Friesen 1978]. The control knobs that allow triggering this behavior are code-named as *librow_WIN* (D1), *librow_CAM* (A1) and *librow_VIP* (B1). The term *librow* is a short and efficient way to rapidly refer to the left inner eyebrow area of the human subject facial model.

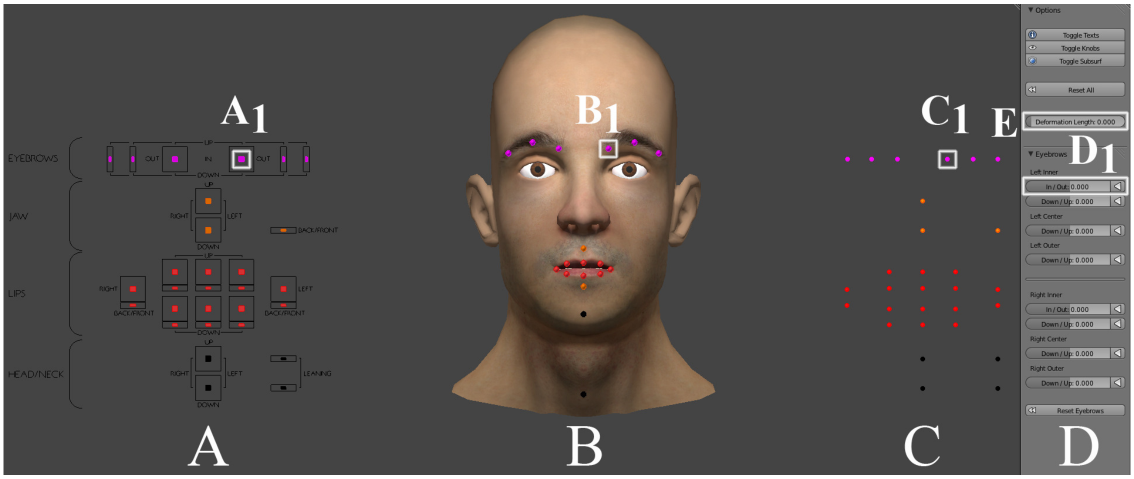


Figure 5.10: Workflow of the user-adapted algorithm for the inner eyebrow controls; left to right: A) camera, B) viewport, C) spatial selectors, D) window mode.

Section A of Figure 5.10 holds the CAM controls, which are composed by osipas and horizontal and vertical sliders. Section B holds the VIP controls, composed by gizmos placed on the surface of the facial model skin. Then section C holds spatial selectors which include the ID properties that represent the WIN controls. Their position in space is not relevant as they are hidden from the user in the running version of the system. Nonetheless, their layout matches that of the CAM and VIP in order to provide to the rigger an easier tracking and correspondence for debugging. Section D holds the WIN controls, which are composed by horizontal numerical sliders and a reset button located to the right of each slider to set the value of the manipulation to its default.

The WIN controls for each facial region are only displayed in case one of the CAM or VIP controls has been or is being used. A1, B1 and D1 are relative to, respectively, the CAM, VIP and WIN control knobs for the left inner eyebrow. C1 is the spatial commuter holding the ID properties for D1, which allows D1 to be synchronized with A1 and B1. In case the user triggers the A1 control knob, then the B1 and D1 control knobs will mimic the A1 control knob. In case the user triggers the B1 control knob, then the A1 and D1 control knobs will mimic the B1 control knob. After triggering either the A1 or B1 control knobs, the system automatically selects the corresponding spatial commuter C1, so that in case the user triggers the D1 control knob, then the A1 and B1 control knobs will mimic C1, which holds the ID properties for the D1 control knob.

The CAM and VIP control knobs are bones which are part of the deformation skeleton used to trigger the behaviors of the human subject facial model. In Blender, bones can also be used as controls with custom shape designs assigned to them while operating in the pose mode of a skeleton, which is the mode that Blender uses to record animation in a skeleton. The WIN controls are custom float properties which are added to the spatial bones seen in section C of Figure 5.10. These bones serve as commuters, meaning that they are selected automatically by a parallel modal operator event when it is detected that either the CAM or VIP control knobs are released by the user. This release event allows to clear the selection of either the CAM or VIP control knobs to let any other control knobs be manipulated by the user, making the synchronization between the different controls possible.

Another property indicated in Figure 5.10 as E allows the user to change the deformation length of the facial behaviors to add flexibility, to achieve cartoon-like deformations. It is available in the main menu called *Options*, located above the WIN controls. It ranges from 0 to 1, where 0 is considered the human deformation length (or the base deformation length) and 1 is considered the maximum increase value to the base deformation. Values tending to 1 will provide a larger deformation length for the control knobs in each facial region.

Despite the WIN, CAM and VIP control modes are available to the animator, only the controls in one mode are required to trigger the underlying rig mechanics that deform the human subject facial model. The controls in the VIP mode were used for this purpose due to being located on the surface of the model, because their closer spatial proximity with the underlying rig mechanics makes it easier for the rigger to debug the facial rig.

Algorithms Implementation: Technical Obstacles

Coding the algorithm methods bared the following technical obstacles: (i) dependency cycles and (ii) the inherent variation of the facial deformation length depending on facial region. The obstacle of dependency cycles, exemplified in software development by [Seemann 2011], originates in the present multimodal system from the need for each control knob to mimic its corresponding control knobs in the other interaction modes and vice-versa without causing a dependency cycle with invalid spatial transformations. Dependency cycles results from a control copying the location of another control and vice-versa, a problem that could not be resolved using regular constraint or driver assignments because the dependency maintains. Instead, Python scripting was used inside Blender to allow each control knob to mimic the reaction of its neighbor control knob only when the neighbor is selected. Hence, each control knob provides its current location to its corresponding control knobs so they can mimic the former while also letting each knob maintain its own transformation properties.

The inherent variation of the facial deformation length is caused by the difference in the motion of the facial skin in each facial region (e.g. the length of deformation of the lower jaw differs from that of the eyebrows). The inherent variation is resolved using a general length parameter to modify the translation limits of the VIP control knobs. These knobs are assigned constraints with minimum and maximum values for their translation limits. These values are then increased according to the general length integer property, provided as a slider parameter called *Deformation Length*, available in the main menu of the control interface.

5.2.2 Algorithm Method for Drawing Menus in the Blender Interface

Four algorithms were developed and are presented in pseudo-code in Sections 5.2.2 to 5.2.5 and in Python in Appendix D. Table 5.2 shows the first algorithm to implement the left inner eyebrow control menu in the Blender UI, as an example extendable to other facial regions.

draw menu options in Blender user interface
    define menu layout as single column
        draw length slider
draw menu brows in Blender user interface
    define menu layout as single column
        draw librow slider **in** left column
        draw librow slider reset button **in** right column
define function reset librow control
    reset value librow WIN slider
    reset translation librow CAM knob
    reset translation librow VIP knob
define function display menu brows
    if librow WIN knob **or** librow CAM knob **or** librow VIP knob **is** selected
        display menu brows
    else
        hide menu brows

Table 5.2: Algorithm method 1: drawing menus in the Blender interface.

The first algorithm method draws two menus in the Blender interface. The first menu has a single column layout holding the slider parameter which allows the user to set the value of the facial deformation length. If required this menu can later include other options, such as buttons to display or hide controls. The second menu holds the WIN controls and is drawn below the first menu with a double column layout. The column in the left holds a slider controlling the upwards and downwards movement of the left inner eyebrow (*librow*) of the facial model. The column in the right holds a button to reset the value of its slider. More sliders can be added here for the different parts of the left and right eyebrows. After drawing the menus a function is defined to reset the left inner eyebrow control, which is called upon pressing the reset button. Lastly, a function is defined to handle the display of the WIN menu of the eyebrows, which hides the menu in case another control not relative to the eyebrows is selected by the user. As a result, the space available in the WIN menu is prevented from being cluttered, since each facial region can easily reach a large number of sliders and buttons. Using this method, in case it would be necessary to add controls for another facial region (e.g. the jaw), it would only be necessary to draw a new menu relative to the jaw and create a new display function for the new jaw controls via the same procedure used for the left inner eyebrow.

5.2.3 Algorithm Method for Synchronized and Scalable Deformation Control

Table 5.3 shows the second algorithm method to implement the synchronized and scalable deformation control of the left inner eyebrow, as an example extendable to other facial regions.

```
variable scale = maximum transformation value / value length slider
define function synchronized and scalable librow controls
    if librow WIN knob is selected
        translation librow CAM knob = value librow WIN slider
        translation librow VIP knob = value librow WIN slider / scale
    if librow CAM knob is selected
        value librow WIN slider = translation librow VIP knob / value length slider
        translation librow VIP knob = translation librow CAM knob * value length slider
    if librow VIP knob is selected
        value librow WIN slider = translation librow VIP knob / value length slider
        translation librow CAM knob = translation librow VIP knob * value length slider
define function brows transform constraint
    transform limits librow VIP knob = value length slider
```

Table 5.3: Algorithm method 2: synchronized and scalable deformation control.

The second algorithm implements the synchronized and scalable deformation control. It divides the maximum transformation value of the WIN and CAM knobs by the value defined by the user in the length slider parameter, to then consider the result in the selection checks performed ahead. Due to the VIP knob being located on the surface of the model, it can translate further or less than the WIN or CAM knobs, which are limited to range from -1 to 1. Given the scale of the subject facial model, the transform limits of the VIP knobs range from -0.3 to 0.3, to be confined to a human-like deformation. These limits can be adjusted by the rigger depending on the morphology of the facial model and its scale in the 3D scene. In case the user increases the deformation scale, the algorithm updates the limits of the VIP knobs and adjusts the WIN and CAM knobs accordingly. In case of the upwards and downwards motion of the left inner eyebrow, if the WIN knob is selected, the CAM knob is equaled to it, whereas the VIP knob is equaled to it and also divided by the result of the division of 1 with the length set by the user. For instance, if the *librow* WIN control value is 0.3, it is divided by the current length value (e.g. 3.3), resulting in an approximate value of 0.1, which is the equivalent in the VIP mode to 0.3 in the WIN mode, because in this case the VIP control is limited to a real 0.3 maximum transformation value in the Z axis. In the same way, if the *librow* CAM control is selected, the WIN control value will correspond to the VIP control value and the VIP control value will correspond to the CAM control value. If the VIP control is selected the WIN and CAM controls will correspond to the VIP control. Lastly, a function limits the transforms of the *librow* VIP knob to what is defined in the length slider parameter. In case controls would be required for another facial region (e.g. the jaw), it is only necessary to duplicate and adapt the previous functions to the WIN, CAM and VIP jaw controls.

5.2.4 Algorithm Method for Automatic Modal Operation

Table 5.4 includes the third algorithm method. It explains the implementation of a modal operator to handle the automatic selection of the different controls of the left inner eyebrow of the facial model, as an example extendable to other facial regions. The third algorithm method builds an automatic modal operation to detect clicks of the left mouse button (LMB) on

```
define function constant running modal event
    if ESC key is pressed
        cancel modal event
    if librow CAM knob or librow VIP knob is selected and dragged using left mouse button
        pass-through
define function timer event
    if left mouse button is released
        clear selections
        select librow WIN knob
```

Table 5.4: Algorithm method 3: automatic modal operation.

the CAM and VIP control knobs of the left inner eyebrow. After selecting and dragging one of these knobs with LMB, a release must occur, which is when the algorithm selects the WIN knob. This waiting process for release events runs continuously in parallel, with the movement of the mouse also being tracked. In case a selection and drag occurs of either CAM or VIP librow control knobs, the process runs without affecting normal Blender behavior. A timer function then checks if LMB has been released. In case it has, (i) the current selection is cleared and (ii) the librow WIN control is selected. Using this method, the extension to other facial regions (e.g. the jaw) only requires a duplication and adaptation to the jaw controls of the if statement relative to the selection and release action of the librow.

5.2.5 Algorithm Method for Multimodal Interaction Extension Management

Table 5.5 includes the fourth algorithm method. It allows the rigger to segment the testing of the code to one or more facial regions, with the benefit to prevent continuous parallel handler events, thus reducing the computing power to not be cumbersome for the rigger. The user interaction for each facial region is programmed in its own script so that the rigger can set independent variables to call each script separately, choosing the scripts to run by calling only their variables for testing.

```
define function to hold facial region scripts
    variable menus = script menus
    variable modal = script modal
    variable brows = script brows
    variable jaw = script jaw
    (...)
```

Table 5.5: Algorithm method 4: multimodal interaction extension management.

5.2.6 Multimodal Facial Rig Control: An International User Experiment

This section presents a final international user experiment to evaluate the user-adapted FAC^{2}RIG system approach. The user experiment was carried out online via TeamViewer [TeamViewer 2014] exclusively with twenty expert 3D digital artists who are long-term professionals in character rigging and animation. The goal was to formally determine if the system eases character facial animation via a synchronized and scalable control of the human subject facial model, which adapts to the interaction preferences of the user.

Experiment Design

The experiment was designed to establish a comparison between the conventional key frame facial animation techniques and the multimodal key frame animation system. The algorithm methods presented in Sections 5.2.2 to 5.2.5 were applied to the upper and lower parts of the face of the human subject facial model, namely to the eyebrows, jaw, lips and head, as an example extendable to other facial regions. The following hypotheses have been formulated for the experiment.

1. A multimodal synchronized control increases precision and reduces the time needed for facial animation.
2. A parameter for extra deformation allows artists a flexible exaggeration of facial poses.

To test hypothesis (i), a sequence of instructions based on the guidelines by [Ekman and Friesen 1975] were provided to reproduce the human facial expressions of anger, fear and surprise, which are illustrated in Figure 5.11.

Figure 5.11: The human facial expressions to simulate in the user experiment; left, middle and right: respectively the anger, fear and surprise expressions applied to the human subject facial model.

To test hypothesis (ii), the cartoon expressions of anger, fear and surprise made by characters Sulley (Monsters Inc., Pixar 2001), Wodody (Toy Story, Pixar 1995) and Shrek (Shrek, DreamWorks 2001), seen in Figure 5.12, were shown and asked to be reproduced.

Figure 5.12: The cartoon facial expressions to simulate in the user experiment; left, middle and right: respectively the anger, fear and surprise expressions of feature film characters Sulley (Monsters Inc., Pixar 2001), Woody (Toy Story, Pixar 1995) and Shrek (Shrek, DreamWorks 2001).

The twenty expert users in character animation were selected after completing an initial questionnaire to confirm their level of expertise (see Appendix C). The questionnaires reveal the high level of expertise of these users: (i) the majority of the participants (90%) have either college or master degrees, (ii) the majority of the TGUs (80%) have attended over 120 hours of extra specific education in animation and the majority of the CGUs (80%) have attended at least 61 to 90 hours of extra specific education, of which more than half have attended over 120 hours, (iii) all the participants have a job in animation, (iv) the professional animation curriculums of the users include character development for feature films, short films, 3D animation, videogames, television series, advertisement campaigns, music clips and projects which are NDA (Non Disclosure Agreements), (v) the users resort to the different major 3D tools used in the industry like 3dsMax, Blender, Maya and Softimage, (vi) all users work in a daily basis with their preferred 3D tool, with 65% of the users working in average at least 5 hours and often over 7 hours daily and lastly (vii) all users have built animation-ready rigs and the majority of the users (80%) have specific experience in building facial rigs.

The users consented to participate via e-mail and they were randomly assigned to two groups, A and B, with each group given 10 users. User assignment to the two groups was done by (i) organizing the names of the users alphabetically, (ii) giving each user an ID number following their alphabetical order and (iii) sorting their IDs into the groups using the PsychicScience website available in the link *http://www.psychicscience.org/random.aspx*.

For a means of comparison, group A was set as the test group users (TGUs) and B as the control group users (CGUs). Group A tested the system and group B tested the conventional techniques. Each group was asked to reproduce the human facial expressions of anger, fear and surprise shown in Figure 5.11 following a sequence of instructions based on the guidelines by [Ekman and Friesen 1975].

The instructions for the facial expressions of anger, fear and surprise are distributed in a total of 7 tasks per expression, which are shown ahead in Table 5.6. The 2nd, 4th and 6th tasks seen in Table 5.6 are better performed using either the WIN or VIP, CAM or WIN and VIP or

Anger

Task 1: inner brows move fully to the center and downwards

Task 2: center brows move downwards by -1.0 (best using either WIN or VIP)

Task 3: outer brows move fully upwards

Task 4: lower jaw thrusts fully backwards (best using either CAM or WIN)

Task 5: upper jaw moves slightly upwards

Task 6: upper row of lips move about half-way upwards in less than 7s (best using either VIP or CAM)

Task 7: head moves slightly downwards

Fear

Task 1: inner brows move fully to the center and upwards

Task 2: center brows move downwards by -1.0 (best using either WIN or VIP)

Task 3: lower jaw drops slightly and moves left about half-way

Task 4: lower jaw thrusts fully backwards (best using either CAM or WIN)

Task 5: lower row of lips move downwards about half-way

Task 6: head moves right about half-way in less than 7s (best using either VIP or CAM)

Task 7: head moves slightly backwards

Surprise

Task 1: inner brows move fully to the center and upwards

Task 2: center brows move upwards by 1.0 (best using either WIN or VIP)

Task 3: lower jaw drops almost fully

Task 4: lower jaw thrusts fully forward (best using either CAM or WIN)

Task 5: upper row of lips move fully upwards

Task 6: head moves slightly backwards in less than 7s (best using either VIP or CAM)

Task 7: head leans slightly to the left

Table 5.6: Tasks relative to the facial expressions of anger, fear and surprise.

CAM modes, respectively. These are not compulsory to use but they are more appropriate. If the user resorts to one of the interaction modes indicated for tasks 2, 4 and 6, there is a higher chance to complete those tasks more rapidly, with more precision and less frustration, for the following reasons. For task 2, the WIN mode allows to type the value required. In this case the VIP mode is the 2nd best option because it allows a faster manipulation, since the minimum motion limit of the knob is -1.0. For task 4, the CAM mode presents the control knob for the jaw thrust separately and in a schematic view. The WIN mode as a 2nd option assures that the deformation is fully accomplished. For task 6 time is more relevant than precision, therefore VIP or CAM are more adequate to rapidly conclude the task instead of searching for the required lip sliders in WIN.

