MORPHABLE 3D FACIAL ANIMATION BASED ON THIN PLATE SPLINES

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

AYSU ERDOĞDU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICAL ENGINEERING

May 2010

Approval of the thesis:

MORPHABLE 3D FACIAL ANIMATION
BASED ON THIN PLATE SPLINES

submitted by **AYSU ERDOĞDU** in partial fulfillment of the requirements for the degree of **Master of Science in Electrical and Electronical Engineering Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen _____
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. İsmet Erkmen _____
Head of Department, **Electrical and Electronical Engineering**

Assist. Prof. Dr. İlkay Ulusoy _____
Supervisor, **Electrical and Electronical Dept., METU**

Examining Committee Members:

Prof. Dr. Semih Bilgen _____
Electrical and Electronical Engineering Dept., METU

Assist. Prof. Dr. İlkay Ulusoy _____
Electrical and Electronical Engineering Dept., METU

Prof. Dr. Uğur Halıcı _____
Electrical and Electronical Engineering Dept., METU

Assoc. Prof. Dr. Veysi İşler _____
Computer Engineering Dept., METU

Erdem Akagündüz _____
Engineer, ASELSAN

Date: _____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : Aysu ERDOĞDU

Signature :

ABSTRACT

MORPHABLE 3D FACIAL ANIMATION

BASED ON THIN PLATE SPLINES

Aysu Erdoğdu

M.S., Department of Electrical and Electronical Engineering

Supervisor: Assist. Prof. Dr. İlkay Ulusoy

May 2010, 106 pages

The aim of this study is to present a novel three dimensional (3D) facial animation method for morphing emotions and facial expressions from one face model to another. For this purpose, smooth and realistic face models were animated with thin plate splines (TPS). Neutral face models were animated and compared with the actual expressive face models. Neutral and expressive face models were obtained from subjects via a 3D face scanner.

The face models were preprocessed for pose and size normalization. Then muscle and wrinkle control points were located to the source face with neutral expression according to the human anatomy. Facial Action Coding System (FACS) was used to determine the control points and the face regions in the underlying model. The final positions of the control points after a facial expression were received from the expressive scan data of the source face. Afterwards control points were transferred to the target face using the facial landmarks and TPS as the morphing function. Finally, the neutral target face was animated with control points by TPS.

In order to visualize the method, face scans with expressions composed of a selected subset of action units found in Bosphorus Database were used. Five lower-face and three-upper face action units are simulated during this study. For experimental results, the facial expressions were created on the 3D neutral face scan data of a human subject and the synthetic faces were compared to the subject's actual 3D scan data with the same facial expressions taken from the dataset.

Keywords: facial animation, facial expression, emotion, thin plate spline, morphing, Facial Action Coding System.

ÖZ

İNCE LEVHA EĞRİSİ TEMELLİ
DÖNÜŞTÜRÜLEBİLEN 3B YÜZ ANİMASYONU

Aysu Erdoğdu
Yüksek Lisans, Elektrik ve Elektronik Mühendisliği Bölümü
Tez Yöneticisi: Yar. Doç. Dr. İlkay Ulusoy

Mayıs 2010, 106 sayfa

Bu çalışmanın amacı, yüzdeki duygu ve ifadelerinin bir yüzden başka bir yüze aktarılabilmesi için yeni bir üç boyutlu (3B) yüz animasyonu yöntemi sunmaktır. Bu amaçla, ince levha eğrileri ile düzgün ve gerçekçi yüz modelleri canlandırılmıştır. İfadesiz yüz modelleri canlandırılmış ve ifadeli gerçek yüz modelleri ile karşılaştırılmıştır. Tüm yüz modelleri, deneklerden 3B yüz tarayıcısı ile elde edilmiştir.

Poz ve boyut normalleştirme için yüz modellerinin önişlemesi sonrasında, kas ve kırışıklık kontrol noktaları ifadesiz kaynak yüzün üzerine insan anatomisine uygun şekilde yerleştirilmiştir. Taban modeldeki kontrol noktalarını ve etkilenen yüz bölgelerini belirlemek için Yüz Hareketi Kodlama Sistemi (Facial Action Coding System - FACS) kullanılmıştır. Kaynak yüzün ifadeli yüz verisinden, kontrol noktalarının ifade sonrası son konumları elde edilmiştir. Sonrasında kontrol noktaları, yüzdeki işaretler ve kesintisiz dönüşüm (morphing) işlevi olarak ince levha eğrileri kullanılarak hedef yüze taşınmıştır. Son olarak ifadesiz hedef yüz, kontrol noktaları ve ince levha eğrileri vasıtasıyla canlandırılmıştır.

Bahsi geçen yöntemi gözde canlandırmak için, Boğaziçi Veritabanı'nda bulunan ve farklı hareket birimlerinden (action unit) oluşan ifadelere sahip tarama verileri kullanılmıştır. Bu çalışma esnasında, beş alt-yüz ve üç üst-yüz hareket birimi canlandırılmıştır. Deneysel sonuçlar için, bir insan deneğinin 3B ifadesiz yüz tarama verisi üzerinde yüz ifadeleri yaratılmış ve yaratılan bu sentetik yüzler, aynı deneğin veri setinde bulunan, aynı yüz ifadeleri ile taranmış gerçek 3B yüz verileri ile karşılaştırılmıştır.

Anahtar kelimeler: yüz animasyonu, yüz ifadesi, duygu, ince levha eğrisi, kesintisiz dönüşüm, Yüz Hareketi Kodlama Sistemi.

ACKNOWLEDGEMENT

I would like to express my appreciation and thankfulness to my supervisor Assist. Prof. Dr. İlkay Ulusoy for her encouragement and support. I wish to thank the thesis committee members for their suggestions and comments.

I thank Erdem Akagündüz for his guidance and encouragement throughout