The experiment expectation was that the conventional techniques would encounter limitations while the multimodal system would be superior, particularly in tasks 2, 4 and 6. To confirm that the multimodal approach is relevant, it was also expected for the user to alternate between the different interaction modes during the execution of the 7 tasks. In case there is an alternation this means that the users resort to the different interaction modes, being positive towards the multimodal approach.

There is also a final 8th task per each facial expression that asks the users to reproduce the cartoon expressions of anger, fear and surprise illustrated previously in Figure 5.12. The purpose was to evaluate if the user resorted to the extra deformation parameter to complete this task and if it actually helped the user achieve the cartoon expressions.

Experiment Stages

The experiment stages include pre and posttest methods [Campbell and Stanley 1963] as follows: (i) pretests for training and tasking using conventional techniques, (ii) posttests for training and performing the same tasks as previously using the system and (iii) a questionnaire via an interview. The expectation was that a repeated finding would occur in which the artists in group A can reproduce the facial expressions in less time, with greater precision and less frustration. Each experiment stage is detailed following.

- *Stage 1: pretests for performing tasks using conventional techniques*

Both groups watch a video (3 minutes) which illustrates and describes the conventional facial animation techniques available in the state of the art. This video helps create a more compelling environment for the experiment because it minimizes the physical distance between the participant and the author as a result of the experiment being carried out online. Then only the users in group B do a brief training of the conventional techniques using Blender (2 minutes each technique) during which users are asked to freely manipulate the controls in each conventional technique to help create ambience and detect any limitations in the interaction which can occur due to internet lag before starting the tasks. Lastly, the group B users are asked to reproduce the seven tasks relative to the anger, fear and surprise human expressions using each conventional technique separately and following the instructions given by the author.

- *Stage 2: posttests for performing tasks using the system features*

After watching the video that illustrates the conventional facial animation techniques, the users in group A watch another video (3 minutes) which introduces the multimodal system and its features. Then only the group A users do a brief training of the system using Blender (2 minutes) to get familiarized with the system. In this training users are asked to freely manipulate the multimodal controls while brief explanations are provided. Then the group A users are asked to reproduce the human facial expressions following the same instructions given to the group B users. Lastly, the group A users are shown the images that illustrate the anger, fear and surprise faces of the cartoon characters seen in Figure 5.12 and are asked to reproduce them using the system freely.

- *Stage 3: questionnaire via interview*

After testing the system, the users in group A are requested to provide scores on a scale of 0 to 10 according to the following parameters: *(i) relevance of each interaction mode* (rate WIN, CAM and VIP as independent/isolated interaction modes); *(ii) system capability* (rate the ability of the system to provide a user-adapted interaction during facial animation in comparison with the independent interaction modes available in the state of the art), *(iii) system learnability* (rate the learning curve of the system, is it intuitive to learn and easy to use even after a period of not using it), *(iv) error tolerance* (rate the reliability of the system, if any errors occurred during the interaction), *(v) synchronization utility* (rate the ability of the system to provide an updated feedback of the different interaction modes when a manipulation occurs), *(vi) deformation utility* (rate the ability of the system to provide extra freedom of manipulation by adding cartoon-like nuances to achieve a more flexible deformation), *(vii) go cross platform* (rate the relevance of implementing the system in other 3D tools considering that it is based in techniques available in the major 3D tools).

Experiment Results

Does the user-adapted system approach benefit character animators? The results of the user experiment are presented following in terms of the (i) time taken, (ii) precision and (iii) users self assessment criteria.

- *Time test*

The time test involved comparing the time taken by group B (CGUs) using each of the conventional techniques separately with group A (TGUs) using the system to reproduce the human facial expressions of anger, fear and surprise. Given the limitations inherent to an online user experiment, in which the host machine needs to continuously send information via the remote connection, a bandwidth measure website (*http://speedmeter.fccn.pt*), certified by

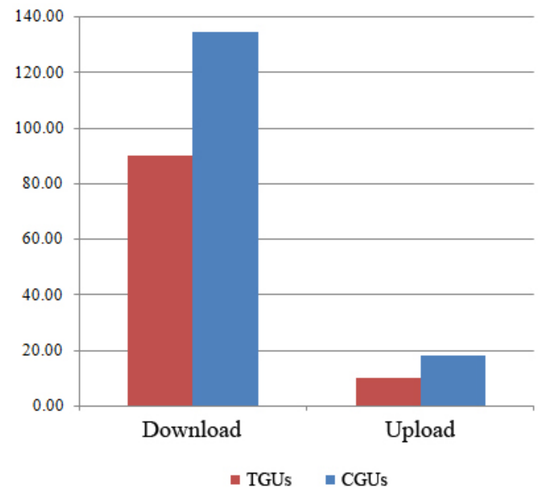


Figure 5.13: Comparison of the sum of the bandwidth values between the TGUs and CGUs for download and upload in megabytes.

the Fundação para a Computação Científica Nacional[6], was used to check both the download and upload speeds of the host and of each user. In general, the TGUs suffered more lag than the CGUs, which influenced their time performance and increased their times in comparison with the times of the CGUs. Figure 5.13 shows a comparison of the sum of the bandwidth values between the TGUs and CGUs in terms of download and upload.

Despite the significant difference in bandwidth, as seen in Figure 5.14, which shows the comparison between the average times taken by the CGUs and TGUs, the time taken by the TGUs was lower than the time taken by the CGUs in anger WIN, fear WIN and surprise VIP.

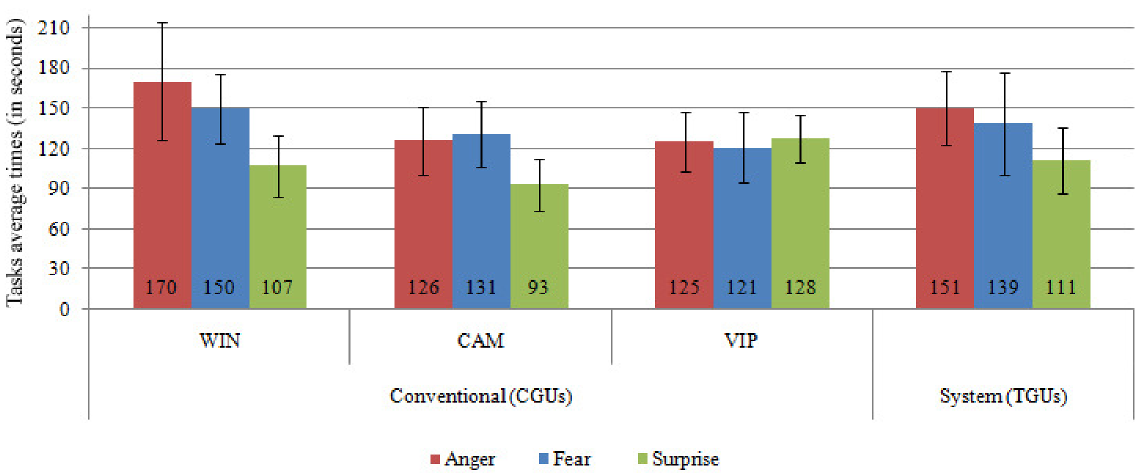


Figure 5.14: Average times comparison in seconds between the CGU and TGU groups with standard deviation error bars.

[6] http://www.fccn.pt/en/

In addition, the time taken by the TGUs in fear and surprise was only slightly higher than the time taken by the CGUs in fear CAM and surprise WIN. It is also significant to mention that the TGUs had more options to choose from, which also led to a slight increase in time consumption due to not being familiarized with the multimodal system. As a final remark, a number of CGUs requested skipping specific tasks using the conventional modes because, as they mentioned, it was difficult to complete those tasks when using a single interaction mode. Therefore, the time taken by the CGUs using the conventional modes would have increased if the skip option was not permitted. In contrast, the TGUs did not request omitting any tasks, which reveals the ability of the system to adapt to the users' needs and decrease their frustration.

- *Precision test*

The precision test involved verifying, via the analysis of the number of clicks and in which interaction mode they occurred, if the TGUs (i) resorted to different interaction modes during the simulation of the human and cartoon expressions, (ii) resorted to the WIN or VIP, CAM or WIN and VIP or CAM modes respectively for the 2nd, 4th and 6th tasks and (iii) resorted to the extra deformation parameter to reproduce the cartoon facial expressions. Figure 5.15 displayed ahead shows the TGUs performance in each interaction mode to reproduce the 7 tasks relative to the human facial expressions. Anger, fear and surprise are respectively illustrated in red, blue and green. The terms w, c and v respectively correspond to the WIN, CAM and VIP modes.

The analysis of Figure 5.15 reveals a significant variation of usage of the different interaction modes for each human facial expression. In spite of there being only seven tasks to accomplish, it is verified that each user varied in his or her approach, and even the most consistent users resorted to another interaction mode at least once. The most pertinent of

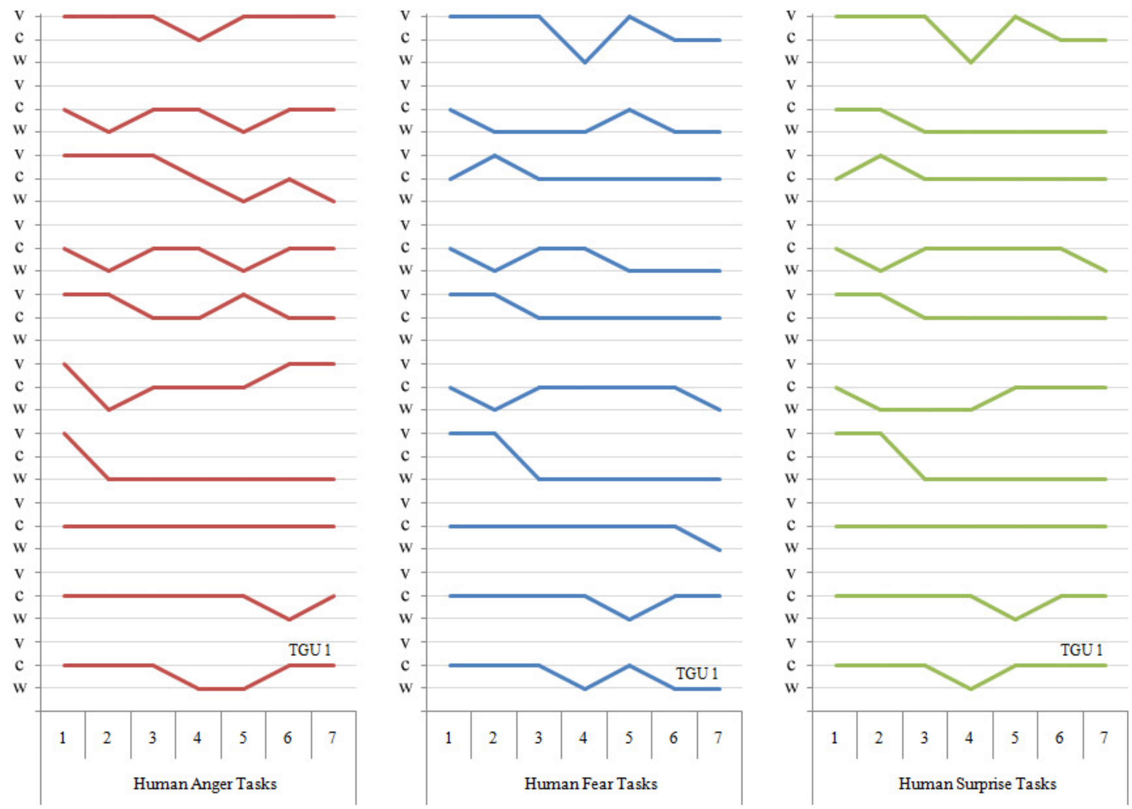


Figure 5.15: Usage variation in each interaction mode made by the TGUs for the human facial expressions of anger, fear and surprise, respectively illustrated in red, blue and green.

these cases is TGU 3, who always resorted to the camera mode except for the expression of fear. It is likely for the variation to increase or maintain in case there is a greater number of tasks to perform in an animation. Data also shows that there are not too many peaks in the interaction, therefore the system does not generate moments of uncertainty. The users generally resort more to one mode and change to another to more easily achieve specific tasks. Also relevant is the fact that there is variation from user to user, while some users resort more to a given interaction mode, others resort more to another, which means that the system adapts to the needs of each user despite their distinctive creativity.

Figure 5.16 displayed ahead shows the TGUs interaction while reproducing the cartoon expressions. The anger, fear and surprise expressions are respectively illustrated in red, blue and green. The WIN, CAM and VIP modes are respectively illustrated in darker, medium and lighter tones. The horizontal axes represent the total number of clicks in the manipulators of the different interaction modes, that each TGU performed to achieve the cartoon facial expressions.

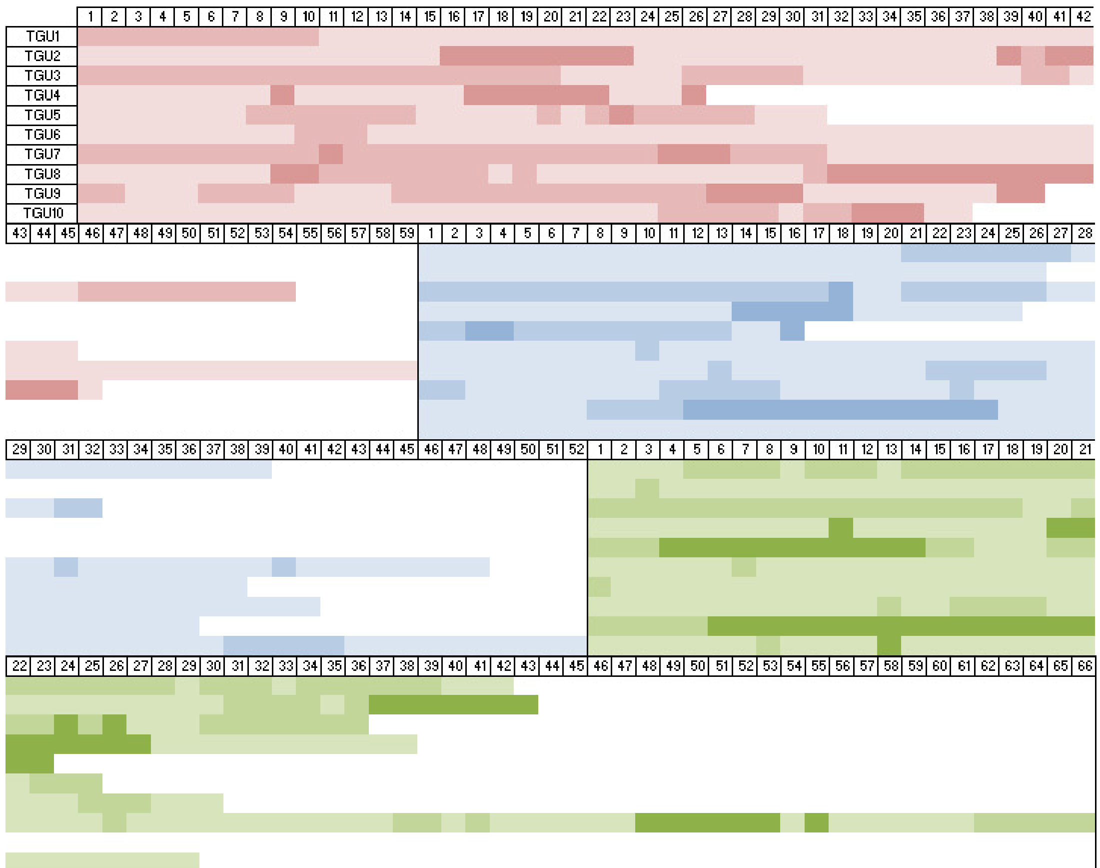


Figure 5.16: Usage variation in each interaction mode made by the TGUs for the cartoon facial expressions of anger, fear and surprise, respectively illustrated in red, blue and green.

The analysis of Figure 5.16 reveals there was also a significant variation of usage of the different interaction modes for each of the cartoon expressions. In spite of the fact that the cartoon expressions do not require as much precision as the human expressions, there still was a significant resorting to the WIN mode when performing the cartoon expressions. In fact, while reproducing both the human and cartoon facial expressions users often start with making an approximation of the required facial movement using one interaction mode and then switch to another mode to adjust the deformation. The WIN mode is usually preferred for adjustments because it allows a numerical deformation per each axis of transformation. In some instances users even select either the CAM or VIP control knobs only to gain access to their corresponding slider knobs in the WIN mode to conclude a deformation, demonstrating that the user-adapted system lets users take the most out of each interaction mode.

Table 5.7 shows the performance of the TGUs in terms of the chosen interaction modes to perform the tasks 2, 4 and 6 per each human facial expression. The slots are marked with bullet points and dashes, which respectively mean that the user did or did not resort to 1 of the 2 most adequate interaction modes for the specified task.

Users/ Tasks		Test group users (TGUs)									
		1	2	3	4	5	6	7	8	9	10
Anger	2	—	—	—	●	●	●	●	●	●	●
	4	●	●	●	●	●	●	●	●	●	●
	6	●	—	●	—	●	●	●	●	●	●
Fear	2	—	—	—	●	●	●	●	●	●	●
	4	●	●	●	●	●	●	●	●	●	●
	6	—	●	●	—	●	●	—	●	—	●
Surprise	2	—	—	—	●	●	●	●	●	—	●
	4	●	●	●	●	●	●	●	●	●	●
	6	●	●	●	—	●	●	●	●	—	●

Table 5.7: Adequacy of the interaction mode usage for tasks 2, 4 and 6 by the TGUs for the human facial expressions of anger, fear and surprise.

As seen in Table 5.7, the TGUs resorted frequently to the most adequate interaction modes for tasks 2, 4 and 6, even though they were not sufficiently familiarized with the tasks and with the system. The ratio of 80% of slots filled with bullet points compared to 20% of slots with dashes is likely to increase as users get more familiarized with the system. With more practice, it can become faster to realize that specific tasks can be performed better using specific manipulators available in specific interaction modes.

Figure 5.17 shows how the multimodal group (TGUs) used the parameter for the extra deformation per each expression. If the colored line is in the outside rim of the circumference the user resorted to the parameter and left it on to perform the cartoon expression. Any instances of the line moving to the center means that the user did not resort to the parameter. The in-between option was also possible, that of the user resorting to the parameter to test it but not really use it to conclude the cartoon expression.

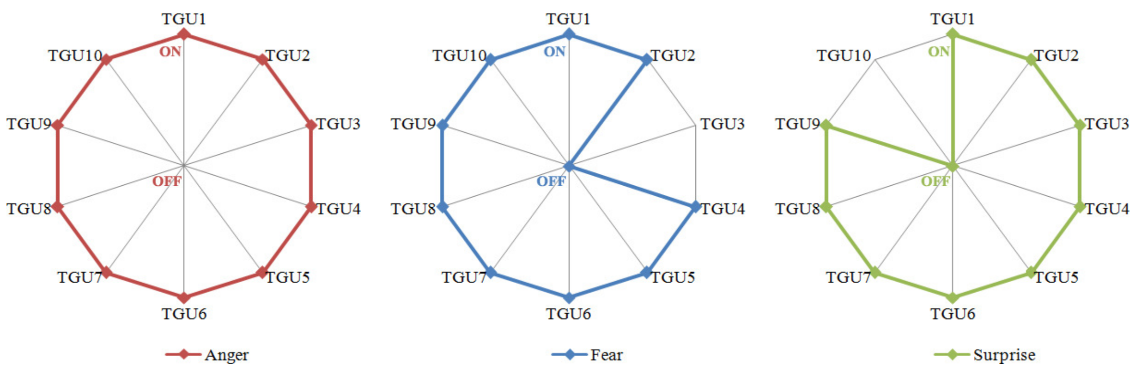


Figure 5.17: Variation in the usage of the extra deformation feature of the system for the anger, fear and surprise expressions.

The extra deformation parameter was not used only twice in a total of 30 instances, which was the case only in the expression of fear simulated by TGU 3 and in the expression of surprise simulated by TGU10. In addition, there was no instance registered of a user resorting to the extra deformation parameter but then not using it to conclude the expressions.

As a final remark, the TGUs managed to perform optimal human and cartoon expressions using the system. Figure 5.18 illustrates the expressions performed by three users of the test group. Not resorting to the extra deformation parameter allowed the TGUs to simulate the human expressions using the behavioral limitations whereas resorting to the parameter allowed them to simulate the cartoon expressions by exceeding the behavioral limitations. In human expressions there was a significant predominance of the CAM mode (120 clicks) over WIN and VIP (respectively 56 and 34 clicks). For the cartoon expressions the VIP mode was

Figure 5.18: Examples of the human and cartoon facial expressions obtained by three TGUs for the facial expressions of anger, fear and surprise.

predominant (706 clicks) over CAM and WIN (respectively 296 and 126 clicks). The CAM mode seems to be preferred for more predictable and precise animation in realistic characters and the VIP mode for more free animation in cartoon characters, with the WIN mode balancing the overall interaction.

- *Self assessment test*

Table 5.8 shows the individual scores and the final average of the individual scores which were awarded to the conventional modes and to the multimodal system by the users of the test group on a scale 0 to 10 and according to the parameters described previously in stage three of the experiment.

A t-test was performed to compare the average scores given to the conventional modes with the average scores given to the multimodal system. The t-test was two-tailed unpaired with 99% confidence and it shows significantly different means with P-value < 0.05. In a total of 60 scoring slots, 36 slots were awarded with a score of 10, followed by 10 slots that were

Parameters/ average scores	Conventional			User-Adapted Multimodal system ***					
	WIN	CAM	VIP	Capa-bility	Lear-ning	Error toler.	Sync. utility	Defor. utility	Cross platf.
TGU 1	8	10	10	9	10	10	10	10	10
TGU 2	5	8	8	10	10	8	8	10	8
TGU 3	7	8	5	10	9	8	10	7	8
TGU 4	9	5	9	10	10	10	10	7	10
TGU 5	5	5	5	10	10	10	7	5	8
TGU 6	3	7	10	10	9	10	10	10	10
TGU 7	6	7	9	8	10	9	10	10	10
TGU 8	7	8	9	9	8	9	10	8	9
TGU 9	10	9	10	10	10	10	9	10	10
TGU 10	4	6	9	9	10	10	10	8	9
Average	6,4	7,3	8,4	9,5	9,6	9,4	9,4	8,5	9,2

Table 5.8: Individual scores and final average of the individual scores provided by the TGUs to the conventional modes and to the user-adapted multimodal system for each parameter on a scale 0 to 10; columns with the scoring of the isolated conventional modes WIN, CAM and VIP; columns with the scoring of the multimodal system in terms of capability, learning, error tolerance, synchronization utility, deformation utility and cross platform.

awarded with a score of 9 and 10 slots with a score of 8, totalizing 56 slots in 60 with scores above 8. There was not a single evaluation parameter in the multimodal system awarded with less than a total of five top scores, with the *learning* parameter and the *synchronization utility* parameter alone receiving seven top scores each. The final average of the scores given to the multimodal system is 9,3 in a scale 0 to 10.

The multimodal interaction was often reported by the TGUs as precise and flexible, able to inspire a feeling of liberty in control and more free of frustration. One user reported that "even if one does not realize a given interaction mode immediately, the other modes provide real-time feedback, so one catches up and has an overall realization". Another user reported: "as I performed more tasks I learned more and in the final tasks I was completely aware of what I was doing". One TGU argued that the extra deformation feature can allow any expression if always left on, which is an interesting topic for further debate. But, one has to consider that a more realistic expression is subject to anatomic limitations, which need to be perceived when animating realistic characters in order to prevent the character to adopt exaggerated poses. The majority of the TGUs consider it extremely relevant to be able to pose both realistic and cartoon expressions, with two users having mentioned that "no artist likes limitations in control" and "you want absolutely to have that".

As a final remark, with 60% of the TGUs using other 3D tools besides Blender (e.g. 3dsMax, Maya and Softimage), there was a general consensus that it would be relevant to extend the system beyond Blender and into other 3D tools.

Chapter 6

Trends in Facial Rig Control Design

This chapter describes a final validation that was carried out to confirm the relevance of the personalized facial rig control approach presented in Chapter 5 and to try to uncover future trends in facial rig control design. This validation was carried out with twenty two undergrad students with experience in 3D modeling to realize if these users welcome the benefits of each interaction mode and also if they can provide extra relevant data to define future guidelines to improve the design of the rig controls in each interaction mode. The user experiment is described in Section 6.1 and the experiment results are presented in Section 6.2.

6.1 User Experiment

Unlike the design of the facial rig mechanics, which is less diverse because it needs to follow specific construction and deformation conditions, which are presented in Section 4.1 of Chapter 4, the design of the facial rig controls is more diverse, because it is more dependent on the visual and interaction preferences of the user. To help shed light on the design of the different conventional facial rig control interfaces, a final experiment was carried out with twenty two students who attend a course of 3D modeling (90h) in the degree in Audiovisual and Multimedia of the Superior School of Social Communication of the Polytechnic Institute of Lisbon (*http://www.escs.ipl.pt*) in December 9th, 2014. The purpose of this experiment was to (i) realize if these users also benefit from using the different interaction modes and (ii) receive their feedback to uncover relevant data to help define future guidelines to improve the design of the facial rig controls in each interaction mode. This experiment can increase chances to obtain creative feedback due to the number of users participating and their predisposal to present and discuss new trends for the design of the facial rig controls, as they are undertaking a degree in which they learn and explore different topics.

Experiment Design

This experiment is intended to confirm the relevance of the personalized rig control approach presented in Chapter 5 and uncover future trends in facial rig control design. It is aimed at realizing if undergrad students will welcome the benefits of each interaction mode and if they can provide information that can help define future guidelines to improve the design of the rig controls. The hypothesis of the experiment are that (i) users will experience difficulties when resorting to a single interaction mode to complete the proposed tasks because each task is better performed using specific interaction modes, thus validating the relevance of the multimodal system presented in Chapter 5, (ii) users may prefer the viewport mode but also realize and value the benefits of the other interaction modes and be able to adapt to each mode and (iii) users can provide relevant feedback to improve the design of the facial rig controls.

Experiment Stages

The experiment occurred in three stages: (i) a group training stage to get users familiarized with the different interaction modes, (ii) a group tasking stage in which the users test the different interaction modes and (iii) a focus group stage in which ten users were randomly selected to participate and engage in a discussion. Each stage is detailed following.

- *Stage 1: Training Stage*

In the training stage the users trained the conventional interaction modes window, camera and viewport in group in a room where each user had a computer. The users trained each mode at a time to get used to them. This stage had a total duration of 30 minutes and was divided in the following moments: (i) experiment explanation and users training (5 + 15 minutes) and (ii) consent sheet and skills questionnaire (10 minutes). In the first moment a brief introduction of the different interaction modes was provided to the users by resorting to a projection in a large screen. Each user then trained each of the interaction modes freely using the computer in front of them in a total of 15 minutes, having 5 minutes per each interaction mode. During this moment the experiment supervisor clarified any doubts that the users may have relative to handling each interaction mode. In the second moment each user signed an experiment participation consent sheet (see Appendix B) and completed a questionnaire to confirm their level of expertise (see Appendix C).

- *Stage 2: Tasking Stage*

In the tasking stage the users simulated three tasks in each conventional interaction mode in group and with time restriction. Users were aware that there was a time limit to complete the tasks. This stage had a total duration of 15 minutes and was divided in the following moments: (i) perform same tasks in each interaction mode (20 seconds per task, 1 minute per mode, 3 minutes total) and (ii) an evaluation questionnaire (12 minutes). In the first moment the users simulated three tasks in each interaction mode that are most fundamental to achieve a basic pose of anger for the human subject facial model by following instructions based on the guidelines by [Ekman and Friesen 1975]. These tasks are presented ahead in Table 6.1 and they involve manipulating controls in the facial regions of the eyebrows, the jaw and the lips to perform a basic expression of anger, being better performed using either the WIN or VIP, CAM or WIN and VIP or CAM modes, respectively. These are not compulsory to use but they are more appropriate for the following reasons. For task 1 the WIN mode allows to type the value required and VIP as a 2nd option allows a faster manipulation. For task 2 the CAM mode presents the jaw thrust control separately in a schematic view and WIN as a 2nd option assures that the deformation is fully accomplished. For task 3 time is more relevant than precision, therefore VIP or CAM are more adequate to rapidly conclude the task instead of searching for the required lip sliders in WIN. In the second moment each user completed a questionnaire to evaluate each conventional mode and also to obtain information that can help define future guidelines to improve the design of the rig controls (see Appendix C).

Anger
Task 1: inner and center brows move downwards by -1 (best using either WIN or VIP)
Task 2: lower jaw thrusts fully backwards (best using either CAM or WIN)
Task 3: upper row of lips move about half-way upwards in less than 7s (best using either VIP or CAM)

Table 6.1: Tasks to achieve a basic facial expression of anger.

- *Stage 3:Focus Group Stage*

In the focus group stage 10 users were randomly selected to participate and engage in a discussion guided by the researcher [Coutinho 2013] with the purpose to further help define future guidelines to improve the design of the facial rig controls. Each user signed a consent sheet to participate in the focus group stage (see Appendix B).

6.2 Experiment Results

To answer the formulated hypothesis and help define future guidelines to improve the design of the facial rig controls, the collected data is presented following in four moments: (i) the precision of the users interaction via the analysis of the number of clicks performed in each interaction mode to complete the tasks, (ii) the time taken by the users in each interaction mode to complete the tasks, (iii) the answers provided by the users in the evaluation questionnaires and lastly (iv) the users feedback provided during the focus group stage. To analyse the statistical significance of the results a one-way ANOVA with post-hoc Bonferroni multiple comparison test with 95% confidence intervals was performed in GraphPad Prism[7].

Precision

In terms of the precision, Figure 6.1 shows the number of valid and invalid clicks and the percentage of over clicks (more than necessary valid clicks) that were made by the users. The necessary clicks for each interaction mode for tasks 1, 2 and 3 is respectively of 4, 1 and 3, as illustrated ahead in each of the graphics in Figure 6.1. The analysis of the average number of clicks performed by the users in each interaction mode to complete each task goes in accordance to the relevance of the most adequate modes to use per each task. For task 1, the valid clicks in WIN vs VIP and CAM vs VIP are significantly different (respectively t=2.546 and t=2.886 with P=0.0099) and the invalid clicks in WIN vs CAM and WIN vs VIP are significantly different (respectively t=4.369 and t=5.319 with P<0.0001). The significant average number of invalid clicks that is verified in the WIN mode (1.4), which originates due to the users having to interpret the higher number of controls listed in this mode, is then compensated by the low average number of invalid clicks verified in the VIP mode for this task (0.1). In task 2, VIP has a higher average number of valid clicks (1.0) than CAM (0.9) or WIN (0.9), but the average number of invalid clicks for this task is higher in VIP (0.8) than in CAM (0.5) or WIN (0.5), which means that the users performed more valid clicks in VIP but also more invalid clicks in VIP to complete this task. Also, the difference in the average number of valid clicks in CAM (0.9) or WIN (0.9) for task 2 is only slightly lower than in VIP (1.0), which supports the more accurate tendency of the users to use CAM or WIN for task 2. For task 3, the valid clicks in WIN vs CAM and WIN vs VIP are significantly different (respectively t=2.727 and t=3.068 with P=0.0055), corroborating VIP or CAM for this task.

---

[7] http://www.graphpad.com/scientific-software/prism/

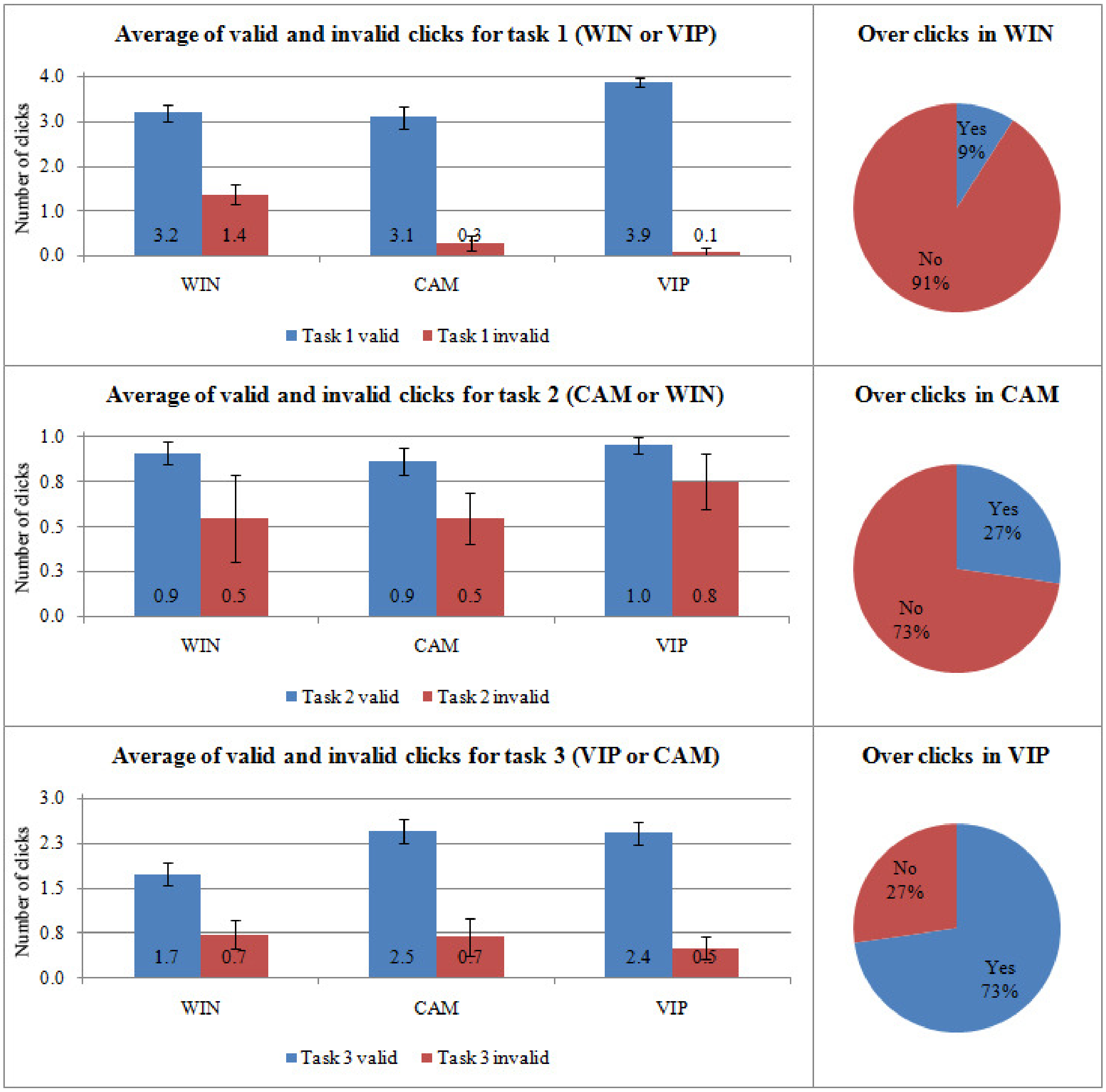


Figure 6.1: Users average number of valid and invalid clicks in each interaction mode to complete tasks 1, 2 and 3 with standard error of the mean (left) and over than necessary valid clicks made by the users in each interaction mode (right).

Despite the fact that the window mode can be more difficult to decode visually, it is more precise. Figure 6.2 shows that 91% of the users resorted to typing values in the WIN mode besides dragging the sliders to complete task 1 and that 36% of the users resorted to orbit the 3D view to complete task 2 in the VIP mode, which means that the users resorted to other functionalities in VIP which are not fundamental in the other modes to better complete task 2.

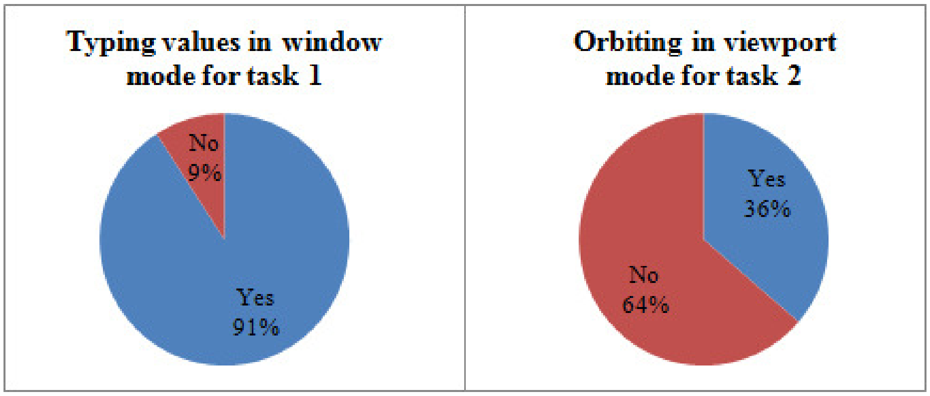


Figure 6.2: Percentage of users typing values in window mode for task 1 (left) and orbiting in viewport mode for task 2 (right).

Time

In terms of the time taken, Figure 6.3 illustrates the average times required by the users to complete each task in each interaction mode.

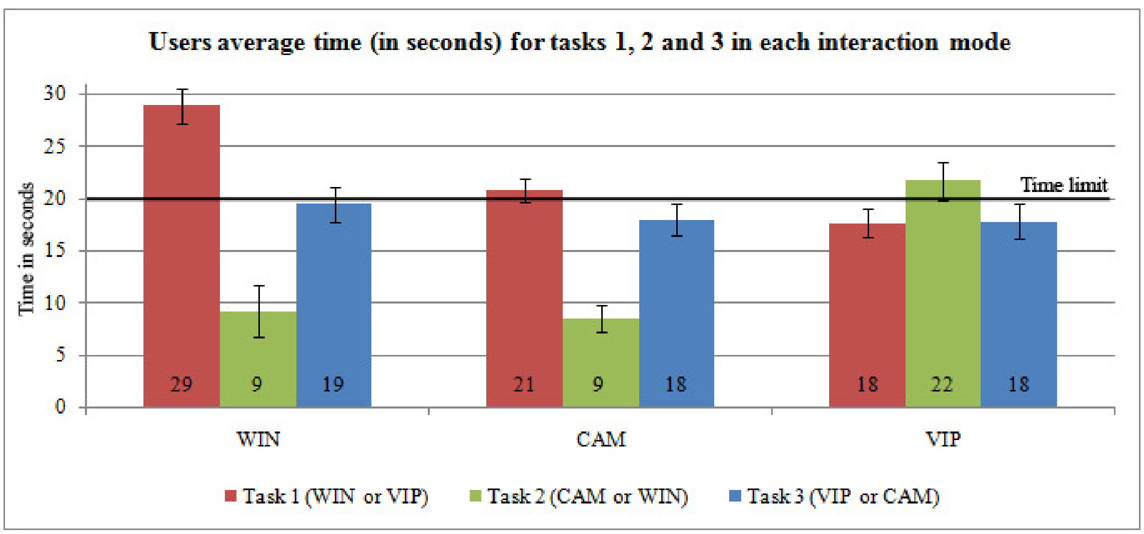


Figure 6.3: Average time taken in seconds by the users in each interaction mode to complete tasks 1, 2 and 3 with standard error of the mean.

There was a time limit of 20 seconds that the users were aware to complete each task in each interaction mode. The analysis of the time taken by the users corroborates the analysis of the precision made previously. For task 1, WIN vs CAM and WIN vs VIP are significantly different (respectively $t=4.167$ and $t=5.768$ with $P<0.0001$), with users requiring more time in WIN (29s) followed by the CAM (21s) and VIP (18s) modes, which suggests an efficient balance between WIN and VIP to better complete this task. For task 2, WIN vs VIP and CAM vs VIP are significantly different (respectively $t=4.542$ and $t=4.791$ with $P<0.0001$), with the CAM or WIN modes (9s each) being clearly more adequate than VIP (22s). It is interesting to verify that the average time is the same for CAM and WIN, which reveals the positive impact of the WIN mode, since WIN is considered the 2^{nd} best option to complete this task. For task 3, the VIP or CAM modes (each 18s) are slightly faster than WIN (19s). It is also interesting to verify that WIN is only slightly slower to use for this task. In overall, the average times only tend to being balanced for task 3, which still is performed in less time using its respective most adequate interaction modes.

Questionnaires

In terms of the evaluation questionnaires (see Appendix C), the users were asked to (i) score each interaction mode in a scale 0 to 10 to perform each task, (ii) if the users agree with the statement that the window mode is more exact, the camera mode is more intuitive and the viewport mode is more direct, (iii) which interaction mode is preferred by the users to transfer values from a character's face to another, to achieve specific deformations without the need to orbit the 3D view and to perform a quick posing of the character's face, (iv) which interaction mode is the most preferred by the users, (v) if the users have felt frustrated when using one interaction mode at a time and lastly (vi) if the users have any suggestions to improve the current control distribution in each interaction mode.

Figure 6.4 shows the results obtained in terms of the users average scores and their agreement with the window mode being more exact, the camera mode being more intuitive and the viewport mode being more direct.

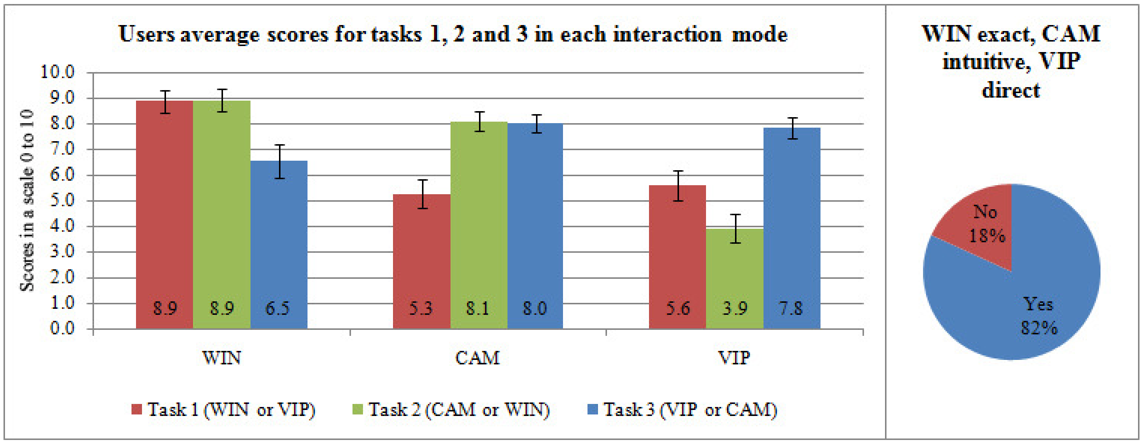


Figure 6.4: Users average scores for each interaction mode to complete tasks 1, 2 and 3 with standard error of the mean (left) and percentage of users who agree with the qualities of each mode (right).

The analysis of Figure 6.4 reflects the relevance of the most adequate interaction modes to achieve each task. For task 1, WIN vs CAM and WIN vs VIP are significantly different (respectively t=4.749 and t=4.328 with P<0.0001), with the users clearly preferring WIN (8.9) over VIP (5.6) and CAM (5.3), with VIP confirmed as 2nd best option. For task 2, WIN vs VIP and CAM vs VIP are significantly different (respectively t=7.710 and t=6.448 with P<0.0001), with users also preferring WIN (8.9), the 2nd best option for this task, but not very above CAM (8.1) and with both CAM or WIN being clearly above VIP (3.9). For task 3, users prefer CAM (8.0), the 2nd best option for this task, but only slightly above VIP (7.8) and with both VIP or CAM being significantly above WIN (6.5). The fact that the majority of the users (82%) agree with the statement that the window mode is more exact, the camera mode is more intuitive and the viewport mode is more direct, also helps to perceive the relevance of each interaction mode for the different tasks required.

Figure 6.5 shows the results obtained in terms of the users preferences to achieve different simulations in each interaction mode. The analysis of Figure 6.5 is also revealing of the benefits of each interaction mode for different simulations.

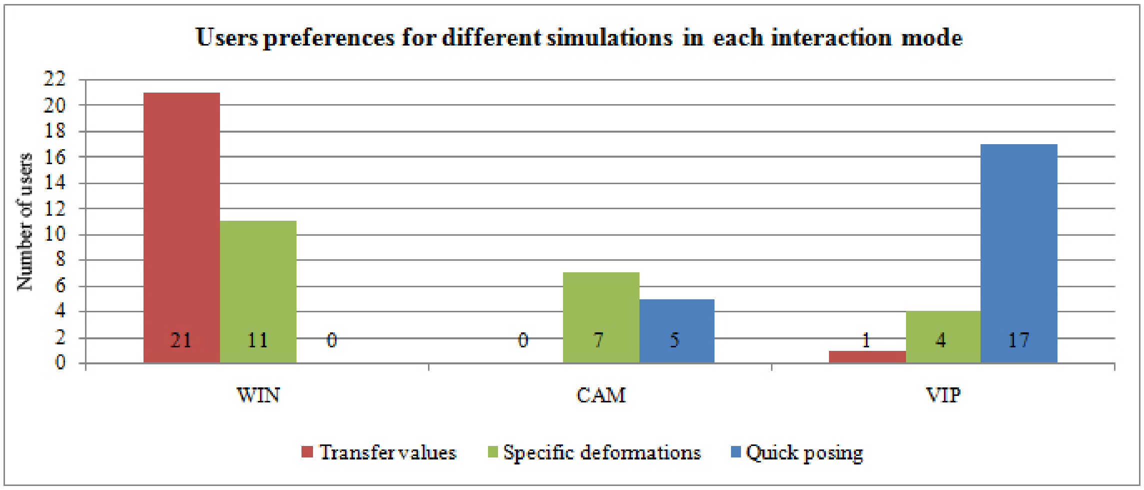


Figure 6.5: Users preferences for different simulations in each interaction mode: transfer values, perform specific deformations and quick posing.

Of the 22 users, 21 choose the window mode to transfer values from a character's face to another, with only one user selecting the viewport mode for that purpose. To achieve specific deformations the users preferences are more distributed, but still 11 users choose the window mode to perform specific deformations without orbiting the 3D view, 7 users choose the camera mode and 4 users choose the viewport mode for that purpose. A total of 17 users choose the viewport mode to perform quick posing, with 5 users selecting the camera mode for that purpose.

Figure 6.6 shows the results obtained in terms of the users preferred interaction mode to achieve different simulations and if the users were frustrated when using one interaction mode at a time to complete the tasks. The analysis of Figure 6.6 shows that there is a preference of the users for the WIN mode (13 users), but also that the VIP (6 users) and CAM (3 users) modes were considered relevant. Despite the majority of the users prefer the window mode, the users preferences are clearly distributed, which shows that the users realize and value the benefits of the other interaction modes. In addition, 95% of the users were frustrated when using a single interaction mode at a time, which corroborates the hypothesis that the users

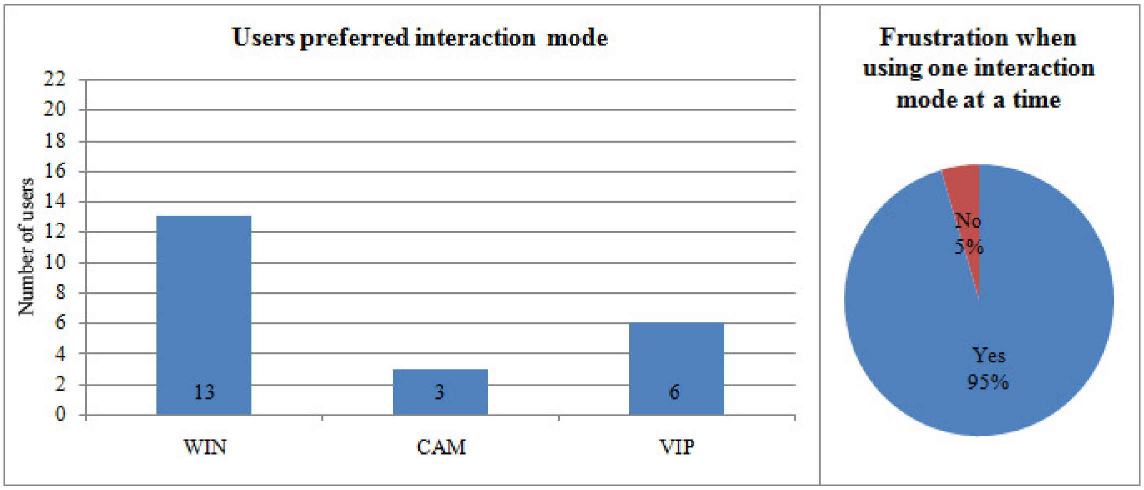


Figure 6.6: Users preferred interaction mode (left) and users frustration when using one interaction mode at a time (right).

experience difficulties when using a single interaction mode and therefore are keen of the different modes together, which helps to confirm the relevance of the multimodal system presented in Chapter 5.

Lastly, the questionnaires also allowed the users to provide suggestions to improve the current distribution of the controls in each interaction mode. The users replies and conclusions retrieved are presented following divided into the different interaction modes and into a set of general conclusions.

In terms of the window mode, the users find it slower to work with due to having to search for the required sliders. Nonetheless, most users prefer the window mode and one user reported to prefer the window mode despite the fact that it is slower to manipulate and that sometimes the user gets lost. To improve this mode, the users have highlighted the importance of providing a better distinction between the different areas within each facial region. For instance, improving the distinction between the upper and lower lip slider control knobs would allow to understand more clearly which area the user is manipulating even before the deformation has occured. It would also be relevant to have an option to move pairs of facial muscles, to prevent having to change the left and right areas of a facial region in case the purpose of the deformation is for those regions to change symmetrically. In terms of the camera mode, the users have highlighted that the nomenclature of the controls could be improved, as sometimes the users were not able to read well the tags that identify each panel due to the colors or even the font used. To improve this mode, the users suggested to add an extra slider to open/close the lips to help facilitate the creation of a considerable number of expressions, instead of modifying the upper and lower lips in separate (such feature can be compared to adding controls that trigger predefined expressions, which seems suitable to be implemented in the camera mode due to its schematic appearance). In terms of the viewport mode, the users have highlighted that the manipulators can sometimes complicate the visualization of the obtained result, in some areas even overlapping each other. To improve this mode, the users have highlighted that the jaw control seemed too sensitive, likely due to triggering multiple deformations. The controls in general could benefit from being visually identified regarding the facial region or area that they deform, due to the proximity that exists between the controls, to prevent the difficulty of having to identify which control is relative to what deformation. In terms of general conclusions, the users have highlighted the relevance of adding to the camera and viewport modes the ability to access information of the deformation values. One user mentioned the possibility for the window and viewport modes to work together in two different windows, another user even recommended using the window and viewport modes simultaneously and another user mentioned that the fusion of the window and viewport modes would allow a greater precision with proximity and a direct control of the facial model. As a final remark, the users report that the different interaction modes are well conceived and that the manipulation of the controls is well achieved with a little patience, which is significantly positive given their experience and the fact that they have had a relative short period of time to use each interaction mode.

Focus Group

The focus group session allowed to obtain further information regarding the improvement of the design and distribution of the facial rig controls. The users provided relevant feedback through a discussion in group that added value to the suggestions they provided in the written questionnaires, with a number of interesting conclusions for future guidelines to improve the design of the facial rig controls being described following, which include (i) control display, (ii) simmetry and order, (iii) generic and specific controls, (iv) multiple window bars and (v) floating windows.

In terms of (i) control display, the users recommended adding more options for facial deformations for each muscle and for each control option or facial region create a specific window which can be sectionable, with the ability to hide/show options to be more organized (such control of the display is available in the system approach presented in Section 5.1). In terms of (ii) simmetry and order, the users recommended to improve the organization of the several panels, for instance instead of being organized from the left to the right side of the face. It would also be interesting to have options for symmetry, to allow deforming both sides of the face when manipulating one side only. In addition, it would be interesting to allow the user to reorganize the order that the panels appear, even going further than reordering panels and into also reordering different sections inside each panel. In terms of (iii) generic and specific controls, the users highlighted the possibility to manipulate multiple controls together, having a generic control that groups some of the more specific controls, depending on each facial region (e.g. to allow the upper lip controls to move together). In terms of (iv) multiple window bars, the users highlighted the possibility in the window mode to drag the menu to increase its size, to better visualize the control sliders and to create more than a single bar to list the several sliders, creating a stack-like approach to avoid having to perform a long scrolling to reach sliders that are available further down the list. In terms of (v) floating windows, the users highlighted the possibility to click in a specific facial region in the viewport to make a floating window appear. This floating window would be a panel of the window mode with control sliders to allow inserting values. To conclude the discussion, the researcher added for debate two questions, the first to realize the users opinions concerning the relevance of a multimodal system that combined the different interaction modes and the second to realize if and what the users would change in the different interaction modes to be able to manipulate cartoon-like characters. The answer of the users to the first question was a general consensus that the interaction would be facilitated by means of a multimodal system combining the different interaction modes. The users also highlighted that this system could make the visualization more difficult due to the several controls being available together and suggested the possibility to show the options relative to the current control selection, especially in the window mode to prevent having to scroll in case a specific option is required which is available further down in the list of sliders. Such control of the display is presented in Section 5.1, as well as displaying the controls of the facial region relative to the current selection, which is presented in Section 5.2, which helps to corroborate the relevance of the findings presented in Chapter 5. To answer the second question the users have highlighted the possibility to extend the minimum and maximum values of the sliders in the window mode, to resort to other deformation scales via a parameter that would allow to go beyond the human deformation limits. Such control parameter is presented in the multimodal system approach described in Section 5.2, which allows an extra deformation via a single parameter to achieve cartoon-like poses.

Chapter 7

Conclusions and Future Perspectives

This dissertation describes optimized facial rig mechanical approaches in Chapter 4 and personalized facial rig control approaches in Chapter 5 that together ease the laborious, disparate and fragmented process of character facial rigging and animation. The mechanical approaches involved an initial realization of the behaviors of the human face and a follow-up construction of the corresponding most adequate facial rig mechanical components to support those behaviors, based on construction and deformation conditions and on validation with expert users who are professors of 3D. The control approaches involved the construction and validation of a multimodal facial rig interface system approach based on multiple seamless control inputs that provide a more free manipulation of the facial deformation. The multimodal approach was validated in an initial user experiment carried out locally with basic, skilled and expert digital artists, and in a final user experiment carried out online with expert digital artists who are long-term professionals in the field. Lastly, a validation was carried out with undergrad students in Chapter 6 to confirm the relevance of the multimodal system and uncover future trends in facial rig control design. This chapter resumes the main conclusions obtained from this research and presents future perspectives for research.

7.1 Conclusions

This dissertation presents a series of approaches to construct and use facial rig mechanics and controls for a human subject facial model as a basis for other facial styles. The reason for the development of a multitude of approaches, which begin in the rig mechanics and end in the rig controls, is their inherent integration with each other. The rig mechanics deform the facial model and the rig controls are the user interface to manipulate the facial model by triggering the underlying mechanics. Although the two stages can be approached separately, they are considered together in character production pipelines to allow animators to accurately deform a facial model. Instead of researching on a single approach, either relative to the facial rig mechanics or to the facial rig controls, the different approaches presented are realized as an overall contribution to the field, because they are linked in a sequence that results in a full optimized rig that stands out from the isolated cases found in the state of the art. The facial rig approaches presented encompass both the mechanics and controls in a complete way. Moreover, the approaches presented can also be realized as independent, with the artist being free to decide whether to retrieve parts of information from this dissertation or learn from it entirely. In this field of research, which is open to significant improvements and has different topics to explore, the research direction presented seems consensual and provides a storyline that makes sense. Other artists, who use other 3D tools rather than Blender, can also benefit from reading this thesis, because the approaches presented are likely to be applicable to other 3D tools. In fact, the approaches have been evaluated in the user experiments by users who resort to Blender and users who resort to other 3D tools in their daily work, who concluded that the approaches presented are highly prone to be recreated in other 3D tools.

In terms of the mechanical approaches presented, it can be mentioned with confidence that these approaches help ease the job of the rigger, because they clarify how to place and configure the main rig mechanical components for each facial region of a human subject facial model: the neck, head, eyebrows, eyelids, eyeballs, nose, cheeks, jaw, lips and the tongue. The rig mechanics are considered generic due to their focus in the human face as a basis for other facial styles. They are also considered as validated guideline proposals to build facial rig mechanics because they are based in construction and deformation conditions which were developed to justify the mechanical approaches based on the study of the human facial behaviors. Within the mechanical definition of the lips, the research has been taken further to support the zipper and sticky lips effects, because these effects are difficult to simulate, rare to find in the state of the art and need to be both realistic and appealing, in particular for realistic faces. In this case, a validation has been carried out with five expert professional users who are also higher education professors in the field. The zipper and sticky lips approaches are accessible and efficiently usable by most digital artists because (i) they are not programmable based, (ii) they are complete, as they include various features (upper and lower jaw control, individual lip control and zipper and sticky lips effects) and (iii) they are developed in the open-source 3D tool Blender. Despite the implementation has been carried out in Blender, the approaches are highly prone to be implemented in other 3D tools, since they are based in techniques which are available in the major 3D production packages. The number of five users who tested the zipper and sticky lips approaches is not realized as a limitation but rather as a significant qualitative study because these users (i) are professional experts and also higher education professors of rigging and animation, (ii) are experts in different major 3D tools (including 3dsMax, Blender and Maya) and (iii) have distinct influences given their different locations, which encompass the current major research and production sites in this field in Portugal (Barcelos, Porto, Lisboa and Viseu) and a foreign location in Brazil (Rio de Janeiro).

In terms of the control approaches presented, it can be mentioned with confidence that these approaches help ease the job of the rigger and also of the animator, because they clarify how to place and configure the controls and how to use the controls in a flexible way. The rig controls include the window, camera and viewport-based key frame animation techniques combined in an interactive multimodal interface system approach. The orientation towards key frame animation is because (i) this is still the most widely used animation technique, (ii) a great deal of clean up and refinement work is done via key frame after motion capture and (iii) it is not feasible to mocap cartoon faces, only realistic human characters. The control approaches are indicated for both realistic and cartoon expressions with potential to be used independently of the character's face and 3D tool, because the system features are suitable to be included in a facial rig by default and the system is based in techniques available in the major 3D tools. The major benefit of the multimodal approach is that it allows the animator to freely choose what control to use for which particular facial behavior or target emotion to be expressed. The multimodal approach is based on not imposing a single control scheme but rather provide a more free control interaction based on multiple controls. For riggers, it acts as an optimized methodology to build the facial rig controls and for animators as a customizable interface that adapts to their preferences. An initial experiment was carried out with a group of digital artists to assess the potential of this approach. Their performance and comments revealed the potential of this approach to decrease the need to adjust the rig and excel the animators' performance and enjoyment. The multimodal system was then upgraded into a user-adapted approach and tested in a final follow-up comparative user validation. This final user study was carried out with twenty expert digital artists who are long-term professionals in character rigging and animation, with 60% of the test group users being international. The user experiment was both qualitative and quantitative in demonstrating that the multimodal user-adapted approach is a relevant contribution for facial rigging and animation based on the results in terms of time, precision and the scores awarded by the users. The multimodal control system was tested with support for the eyebrows, jaw, lips and head as an example extendable to other facial regions. In case it is necessary for the rigger to include support for other facial regions (e.g. eyelids, eyeballs, nose or cheeks), three main steps need to be carried out: (i) declare in the first algorithm (Section 5.2.2) the new facial region operator to appear in the main panel of the rig interface, (ii) code instances of the second algorithm (Section 5.2.3) per each facial region, adapting and extending these instances according to the transformations required, so that the controls in each region can trigger their corresponding deformations and (iii) code instances of the third algorithm (Section 5.2.4) per each facial region, adapting and extending these instances to detect clicks of the left mouse button performed by the user in the different manipulators. Lastly, the results of the experiment described in Chapter 6 are in accordance with the results obtained previously and thus help confirm the multimodal system as a relevant trend for facial rigging. The user data in terms of number of clicks, time and the users replies in the questionnaires and in the focus group session, show that the users realized the benefits of each interaction mode and also provided relevant suggestions to improve the distribution of the controls in each interaction mode, contributing to define the future perspectives presented ahead.

7.2 Future Perspectives

Facial rigging and animation continues to be a vast research and production field with a number of interesting directions that can be further explored after this research, which are described following in terms of both the facial rig mechanics and the facial rig controls.

- Due to their intrinsic nature, the facial rig mechanics require experienced artists in rigging to be validated. But, there are less riggers than animators in the industry. Therefore, it can be relevant to research new methods to validate the rig mechanics by consulting with riggers. The creation of a database to gather artists can aid in contacting them, as well as disseminate new ideas in character rigging and animation;
- Rigging the tongue is a topic little approached and described by artists and an area to explore further, both in terms of behaviors as well as in terms of rig approaches;

- Additional research can be conducted in the integration of the presented multimodal rig interaction model with cartoon faces, taking care to identify their behaviors in order to extend the multimodal system approach accordingly and if necessary;
- Researching new ways to merge the camera and viewport modes in the multimodal rig, as a follow-up of the preliminary result described in Section 3.3.1 of Chapter 3. An example is to adjust the camera-based widgets to the current angle of view of the character's face, in order to facilitate triggering certain behaviors like the jaw thrust or the movement of the tongue, which are better perceived from a profile view;
- The extension of the multimodal rig control approach to a character's body can help the artists manipulate certain body parts (e.g. a hand or a creature tail). In this case, the combination of synchronized gizmos, osipas and sliders can respectively provide a faster, a schematic and a more detailed posing;
- Although more valid for human characters than cartoon-like ones, the combination of the multimodal approach with motion capture solutions can also be a relevant topic to explore in order to speed up the animation process using the multimodal rig interface;
- In the follow-up of the users recommendations in Chapter 6, the development of an interaction based on floating windows in the viewport mode, which are accessible by clicking in specific regions of the facial model in the viewport. The floating windows would be panels of the window mode with control sliders to allow inserting values;
- The addition of other types of control widgets to the multimodal interface to extend its current functionality (e.g. osipas and gizmos with other custom shapes and different behavior). These can be coupled with the possibility for the animator to include predefined widget designs that would be available in a widget library, as well as provide the ability to include personalized widget designs and including multiple window bars, in the follow-up of the users recommendations in Chapter 6, who have

highlighted the possibility in the window mode to drag the menu to increase its size, to better visualize the control sliders and to create more than a single bar to list the several sliders to avoid having to perform a long scrolling to reach sliders that are available further down the list;

- Lastly, the implementation of an automatic facial rig system compatible with different styles of character faces. This automatic approach would be based on the mechanical and control approaches presented in this thesis, which can be deployed as an add-on for Blender or a plug-in for other 3D tools to automatically generate a more friendly animation-ready facial rig.

Facial rigging and animation will continue to be a vast field of research and production capable of presenting more challenges in the upcoming years. The author hopes that the approaches presented in this dissertation can generate future research to excel the field of character facial rigging and animation.

References

ABONITALLA, R. 2010. Hulk Hogan Face Rig Demo. In URL: https://vimeo.com/15157374. Checked 12/2014.

AITKEN, M., BUTLER, G., LEMMON, D., SAINDON, E., PETERS, D. AND WILLIAMS, G. 2004. The Lord of the Rings: The Visual Effects that Brought Middle Earth to the Screen. In *Proc. of SIGGRAPH Course Notes*.

ALEXANDER, O., ROGERS, M., LAMBETH, W., CHIANG, M., AND DEBEVEC, P. 2009. Creating a Photoreal Digital Actor: The Digital Emily Project. In *Proc. SIGGRAPH Course Notes*, pp. 176-187.

ANIMATION INSTITUTE. 2014. Facial Animation Toolset. In URL: http://research.animationsinstitut.de/facial-research-tools/facial-animation-toolset-2015/. Checked 12/2014.

ANDERSON, R., STENGER, B., AND CIPOLLA, R. 2013. Lip Tracking for 3D Face Registration. In *Proc. of the 13th International Conference on Machine Vision Applications*, May 20-23, Kyoto, Japan.

ANZOVIN STUDIO. 2014. The Face Machine. In URL: http://www.anzovin.com/tfm. Checked 12/2014.

AQUARO, M. 2009. Mario Aquaro Character Rigging Reel. In URL: http://www.youtube.com/watch?v=zximkpu9D0I. Checked 12/2014.

ARGHINENTI, A. 2011. Animation Workflow in Killzone3: A Fast Facial Retargeting System for Game Characters. In *Proc. SIGGRAPH Talks*.

ATHIAS, D. 2006. *Facial Rigging in Maya*. Digital-Tutors Training DVD. In URL: http://www.digitaltutors.com/tutorial/1593-Facial-Rigging-in-Maya-7. Checked 12/2014.

ATHIAS, D. 2010. *Curve-Based Facial Rigging in Maya*. Digital-Tutors Training DVD. In URL: http://www.digitaltutors.com/tutorial/389-Curve-Based-Facial-Rigging-in-Maya. Checked 12/2014.

ATHIAS, D. 2010. *Character Rigging for Production in Maya*. Digital-Tutors Training DVD. In URL: http://www.digitaltutors.com/tutorial/321-Character-Rigging-for-Production-in-Maya. Checked 12/2014.

ATHIAS, D. 2011. *Rigging the Human Face in Maya*. Digital-Tutors Training DVD. In URL: http://www.digitaltutors.com/tutorial/27-Rigging-the-Human-Face-in-Maya. Checked 12/2014.

AU, A. 2009. Friendly Facial Rigging. In URL: http://www.youtube.com/watch?v=Q6QSEo6aqao. Checked 12/2014.

AUTODESK. 2014. Autodesk 3D Studio Max - 3D modeling, animation, and rendering software. In URL: http://www.autodesk.com/products/autodesk-3ds-max/overview. Checked 12/2014.

AUTODESK. 2014. Autodesk Maya - Comprehensive 3D animation software. In URL: http://www.autodesk.com/products/autodesk-maya/overview. Checked 12/2014.

AUTODESK. 2014. Autodesk Softimage - Final Release Announcement. In URL: http://www.autodesk.com/products/autodesk-softimage/overview. Checked 12/2014.

AVENDRA, R. 2008. Ragh Rigging Reel 2008. In URL: http://www.youtube.com/watch?v=_tJXELEbeeM. Checked 12/2014.

BADLER, N. I. 1995. A Workshop on Standards for Facial Animation. In ACM SIGGRAPH Newsletter, Computer Graphics Modular Visualization Environments.

BADLER, N. I., MANOOCHEHRI, K. H., AND BARAFF, D. 1986. Multi-Dimensional Input Techniques and Articulated Figure Positioning by Multiple Constraints. In *Proc. Workshop on Interactive 3D Graphics*, pp. 151-159.

BASTOS, P., ALVAREZ, X., AND ORVALHO, V. 2011. A Demo of a Dynamic Facial UI for Digital Artists. In *Proc. of the 13th International Conference on Human-Computer Interaction - INTERACT*, September 5-9, Lisbon, Portugal, pp. 358-359.

BASTOS, P. 2012. Generic and Certified Facial Rig Mechanical Approaches for Key Frame Character Animation. In *Proc. of the 1st International Conference in Illustration and Animation*, Ofir, Portugal, November 30 - December 1, pp. 159-172.

BASTOS, P., ALVAREZ, X., AND ORVALHO, V. 2012. A Demo of a Facial UI Design Approach for Digital Artists. In *Proc. of the 17th International Conference on Intelligent User Interfaces - IUI*, February 14-17, Lisbon, Portugal, pp. 307-308.

BASTOS, P., ALVAREZ, X., AND BARBOSA, N. 2012. Facial Rigging for Key Frame Animation. In *Proc. of the 6th International Conference on Digital Arts*, Faro, Portugal, November 8-9, pp. 337-340.

BASTOS, P. 2013. Easy Character Facial Rig Mechanical Approaches to Animate Zipper and Sticky Lips Effects. In *Proc. of the 2nd International Conference in Illustration and Animation*, Porto, Portugal, November 29 - 30, pp. 81-94.

BASTOS, P., AND FERNANDES, L. 2013. FAC^2RIG: A Complete and Customizable Rig Interface System Approach for Facial Animation. In *International Journal of Advanced Computer Science - IJACSci*, 3(11): 561–569.

BASTOS, P., FERNANDES, L., AND STROVER, S. 2014. Multimodal Facial Rigging: Towards a User-Adapted Animation Interface. In *Proc. of the 8th International Conference on Computer Graphics, Visualization, Computer Vision and Image Processing*, Lisboa, Portugal, July 17 - 19, pp. 197-205.

BASTOS, P., PEREIRA, R. J., PEREIRA, R. A. AND SOUSA, R. L. 2014. Conferência Blender PT 2013 - Promoção, Intercâmbio e Partilha Digital 3D. In *Proc. of Conferência Blender PT 2013,* Universidade Católica Editora - Porto, ISBN 978-989-8366-65-8.

BEELER, T., BICKEL B., BEARDSLEY, P., SUMNER, B., AND GROSS, M. 2010. High-quality Single-shot Capture of Facial Geometry. In *Proc. ACM SIGGRAPH*.

BEESON, C., AND BJORKE, K. 2004. GPU Gems. No. 3. AddisonWesley, Ch. Skin in the “Dawn” Demo.

BERGERON, P., AND LACHAPELLE, P. 1985. Controlling Facial Expressions and Body Movements in the Computer-generated Animated Short 'Tony de Peltrie'. In *Proc. SIGGRAPH Papers*.

BEST, D. 2011. Bone Based Facial Rig. In URL: https://vimeo.com/29832863. Checked 12/2014.

BETHEL, E. W., AND USELTON, S. P. 1989. Shape Distortion in Computer-assisted Key Frame Animation. In *Proc. Computer Graphics International*, pp. 215-224.

BLAIR, P. 1996. Cartoon Animation. Walter Foster Publishing.

BLENDER NATION. 2009. Blender: 3.4 Million downloads per year. In URL: http://www.blendernation.com/2009/05/06/blender-34-million-per-year/. Checked 12/2014.

BLENDER FOUNDATION. 2011. Website statistics for blender.org. In URL: http://www.blender.org/about/website/statistics/. Checked 12/2014.

BLENDER FOUNDATION. 2012. Blender vs the rest - by Ton Roosendaal. In URL: http://www.youtube.com/watch?v=VJIi7IZJc8I. Checked 12/2014.

BLENDER FOUNDATION. 2014. Blender. URL: http://www.blender.org. Checked 12/2014.

BLENDER FOUNDATION. 2014. Blender Documentation Contents. In URL: http://www.blender.org/documentation/250PythonDoc/. Checked 12/2014.

BIBLIOWICZ, J. 2005. *An Automated Rigging System for Facial Animation*. Master's Thesis, Cornell University.

BLUM D. 1998. *Face It!*. Psychology Today, Sussex Publishers, pp. 32-39.

BREDOW, R., SCHAUB, D., KRAMER, D., HAUSMAN, M., DIMIAN, D., AND DUGUID, R. S. 2007. Surf’s Up: The Making of an Animated Documentary. In *Proc. SIGGRAPH Courses*.

BROOKS, D. 2012. David Brooks Character Facial Rigger/TD Aardman 'Pirates: Band of Misfits' Showreel. In URL: https://vimeo.com/36362745. Checked 12/2014.

BUNNELL, M. 2005. *GPU Gems 2: Programming Techniques for High-Performance Graphics and General-Purpose Computation*. No. 7. NVIDIA, Addison-Wesley Professional, Ch. Adaptive Tessellation of Subdivision Surfaces with Displacement Mapping.

BURTNYK, N., AND WEIN, M. 1970. Computer-generated Key Frame Animation. In *Society of Motion Picture and Television Engineers, Inc*. - *SMPTE*, 80(3): 149-153.

BURTNYK, N., AND WEIN, M. 1976. Interactive Skeleton Techniques for Enhancing Motion Dynamics in Key Frame Animation. In *Magazine Communications of ACM*, pp. 564-569.

BURTON, J. 2011. Demo Reel - Summer 2011. In URL: https://vimeo.com/27384539. Checked 12/2014.

CABRERA, C. 2008. *An Essential Introduction to Maya Character Rigging with DVD*. Focal Press Elsevier.

CAMILO, A. 2009. Brown Rig. In URL: https://vimeo.com/3301362#at=0. Checked 12/2014.

CAMPBELL, D. T., AND STANLEY, J. 1963. *Experimental and Quasi-Experimental Designs for Research*. Cengage Learning, Ch. 5, pp. 34-61.

CATMULL, E., AND CLARK, J. 1978. Recursively Generated B-spline Surfaces on Arbitrary Topological Meshes. In *Computer Aided Design*.

CHANG, Y., AND HABIB, A. 2008. Reconstruction Registration and Matching of 3D Facial Models from Stereo-images. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*.

CHEN, F. 2011. Maya Character Facial Rigging Reel Blend Shape. In URL:

http://www.youtube.com/watch?v=Wv-0RbHNxvE. Checked 12/2014.

CHEUNG, C. H. 2009. Chi Hung Cheung Motion Builder Facial Rig. In URL: http://www.youtube.com/watch?v=sbLS-9utqhY. Checked 12/2014.

CONNER, B., SNIBBE, S., HERNDON, K., ROBBINS, D., ZELEZNIK, R., AND VAN DAM A. 1992. Three-dimensional Widgets. In *Proc. of the Symposium on Interactive 3D Graphics, ACM*, pp. 183–188.

COUTINHO, C. P. 2013. *Metodologia de Investigação em Ciências Sociais e Humanas: Teoria e Prática*. 2ª Ed., Coimbra: Almedina, pp. 142-143.

DARWIN, C. 1872. *The Expression of the Emotions in Man and Animals*. Kessinger Publishing, LLC.

DEGRAF, B. 1988. Mike the Talking Head. In *SIGGRAPH Electronic Theatre*.

DEROSE, T., KASS M., AND TRUONG, T. 1998. Subdivision Surfaces in Character Animation. In *Proc. SIGGRAPH*. pp. 85-94.

DEROSE, T. 2013. In URL: http://www.theverge.com/2013/3/7/4074956/pixar-senior-scientist -derose-explains-how-math-makes-movies-games. Checked 12/2014.

DENG, Z., AND NOH, J. 2007. *Computer Facial Animation: A survey*. In Data-Driven 3D Facial Animation, Eds. Springer London, pp. 4-5.

DIAMANT, R., AND SIMANTOV, J. 2010. Uncharted 2 Character Pipeline: An In-depth Look at the Creation of U2's Characters. In *Proc. Game Developers Conference*.

DICKENS, M., SMITH, W., WU, J., AND HANCOCK, E. 2007. Face Recognition Using Principal Geodesic Analysis and Manifold Learning. In *Proc. of the 3rd Iberian Conference on Pattern Recognition and Image Analysis*.

DIKE, C. 2013. Chris Dike Demo Reel 2013 (Animation & Rigging). In URL: http://www.youtube.com/watch?v=gaRKFEF3n84. Checked 12/2014.

DIX, A., FINLAY, J., ABOWD, G. D., AND BEALE, R. 2004. *Human-Computer Interaction*. 3rd

Ed., Harlow: Pearson Education, pp. 240.

DUTREVE, L., MEYER, A., ORVALHO, V., AND BOUAKAZ, S. 2010. Easy Rigging of Face by Automatic Registration and Transfer of Skinning Parameters. In *Proc. International Conference on Computer Vision and Graphics*.

ELYAN, E., AND UGAIL, H. 2007. Reconstruction of 3D Human Facial Images Using Partial Differential Equations. In *Proc. Journal of Computers*.

EKMAN, P., AND FRIESEN, W. 1975. *Unmasking the Face: A Guide to Recognizing Emotions from Facial Clues*. Prentice-Hall, pp. 34-128.

EKMAN P., AND FRIESEN, W. 1978. *Manual of The Facial Action Coding System*. Consulting Psychologist Press, Palo Alto, CA, USA, pp. 15-347.

EREN, K. 2011. FaceRig. In URL: https://vimeo.com/22571532. Checked 12/2014.

EVANS, C., AND HAGEDORN, H. 2007. Character Pipeline Advances for Next-Gen Games. In *Proc. Game Developers Conference*.

FAIGIN, G. 1990. *The Artist's Complete Guide to Facial Expressions*. Watson-Guptill. pp. 95.

FAILES, I. 2014. Age of Extinction: ILM turns up its Transformers toolset. In URL: http://www.fxguide.com/featured/age-of-extinction-ilm-turns-up-its-transformers-toolset/. Checked 12/2014.

FALK, R., MINTER, D., VERNON, C., ARETOS, G., MODESTO, L., LAM-ORLETTE, A., WALKER, N., CHEUNG, T., RENTEL-LAVIN, J., AND MAX, H. 2004. Art-directed Technology: Anatomy of a Shrek 2 Sequence. In SIGGRAPH Course Notes, pp. 37.

FARKAS, M. D., AND LESLIE, G. 1994. *Anthropometry of the Head and Face*. Raven Press.

FEDOR, P., SUH, H., STAHLBERG, S., HARTLE, M., AND LEVIUS, P. 2007. *Essence: The Face: Modeling and Texturing*. Ballistic Publishing.

FLEMING, B. AND DOBBS, D. 1998. *Animating Facial Features and Expressions*, Charles River Media.

GEORGIEV, D. 2007. Face Rig. In URL: http://www.youtube.com/watch?v=llDSg5JKHas. Checked 12/2014.

GIBBS, J. K., AND DERAKHSHANI, D. 2005. *Maya: Secrets of the Pros*, 2nd Ed. Sybex.

GLEICHER, M. 1998. Retargeting Motion to New Characters. In *Proc. SIGGRAPH, ACM Press*, New York, USA, pp. 33-42.

GORDEN, J. 2005. *LightWave 3D (8) Cartoon Character Creation Volume 1: Modeling and Texturing*. Wordware Publishing.

GORDEN, J. 2005. *LightWave 3D (8) Cartoon Character Creation Volume 2: Rigging and Animation*. Wordware Publishing.

GRAFT, L., DWELLY, B., RIESBERG, D., AND MOGK, C. 2002. *Learning Maya: Character Rigging and Animation*. Alias Wavefront Education.

GRANBERG, C. 2009. *Character Animation with Direct 3D*. Charles River Media.

GRAY, H. 1974. *Gray's Anatomy*. Running Press, U.S.

GRUBB, W. 2009. Facial Animation Rig for Delgo. In URL: http://www.creativecrash.com/maya/tutorials/character/c/facial-animation-rig-for-delgo. Checked 12/2014.

HAMMOND, M. 2009. Blender Girl Rig Demo. In URL: http://vimeo.com/channels/rigging/3804230. Checked 12/2014.

HARKINS, J. 2005. *Maya Techniques: Super Toon Facial Rigging*. Alias, Training DVD. In URL: http://forums.cgsociety.org/showthread.php?t=295301 and in URL: http://downloads.alias.com/gcs/edu-mt_st_facial_rigging_sample.mov. Checked 12/2014.

HAUSER, T. 2011. *The Pixar Treasures*. Disney Editions, Deluxe Scrapbook.

HENNE, M., HICKEL, H., JOHNSON, E., AND KONISHI, S. 1996. The Making of Toy Story. In *IEEE International Computer Conference*, pp. 463-468.

HEWITT, K. 2011. Female Facial Rig Demo. In URL: https://vimeo.com/18057764. Checked

12/2014.

HOLLY, A. 2006. *Facial Rigging for Feature Animation*. Fahrenheit Digital Training DVD. In URL: http://www.verycd.com/topics/88913/. Checked 12/2014.

HOLLYWOOD. 2014. How They Created 'The Lego Movie'. In URL: http://www.youtube.com/watch?v=aq_tqsasgoI. Checked 12/2014.

JACK, R. E., GARROD, O. G., AND SCHYNS, P. G. 2014. Dynamic Facial Expressions of Emotion Transmit an Evolving Hierarchy of Signals Over Time. In *Current Biology Journal*, Vol. 24, Issue 2, pp. 187.

KAISER, M., STORMER, A., ARSIC, D., AND RIGOLL, G. 2010. Non-rigid Registration of 3D Facial Surfaces with Robust Outlier Detection. In *Applications of Computer Vision*.

KAZUO, R., AND PINHEIRO, D. 2011. *3D Total's Introduction to Rigging - Maya*. 3D Total eBooks Series. In URL: http://shop.3dtotal.com/training/training-ebooks-maya/intro-to-rigging-maya-ebook.html#t2 and in URL: http://pt.scribd.com/doc/57410230/3D-Total-s-Introduction-to-Rigging-MAYA. Checked 12/2014.

KLEIN, A. 2010. Thirtyfour facial rig. In URL: https://vimeo.com/14950162. Checked 12/2014.

KOCHANEK, D., AND BARTELS, R. 1984. Interpolating Splines for Key Frame Animation. In *National Computer Graphics Association of Canada Conference – Graphics Interface*, pp. 41–42.

KOMATSU, K. 1988. Human Skin Model Capable of Natural Shape Variation. In *The Visual Computer*.

KOMBAN, L. 2013. Elliot Face Rig. In URL: http://www.youtube.com/watch?v=LK-akOSvvpg. Checked 12/2014.

KOMOROWSKI, D., MELAPUDI, V., MORTILLARO, D., AND LEE, G. S. 2010. A Hybrid Approach to Facial Rigging. In *Proc. SIGGRAPH Sketches*.

KUMAR, B. 2013. Facial Rigging Maya 2013. In URL: http://www.youtube.com/watch?v=sLw2yN6vD24. Checked 12/2014.

LABOURLETTE, C., CANI, M.-P., AND ARNALDI, B. 2005. Dynamic Skinning: Adding Real-Time Dynamic Effects to an Existing Character Animation. In *Spring Conference on Computer Graphics*, pp. 87-93.

LABS, N. 2010. Face Rig Maya - Designed for a Game Engine. In URL: http://www.youtube.com/watch?v=pYDuF8DTOJM. Checked 12/2014.

LALLY, J. 2013. John Lally's Character TD Reel. In URL: http://www.youtube.com/watch?v=Gq0ZkblAySg#t=109. Checked 12/2014.

LASSETER, J. 1987. Principles of Traditional Animation Applied to 3D Computer Animation. In *Proc. SIGGRAPH*, pp. 35-44.

LEE, Y., TERZOPOULOS, D., AND WATERS, K. 1995. Realistic Modeling for Facial Animation. In *Proc. SIGGRAPH*.

LEITE, L. 2007. Marionetas Virtuais - Animação Interativa em Tempo Real. Master's Thesis. Faculdade de Engenharia da Universidade do Porto.

LEWIS, J. P., CORDNER, M., AND FONG, N. 2000. Pose Space Deformation: A Unified Approach to Shape Interpolation and Skeleton-driven Deformation. In *Proc. SIGGRAPH*, pp. 165-172.

LEWIS, J. P., AND ANJYO, K. 2010. Direct Manipulation Blendshapes. In *Journal IEEE on Computer Graphics and Applications*, Vol. 30, Issue 4, pp. 42-50.

LI, H., WEISE, T., AND PAULY, M. 2010. Example-based Facial Rigging. In *Proc. SIGGRAPH Papers*.

LIM, P. 2007. Simple Facial Rigging using Morph Blending. In URL: http://www.youtube.com/watch?v=NlxdDXd1kyI. Checked 12/2014.

LOMBARD, P. 2013. Blender Hybrid Face Rigging. In URL:

http://www.youtube.com/watch?v=flKz3h5m8HU. Checked 12/2014.

MADDOCK, S., EDGE, J., AND SANCHEZ, M. 2005. Movement Realism in Computer Facial Animation. In *Proc. of the 19th British HCI Group Annual Conference, Workshop on Human-animated Characters Interaction*.

MAESTRI, G. 1996. *Digital Character Animation*. New Riders Publishing.

MAESTRI, G. 1999. *Digital Character Animation 2, Volume I: Essential Techniques*. New Riders Publishing.

MAESTRI, G. 2001. *Digital Character Animation 2, Volume II: Advanced Techniques*. New Riders Publishing.

MAESTRI, G. 2006. *Digital Character Animation 3*. New Riders Publishing.

MAHLER, T. 2010. Basil Facial Animation Test. In URL: http://www.youtube.com/watch?feature=player_embedded&v=0nj1VbN6HV8. Checked 12/2014.

MARAFFI, C. 2003. *Maya Character Creation: Modeling and Animation Controls*. New Riders Publishing.

MARMOR, D. 2011. Blend shapes. In URL: http://www.youtube.com/watch?v=KPDfMpuK2fQ. Checked 12/2014.

MARTE, D. 2009. The Librarian - Rig Demonstration. In URL: http://www.youtube.com/watch?v=Esn-cI-AbqQ. Checked 12/2014.

MAXON. 2014. Cinema 4D. In URL: http://www.maxon.net. Checked 12/2014.

MCCLOUD, S. 1993. *Understanding Comics*. Tundra Publishing.

MCLAUGHLIN, T. 2006. Taxonomy of Digital Creatures: Defining Character Development Techniques Based upon Scope of Use. In *Proc. SIGGRAPH Courses*.

MCLAUGHLIN T., AND SUMIDA, S. 2007. The Morphology of Digital Creatures. In *Proc. SIGGRAPH Courses*.

MCLAUGHLIN, T., CUTLER, L., AND COLEMAN, D. 2011. Character Rigging, Deformations and Simulations in Film and Game Production. In *Proc. SIGGRAPH Courses*.

MEYER, P. 2013. Investing in the Future of Animation. In URL: http://www.moneyweb.co.za/moneyweb-soapbox/investing-in-the-future-of-animation. Checked 12/2014.

MILLER, E. 2004. *Hyper Real Facial Setup*. Alias Training DVD.

MILLER, E., AND GROSSMAN, R. 2006. *Autodesk Maya Techniques: Hyper Realistic Creature Anatomy I*. Autodesk Training DVD.

MILLER, M., ARIKAN, O., AND FUSSELL, D. 2010. Frankenrigs: Building Character Rigs from Multiple Sources. In *Proc. SIGGRAPH Symposium on Interactive 3D Graphics and Games*.

MONTESARCHIO, D. 2012. Face Rigging. In URL: https://vimeo.com/46006859. Checked 12/2014.

MORI, M. 1970. *The Uncanny Valley*. In Energy, 7(4).

MULLEN, T. 2007. *Introducing Character Animation with Blender*. Wiley Publishing, pp. viii.

NEALE, P. 2008. *Intermediate Rigging 4: Facial Rigging Techniques*. CG Academy, Training DVD. In URL: http://www.youtube.com/watch?v=FST2Bb76OdI&list=PL3C6BDFAA6 84F5EEC. Checked 12/2014.

NEBULA, C. 2013. Colin Nebula Facial Rigging Animation Demo. In URL: http://www.youtube.com/watch?v=lIrnDytiNxA. Checked 12/2014.

NEUFFER, D. 2011. Facial Rigging. In URL: https://vimeo.com/28916898. Checked 12/2014.

NEWTEK. 2014. LightWave. In URL: https://www.lightwave3d.com/. Checked 12/2014.

NIELSON, G., AND OLSEN, D. 1986. Direct Manipulation Techniques for 3D Objects Using 2D Locator Devices. In *Proc. Symposium on Interactive 3D Graphics*, pp. 175-182.

NIELSEN, J. 1992. *Usability Engineering*. Academic Press, New York.

NITTI, A. 2009. Character/TD Rigging Reel - Alessio Nitti 2009. In URL: http://www.youtube.com/watch?v=XFtGQEzGQuc. Checked 12/2014.

NOGUEIRA, F. 2010. DemoReel 2010. In URL: https://vimeo.com/31986339. Checked 12/2014.

OLIVER, P., PERNA, C., WALKER, J., CLAUSEN, M., BELICA, Z., MORAN, G., AND HOSFELT, J. 2011. Samaritan Real-time Demo. In *Proc. SIGGRAPH Computer Animation Festival*.

OLIVERIO, G. 2007. *Maya 8 Character Modeling*. Wordware Publishing.

O'MALLEY, C. D. 1983. *Leonardo Da Vinci on the Human Body: The Anatomical, the Physiological and Embryological Drawings of Leonardo Da Vinci*. Crown Publications.

O'NEILL, R. 2008. *Digital Character Development: Theory and Practice*. Morgan Kaufmann, Elsevier.

ORVALHO, V. 2007. *Reusable Facial Rigging and Animation: Create Once, Use Many*. Ph.D. Thesis in Computer Graphics, Universitat Politècnica de Catalunya, Barcelona, Spain.

ORVALHO, V., BASTOS, P., OLIVEIRA, B., AND ALVAREZ, X. 2012. A Facial Rigging Survey. In *Proc. of the 33rd Annual Conference of the European Association for Computer Graphics - EUROGRAPHICS*, May 13-18, Cagliari, Italy, Volume 32, pp. 10-32.

OSIPA, J. 2003. *Stop Staring: Facial Modeling and Animation Done Right*, 1st Ed. Wiley Publishing, Sybex.

OSIPA, J. 2007. *Stop Staring: Facial Modeling and Animation Done Right*, 2nd Ed. Wiley Publishing, Sybex.

OSIPA, J. 2010. *Stop Staring: Facial Modeling and Animation Done Right*, 3rd Ed. Wiley Publishing, Sybex.

OZTURK, H. 2007. Huseyin Ozturk Demoreel 2007. In URL: http://www.youtube.com/watch?v=RBlnq5ggA9s. Checked 12/2014.

PAGORIA, C. 2009. Chris Pagoria Facial Rigging Reel. In URL: https://vimeo.com/3899752.

Checked 12/2014.

PANDZIC, I. S., AND FORCHHEIMER, R. 2002. *MPEG-4 Facial Animation: The Standard, Implementation and Applications*. John Wiley & Sons, pp. 20-21.

PARDEW, L. 2008. *Character Emotion in 2D and 3D Animation*. Delmar Cengage Learning.

PARK, I. K., ZHANG, H., AND VEZHNEVETS, V. 2005. Image-based 3D Face Modeling System. In *Journal on Applied Signal Processing*.

PARKE, F. I. 1972. Computer Generated Animation of Faces. In *Proc. ACM National Conference*.

PARKE, F. I. 1974. *A Parametric Model for Human Faces*. Ph.D. Thesis, University of Utah, Salt Lake City.

PARKE, F. I., AND WATERS, K. 2008. *Computer Facial Animation*. A.K. Peters. 2nd Edition.

PARKS, C. 2003. *Secrets to Drawing Realistic Faces*. North Light Books.

PATNODE, J. 2008. *Character Modeling with Maya and ZBrush*. Focal Press.

PELACHAUD, C., BADLER, N. I., AND VIAUD, M-L. 1994. Final Report to NSF of the Standards for Facial Animation Workshop. Technical Report, University of Pennsylvania.

PERLIN, K. 1997. Layered Compositing of Facial Expression. In *Proc. SIGGRAPH*, URL: http://mrl.nyu.edu/~perlin/experiments/facedemo. Checked 12/2014.

PESSA, J. 2012. *Facial Topography: Clinical Anatomy of the Face*. Quality Medical Publishing.

PLATT, S. M. AND BADLER, N. I. 1981. Animating Facial Expressions. In *Proc. SIGGRAPH, ACM Press*, New York, NY, USA, pp. 245-252.

PORTER, T. 1997. Creating Lifelike Characters in Toy Story. In *Newsletter of ACM SIGART Bulletin*.

PRICE, D. 2008. *The Pixar Touch: The Making of a Company*. Knopf Publishers.

RAITT, B., AND MINTER, G. 2000. How to Apply the Principles of Traditional Sculpture to

Make Stunning 3D Characters. In URL: http://theminters.com/misc/articles/derived-surfaces/. Checked 12/2014.

RAITT, B. 2004. The Making of Gollum. In *Frontiers of Facial Animation Workshop*, University of Southern California, Institute for Creative Technologies.

REEVES, W., OSTBY, E., AND LEFFLER, S. 1990. The Menv Modelling and Animation Environment. In *The Journal of Visualization and Computer Animation*, Vol. 1, pp. 33-40.

RITCHIE, K., CALLERY, J., AND BIRI, K. 2005. *The Art of Rigging: A Definitive Guide to Character Technical Direction with Alias Maya*, Vol. 1. CG ToolKit, Alias. In URL: http://pt.scribd.com/doc/35580059/Art-of-Rigging-Vol-I-Book-CG-Toolkit and in URL: http://www.kenstojevich.com/rigging.html. Checked 12/2014.

ROSA, G. 2011. Exemplo de Rigging Facial Simples usado em aula - Morph Version. In URL: http://www.youtube.com/watch?v=yJCGUgrruT0. Checked 7/2014.

ROSE, R., JUTAN, M., AND DOUBLESTEIN, J. 2013. BlockParty 2: Visual Procedural Rigging for Film, TV and Games. In *Proc. SIGGRAPH Talks*.

SANGUIGNI, T. 2010. Face Rigging Setup. In URL: http://www.youtube.com/watch?v=IHyKBjjZd7E. Checked 12/2014.

SANTANDER, I. 2010. Mario Facial Rig + morphConnector. In URL: https://vimeo.com/14949099. Checked 12/2014.

SAWALHA, A. 2010. Facial Rigging Maya 2011. In URL: http://www.youtube.com/watch?v=0GIW8fUzuA4. Checked 12/2014.

SCHERER, K. AND EKMAN, P. 1984. *Approaches to Emotion*. Hillsdale, NJ: Lawrence Erlbaum.

SCHLEIFER, J., SCADUTO-MENDOLA, R., CANETTI, Y., AND PIRETTI, M. 2002. Character Setup from Rig Mechanics to Skin Deformations: A Practical Approach. In *Proc. SIGGRAPH*

Course Notes.

SCHILLER, M. 2009. Matt Schiller Rigging Reel. In URL: http://www.youtube.com/watch?v=ytW8UW8yASc. Checked 12/2014.

SCRIPTATTACK, AND FX. 2014. Bony Face: Facial Animation System for 3dsMax. In URL: http://www.scriptattack.com/maxscripts/bonyface/index_eng.html. Checked 12/2014.

SEDERBERG, T. W., AND PARRY, S. R. 1986. Free-form Deformation of Solid Geometric Models. In *Proc. SIGGRAPH*, pp. 151-160.

SEEMANN, M. 2011. *Dependency Injection in .Net*. Manning Publications Co., pp. 175-178.

SENG, E. 2013. Rigging and Animation Demo Reel 2013. In URL: http://www.youtube.com/watch?v=jBOEzXYMugE. Checked 12/2014.

SEPULVEDA, P. 2012. Intro Rigging - Animschool. In URL: https://vimeo.com/29491242. Checked 12/2014.

SHENG, Y., KUCHARSKI, K., SADKA, A., AND SKARBEK, W. 2006. Automatic Face Synthesis and Analysis: A Quick Survey. In *Proc. International Conference on Computer Vision and Graphics*.

SIMANTOV, J. 2013. Character Rigging and Modeling in Naughty Dog's The Last of Us. In URL: http://www.youtube.com/watch?v=myZcUvU8YWc. Checked 12/2014.

SINGH, K., AND FIUME, E. 1998. Wires: A Geometric Deformation Technique. In *Proc. SIGGRAPH*, pp. 405-414.

SKONICKI, M. 2008. *Facial Rigging for Games*. Eat3D, Training DVD. In URL: http://www.youtube.com/watch?v=Y_q_2GS9akY. Checked 12/2014.

SNIBBE S. 1995. A Direct Manipulation Interface for 3D Computer Animation. In *Computer Graphics Forum*, Vol. 14, Wiley Online Library, pp. 271-283.

SOTO, J. 2012. Facial Rig Test. In URL: https://vimeo.com/46499529. Checked 12/2014.

STERLING, S. 2012. Alf Facial Rig. In URL: https://vimeo.com/45228727. Checked 12/2014.

STOIBER, N., SEGUIER, R., AND BRETON, G. 2009. Automatic Design of a Control Interface for a Synthetic Face. In *Proc. of the 14th International Conference on Intelligent User Interfaces - IUI*.

STRATEN, A. 2011. PSD Tool Demo. In URL: https://vimeo.com/25223035. Checked 12/2014.

STRAUSS, P., AND CAREY, R. 1992. An Object-oriented 3D Graphics Toolkit. In *ACM SIGGRAPH Computer Graphics*, 26, 2, pp. 341-349.

STRAUSS, P. S. 2002. The Design and Implementation of Direct Manipulation in 3D. In *Proc. SIGGRAPH Course Notes*.

STURMAN, D. 1998. The State of Computer Animation. In *ACM SIGGRAPH Computer Graphics*, 32, 1, pp. 57-61.

SUISSA, G. 2013. Gary Suissa Facial Rigging. In URL: http://www.youtube.com/watch?v=lwzPsWGanNU. Checked 12/2014.

TALBOT, J. 2003. Character Articulation Facial Setup. Lucas Entertainment Company.

TEAMVIEWER. 2014. TeamViewer - Free Remote Control, Remote Access & Online Meetings. In URL: http://www.teamviewer.com/. Checked 12/2014.

THOMAS, F., AND JOHNSTON, O. 1981. *Disney Animation, The Illusion of Life*. Abbeville Press.

TIAN, Y., KANADE, T., AND COHN, J. F. 2001. Recognizing Action Units for Facial Expression Analysis. In *Proc. of the IEEE Transactions on Pattern Analysis and Machine Intelligence*.

TORRES, J. 2012. Face rig animation setup 50 seconds. In URL: http://www.youtube.com/watch?v=p3PFxjLFgjo. Checked 12/2014.

TRAZOS, L. 2012. Facial Rig Test de Eideann, trabajo de Trazos. In URL: https://vimeo.com/39620711. Checked 12/2014.

UNAY, J. 2004. Hyper-real Modeling. In *Proc. SIGGRAPH Course Notes*.

UNAY, J., AND GROSSMAN, R. 2005. Hyper-real Advanced Facial Blendshape Techniques and Tools for Production. In *Proc. SIGGRAPH Master Classes*.

VASCONCELOS, V. 2011. *Blender 2.5 Character Animation Cookbook*. Packt Publishing. In URL: http://virgiliovasconcelos.com/blender-animation-cookbook/. Checked 12/2014.

VAUCHER, R. 2011. Face Rig Demo. In URL: https://vimeo.com/3605280. Checked 12/2014.

VEGDAHL, N. 2010. Durian Project: Sintel's Face Rig. In URL: http://www.youtube.com/watch?v=o3LFZJjfRI8. Checked 12/2014.

VERNON, C. 2010. Updated Sticky Lips Deformer. In URL: http://www.youtube.com/watch?v=BzOdNmS9aRo. Checked 12/2014.

VILLAGRASA, S., AND SUSIN, A. 2009. Face! 3D Facial Animation System Based on FACS. In *IV Ibero-American Symposium in Computer Graphics*.

WATERS, K. 1987. A Muscle Model for Animating Three Dimensional Facial Expression. In *Proc. SIGGRAPH*.

WAVEFRONT. 2003. *The Art of Maya: An Introduction to 3D Computer Graphics*. Sybex, 1st Ed.

WEISE, T., BOUAZIZ, S., LI, H., AND PAULY, M. 2011. Realtime Performance-based Facial Animation. In *Proc. SIGGRAPH Papers*.

WIDANAGAMAACHCHI, W., AND DHARMARATNE, A. 2007. Single-patch NURBS Face. In *Signal-Image Technologies and Internet-Based Systems*.

WIDANAGAMAACHCHI, W., AND DHARMARATNE, A. 2008. 3D Face Reconstruction from 2D Images A survey. In *Computing: Techniques and Applications*.

WILLIAMS, L. 1990. Performance-driven Facial Animation. In *Proc. of the 17th Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH*, pp. 235-242.

WILLIAMS, R. 2001. *The Animator's Survival Kit*. Faber and Faber.

WITKIN, A., AND KASS, M. 1988. Spacetime Constraints. In *Proc. SIGGRAPH*.

YANG, X., AND ZHANG, J. 2006. Stretch It - Realistic Smooth Skinning. In *Proc. of the International Conference on Computer Graphics, Imaging and Visualization, IEEE Computer Society*.

ZELTZER, D. 1982. Motor Control Techniques for Figure Animation. In *Journal IEEE on Computer Graphics and Applications*, Vol. 2, Issue 9, pp. 53-59.

ZHENG, K. 2009. Character Rigging Demo Reel. In URL: http://www.youtube.com/watch?v=Hg6jvHjvHGk. Checked 12/2014.

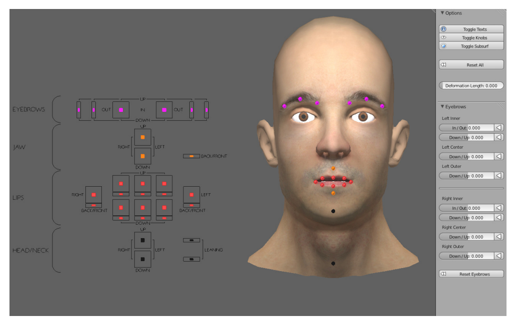
Appendix

A) Publications

The following pages present the list of publications and respective abstracts that resulted from this research. These results are available in PDF format in the accompanying CD in the folder *publications*. This folder also contains videos that illustrate the papers A-2014, A-2013-1, A-2013-2, A-2012-4 and A-2011.

2014

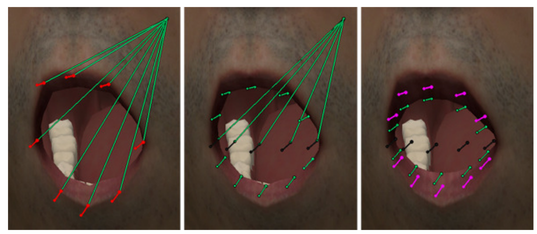
BASTOS, P., FERNANDES, L., AND STROVER, S. 2014. Multimodal Facial Rigging: Towards a User-Adapted Animation Interface. In *Proc. of the 8th International Conference on Computer Graphics, Visualization, Computer Vision and Image Processing*, Lisboa, Portugal, July 17 - 19, pp. 197-205. (see A-2014-video)



Animating a character's face is a strenuous task due to a lack of a consensus among 3D artists on how to build and provide the controls in an animation-ready facial rig. To help provide a uniform integrative of control interaction to facial animation, a multimodal rig interface system approach was developed and tested in an initial user experiment with 15 digital artists. The users requested improvements and the system was upgraded into a user-adapted multimodal approach that supports a synchronized input of its different interaction modes and a scalable control of the facial deformation. This paper describes the algorithms behind the user-adapted system and a follow-up user experiment with 20 professional expert character animators. The results and user scores help clarify the relevance of the user-adapted multimodal rig interface system approach as an interaction model that decreases the effort required to animate a character's face.

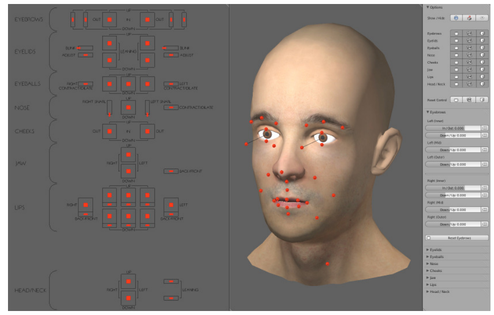
2013

[1] BASTOS, P. 2013. Easy Character Facial Rig Mechanical Approaches to Animate Zipper and Sticky Lips Effects. In *Proc. of the 2nd International Conference in Illustration and Animation*, Porto, Portugal, November 29 - 30, pp. 81-94. (see A-2013-1-video)



Facial rig mechanics are the underlying structures in a character's face that allow it to be animated for use in digital media such as films, videogames or advertising. These structures are built and maintained by riggers so that animators can deform a character's face by triggering user interface controls. The increasing demand for more quality facial animation in the entertainment industry led digital artists to develop a diversity of rig mechanical approaches over the last 15 to 20 years. But there are special cases which are rarely developed, namely the zipper and sticky lips deformations, which are fundamental effects for realistic character facial animation. This paper presents easy rig mechanical approaches for these effects, which are realized as optimized procedures to provide a consensus to their rig construction process. These approaches were presented to expert digital artists who have evaluated them according to a set of parameters. The results reveal that these approaches ease and improve the rigging of the zipper and sticky lips effects.

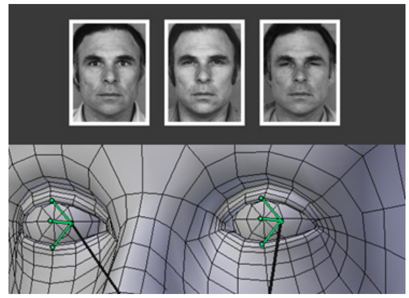
[2] BASTOS, P., AND FERNANDES, L. 2013. FAC^{2}RIG: A Complete and Customizable Rig Interface System Approach for Facial Animation. In *International Journal of Advanced Computer Science - IJACSci*, 3(11):561-569. (see A-2013-2-video)



A character's face can be manipulated by an animator via complex rig controls that allow to pose and to key frame facial behaviors in the timeline of a 3D tool. Setting the facial rig controls is a key process that riggers execute to ensure animators have a pleasant and precise control of a character's face. But the process of setting the controls is often laborious, time-consuming, diverse and erratic, because controls need to be effective in terms of the deformation generated in the facial skin and intuitive in terms of the animator interaction. In addition, there is uncertainty amongst digital artists regarding how a facial rig can be built for animation; each rigger has its own approach to build and to provide the control of a facial rig. This poses an interesting question: is it possible to optimize the setup of facial rig controls in order to achieve appealing and accurate key frame facial animation? This paper presents a solution: a complete and user customizable facial rig interface system based on a multimodal approach as a result of previous research demonstrating that this topic is open to improvements. A validation of this system has been carried out with 15 users who are 3D digital artists and the results reveal that this approach can be a relevant contribution for character facial rigging and animation.

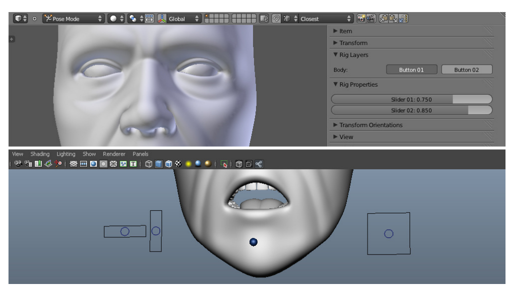
2012

[1] Bastos, P. 2012. Generic and Certified Facial Rig Mechanical Approaches for Key Frame Character Animation. In *Proc. of the 1st International Conf. in Illustration and Animation*, Ofir, Portugal, November 30 - December 1, pp. 159-172.



A character facial rig is a set of mechanical and control structures that allow an animator to pose and key frame a character's face in a timeline. The setup of the rig mechanics is a key task that digital artists perform to assure subtle facial skin deformations resulting from the manipulation of the rig controls. But setting the facial rig mechanics is a laborious and time-consuming process because artists need to prevent awkward facial deformations. This poses an interesting question: is it possible to improve the setup of facial rig mechanics for key frame character facial animation? A generic and certified approach is presented in this paper.

[2] Bastos, P., Alvarez, X., and Barbosa, N. 2012. Facial Rigging for Key Frame Animation. In *Proc. of the 6th International Conf. on Digital Arts*. Faro, Portugal, November 8-9, pp. 337-340.

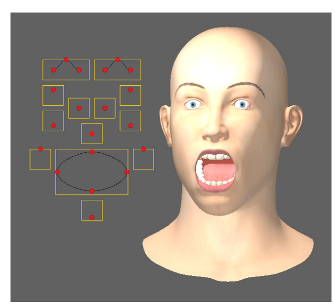


Facial rigging involves planning and setting the mechanics and controls to animate a character's face for films and videogames. This paper presents the main concepts and techniques involved in the process of facial rigging and explains how artists use them to rig the face of 3D characters for key frame animation.

[3] ORVALHO, V., BASTOS, P., OLIVEIRA, B., AND ALVAREZ, X. 2012. A Facial Rigging Survey. In *Proc. of the 33rd Annual Conf. of the European Association for Computer Graphics - EUROGRAPHICS*, May 13-18, Cagliari, Italy, Volume 32, pp. 10-32.

Rigging is the process of setting up a group of controls to operate a 3D model, analogous to the strings of a puppet. It plays a fundamental role in the animation process as it eases the manipulation and editing of expressions, but rigging can be very laborious and cumbersome for an artist. This difficulty arises from the lack of a standard definition of what is a rig and the multitude approaches on how to setup a face. This survey presents a critical review on the fundamentals of rigging, with an outlook of the different techniques, their uses and problems. It describes the main problems that appear when preparing a character for animation. This paper also gives an overview of the role and relationship between the rigger and the animator. Continues with an exhaustive analysis of the published literature and previous work, centered on the facial rigging pipeline. Finally, the survey discusses future directions of facial rigging.

[4] BASTOS, P., ALVAREZ, X., AND ORVALHO, V. 2012. A Demo of a Facial UI Design Approach for Digital Artists. In *Proc. of the 17th International Conf. on Intelligent User Interfaces - IUI*, February 14-17, Lisbon, Portugal, pp. 307-308. (see A-2012-4-video)



In the character animation industry, animators use facial UI's to animate a character's face. A facial UI provides widgets and handles that the animator interacts with to control the character's facial regions. This paper presents a facial UI design approach to control the animation of the six basic facial expressions of the anthropomorphic face. The design is based in square shaped widgets holding circular handles that allow the animator to produce the muscular activity relative to the basic facial expressions. We have implemented a prototype of the facial UI design in the Blender open-source animation software and did a preliminary pilot study with three animators. Two parameters were evaluated: the number of clicks and the time taken to animate the six basic facial expressions. The study reveals there was little variation in the values each animator marked for both parameters, despite the natural difference in their creative performance.

2011

BASTOS, P., ALVAREZ, X., AND ORVALHO, V. 2011. A Demo of a Dynamic Facial UI for Digital Artists. In *Proc. of the 13th International Conf. on Human-Computer Interaction - INTERACT*, September 5-9, Lisbon, Portugal, pp. 358-359. (see A-2011-video)

Character facial animation is difficult because the face of a character assumes many complex expressions. To achieve convincing visual results for animation, 3D digital artists need to prepare their characters with sophisticated control structures. One of the most important techniques to achieve good facial animation is to use facial control interfaces, also called facial user interfaces, or facial UI' s. But facial UI' s are usually dull and often confusing, with limited user interaction and no flexibility. We developed a concept and a working prototype of a dynamic facial UI inside the Blender open-source software to allow their large community of digital artists to better control and organize the facial animation of a character. Our interactive system is running stable in the latest version of Blender and we started to build a full-face dynamic UI to show its interactive potential in a character' s face.

B) Consent Sheets

Following is the contents of the consent sheet signed by the users in order to participate in the user experiment described in Section 5.1 of Chapter 5.

This consent is part of the Ph.D. research of Pedro Bastos of the Doctoral Program in Digital Media (PDMD) of the Faculty of Engineering of the University of Porto (FEUP). The reader is hereby invited to participate in a scientific experiment to test a novel facial rigging and animation system. Read the following information and feel free to ask questions before you decide if you are willing to participate.

- This experiment is voluntary and you have the right to interrupt it at any moment and for any reason. The total duration expected for the experiment is of **30** (thirty) minutes in the training stage and **10** (ten) minutes in the tasking stage.
- This experiment is recorded in video and audio only as a reference for the purpose of this investigation. You have the right to revoke the permission of recording and/or terminate the experiment at any moment.

The reader declares to be aware of the conditions described above, that any questions the reader may have were clarified, agrees to participate in this study and gives permission for the experiment to be recorded in video and audio.

User name ____

User signature ___________________________________ Date __________

Researcher name ____

Researcher signature ______________________________ Date __________

Following is the contents of the consent sheet signed by the users in order to participate in the user experiment described in Chapter 6.

This consent is part of the Ph.D. research of Pedro Bastos of the Doctoral Program in Digital Media (PDMD) of the Faculty of Engineering of the University of Porto (FEUP). The reader is hereby invited to participate in a scientific experiment to test conventional rig interaction modes for 3D character facial animation. Read the following information and feel free to ask questions before you decide if you are willing to participate.

- This experiment is voluntary and you have the right to interrupt it at any moment and for any reason. The total duration expected for the experiment is of **30** (thirty) minutes in the training stage and **15** (fifteen) minutes in the tasking stage.

The reader declares to be aware of the conditions described above, that any questions the reader may have were clarified and agrees to participate in this study.

User name ____

User signature ____ Date __________

Researcher name ____

Researcher signature ____ Date __________

Following is the contents of the consent sheet signed by the users in order to participate in the focus group session of the user experiment described in Chapter 6.

This consent is part of the Ph.D. research of Pedro Bastos of the Doctoral Program in Digital Media (PDMD) of the Faculty of Engineering of the University of Porto (FEUP). The reader is hereby invited to participate in a focus group session with the purpose to further help define future guidelines to improve the design of facial rig controls. Read the following information and feel free to ask questions before you decide if you are willing to participate.

- This session is voluntary and you have the right to interrupt it at any moment and for any reason. The total duration expected for this session is of **45** (forty five) minutes.
- This group session is recorded in audio only as a reference for the purpose of this investigation. You have the right to revoke the permission of audio recording and/or terminate the experiment at any moment.

The reader declares to be aware of the conditions described above, that any questions the reader may have were clarified, agrees to participate in this study and gives permission for this session to be recorded in audio.

User name ____

User signature ___________________________________ Date __________

Researcher name ____

Researcher signature _______________________________ Date __________

C) Questionnaires

Following is the questionnaire completed by the users to collect their professional data before participating in the user experiments described in Sections 5.1 and 5.2 of Chapter 5.

This research is part of the Doctoral Program in Digital Media (PDMD) of the Faculty of Engineering of the University of Porto (FEUP). The purpose of this questionnaire is to collect data from the users who will test a novel facial rigging and animation system.

Name: ____

Nationality: ______________________________ Age: __________

1. What is your education level?

☐ School ☐ High School ☐ College Degree ☐ Masters Degree ☐ Ph.D. Degree

2. Have you attended any specific education in the area of 3D animation?

☐ Yes ☐ No

3. If you answered "**Yes**" in the previous question: what was the duration (hours)?

☐ Until 30 ☐ From 31 to 60 ☐ From 61 to 90 ☐ From 91 to 120 ☐ > 120 hours

4. Do you have a job in the field of 3D animation?

☐ Yes ☐ No

5. Name one or more animation projects you were involved which you consider most significant.____

____

6. Of the following 3D applications select the **ONE** which you **most use**:

☐ 3ds Max ☐ Blender ☐ Cinema 4D ☐ Maya ☐ Softimage Other: __________

____

7. In average, how many **hours** do you spend using a 3D application **per day**?

☐ Until 1 ☐ From 1 to 3 ☐ From 3 to 5 ☐ From 5 to 7 ☐ > 7 hours

8. Which of the following sentences <u>better describes</u> **rigging**?

☐ Building interfaces for animation.

☐ Preparing characters for animation using sophisticated manipulation structures.

☐ Planning and building the deformation mechanics and the control interfaces to animate an object or a character.

9. How many animation-ready **rigs** have you built for an object or a character?

☐ None ☐ 1 to 5 ☐ 6 to 10 ☐ 11 to 20 ☐ More than 20

10. Have you built any type of **rig** for a character's <u>face</u>?

☐ Yes

☐ No If you answered “**No**” state here the reason why ______________________________

____

Following is the questionnaire completed by the users to collect their data before participating in the user experiment described in Chapter 6.

This research is part of the Doctoral Program in Digital Media (PDMD) of the Faculty of Engineering of the University of Porto (FEUP). The purpose of this questionnaire is to collect data from the users who will test conventional rig interaction modes.

Name: ____

Nationality: ______________________________________ Age: _____________

1. What degree are you attending?

____

2. Do you attend a course of 3D?

☐ Yes ☐ No

3. What is the duration of the course?

_______ (hours)

4. Indicate if you have ever used or currently use any of the following 3D tools:

☐ 3ds Max ☐ Blender ☐ Cinema 4D ☐ Maya ☐ Softimage Other: _________

5. How often do you use the 3D tools that you have selected in the previous question:

☐ Occasionally ☐ A few hours per week ☐ Daily

Following is the questionnaire completed by the users to collect their data after participating in the user experiment described in Chapter 6.

This research is part of the Doctoral Program in Digital Media (PDMD) of the Faculty of Engineering of the University of Porto (FEUP). The purpose of this questionnaire is to collect data from the users who have tested conventional rig interaction modes.

Name: ____

Nationality: ______________________________ Age: __________

1. To perform the task "Inner and center brows move downwards by -1", rate each interaction mode in a scale 0 to 10. (0 - not important / 10 - very important). Mark the value with an X.

Window										
0	1	2	3	4	5	6	7	8	9	10

Camera										
0	1	2	3	4	5	6	7	8	9	10

Viewport										
0	1	2	3	4	5	6	7	8	9	10

2. To perform the task "Lower jaw thrusts fully backwards", rate each interaction mode in a scale 0 to 10. (0 - not important / 10 - very important). Mark the value with an X.

Window										
0	1	2	3	4	5	6	7	8	9	10

Camera										
0	1	2	3	4	5	6	7	8	9	10

Viewport										
0	1	2	3	4	5	6	7	8	9	10

3. To perform the task "Upper row of lips move about half-way upwards in less than 7s", rate each interaction mode in a scale 0 to 10. (0 - not important / 10 - very important). Mark the value with an X.

Window										
0	1	2	3	4	5	6	7	8	9	10

Camera										
0	1	2	3	4	5	6	7	8	9	10

Viewport										
0	1	2	3	4	5	6	7	8	9	10

4. Would you agree that the window mode is more exact, the camera mode is more intuitive and the viewport mode is more direct?

☐ Yes ☐ No

4.1. If you answered "No" in the previous question please answer why.___________________

____

____

5. If you would need to transfer the deformation values of a facial pose from a character to another, which interaction mode would you use:

☐ Window ☐ Camera ☐ Viewport

6. If you would require specific facial deformations without the need to orbit the 3D view, which mode would you use:

☐ Window ☐ Camera ☐ Viewport

7. If you would need to build a facial pose quickly, which mode would you use:

☐ Window ☐ Camera ☐ Viewport

8. From the three different interaction modes, please state which is the one that you most prefer:

☐ Window ☐ Camera ☐ Viewport

9. During the use of each interaction mode, did you feel frustrated at any time in terms of manipulating the facial model due to the fact of being limited to a single interaction mode at a time?

☐ Yes ☐ No

10. Is there any recommendation that you might have to improve the current distribution of the controls in each interaction mode?

____

____

____

D) Python Implementation Codes

Following are the Python implementation codes relative to the algorithm methods 1 to 4, which are presented in Section 5.2 of Chapter 5. The algorithms are also found and can be tested inside the Blender file entitled *section_5.2* available in the folder *blend* in the accompanying CD.

Algorithm Method 1

```
import bpy
perc = 0.875
scaley = 1.0
posebones = bpy.data.objects['skeleton'].pose.bones
master = posebones['master']
librow_WIN = posebones['librow_WIN']
librow_CAM = posebones['librow_CAM']
librow_VIP = posebones['librow_VIP']
class options(bpy.types.Panel):
    bl_label = 'Options'
    bl_space_type = 'VIEW_3D'
    bl_region_type = 'UI'
    def draw(self, context):
        layout = self.layout
        col = layout.column()
        col.prop(master, '["length"]', text='Deformation Length', slider=True)
class brows(bpy.types.Panel):
    bl_label = 'Eyebrows'
    bl_space_type = 'VIEW_3D'
    bl_region_type = 'UI'
    def draw(self, context):
        layout = self.layout
        col = layout.column()
        split = layout.split(perc, align=True)
```

```
        split.scale_y = scaley
        split.prop(librow_WIN, '["librow_NUM"]', text='Down / Up', slider=True)
        split.operator('librow_reset.op', text='', icon='PLAY_REVERSE')
class librow_reset(bpy.types.Operator):
    bl_description = 'Reset'
    bl_idname = 'librow_reset.op'
    bl_label = 'librow_reset'
    def execute(self, context):
        librow_WIN['librow_NUM'] = 0.0
        librow_CAM.location.z = 0.0
        librow_VIP.location.z = 0.0
        bpy.ops.screen.frame_jump(end=False)
        return {'FINISHED'}
def brows_display(context):
    if librow_WIN.bone.select or librow_CAM.bone.select or librow_VIP.bone.select is True:
        bpy.utils.register_class(brows)
    else:
        bpy.utils.unregister_class(brows)
bpy.app.handlers.scene_update_post.append(brows_display)
bpy.utils.register_module(__name__)
```

Algorithm Method 2

```
import bpy
posebones = bpy.data.objects['skeleton'].pose.bones
master = posebones['master']
librow_WIN = posebones['librow_WIN']
librow_CAM = posebones['librow_CAM']
librow_VIP = posebones['librow_VIP']
def handler_brows(context):
    equival = 1 / master['length']
    if librow_WIN.bone.select is True:
        librow_CAM.location.z = librow_WIN['librow_NUM']
```

```
        librow_VIP.location.z = librow_WIN['librow_NUM'] / equival
    if librow_CAM.bone.select is True:
        librow_WIN['librow_NUM'] = librow_VIP.location.z / master['length']
        librow_VIP.location.z = librow_CAM.location.z * master['length']
    if librow_VIP.bone.select is True:
        librow_WIN['librow_NUM'] = librow_VIP.location.z / master['length']
        librow_CAM.location.z = librow_VIP.location.z / master['length']
def handler_brows_length(context):
    librow_VIP.constraints['limit_location'].min_z = -master['length']
    librow_VIP.constraints['limit_location'].max_z = master['length']
bpy.app.handlers.scene_update_post.append(handler_brows)
bpy.app.handlers.scene_update_post.append(handler_brows_length)
```

Algorithm Method 3

```
import bpy
armature = bpy.data.armatures['skeleton']
posebones = bpy.data.objects['skeleton'].pose.bones
librow_WIN = posebones['librow_WIN']
librow_CAM = posebones['librow_CAM']
librow_VIP = posebones['librow_VIP']
class modals(bpy.types.Operator):
    bl_idname = 'modals.op'
    bl_label = 'modals'
    timer = None
    def modal(self, context, event):
        if event.type == 'ESC':
            return self.cancel(context)
        if event.type == 'MOUSEMOVE' and event.value == 'RELEASE':
            if librow_CAM.bone.select or librow_VIP.bone.select is True:
                bpy.ops.pose.select_all(action='TOGGLE')
                librow_WIN.bone.select = True
                armature.bones.active = librow_WIN.bone
        return {'PASS_THROUGH'}
```

```
    def execute(self, context):
        context.window_manager.modal_handler_add(self)
        self.timer = context.window_manager.event_timer_add(0.1, context.window)
        return {'RUNNING_MODAL'}
    def cancel(self, context):
        context.window_manager.event_timer_remove(self.timer)
        return {'CANCELLED'}
bpy.utils.register_module(__name__)
bpy.ops.modals.op()
```

Algorithm Method 4

```
class compile_all(bpy.types.Operator):
    bl_description = 'compile_all'
    bl_idname = 'compile_all.op'
    bl_label = 'compile_all'
    def execute(self, context):
        menus = {'edit_text': bpy.data.texts['02_menus.py']}
        modals = {'edit_text': bpy.data.texts['03_modals.py']}
        brows = {'edit_text': bpy.data.texts['04_brows.py']}
        jaw = {'edit_text': bpy.data.texts['05_jaw.py']}
        operators = {'edit_text': bpy.data.texts['06_operators.py']}
        text.run_script(menus)
        text.run_script(modals)
        text.run_script(brows)
        text.run_script(jaw)
        text.run_script(operators)
        bpy.ops.screen.frame_jump(end=False)
        return {'FINISHED'}
```

E) Accompanying CD-ROM

Accompanying the printed version of this dissertation is a CD-ROM which includes three folders entitled *blender, publications, sections blend files*, a text file entitled *read me* that describes each folder's contents and also a copy of the thesis and of the author's CV in PDF format.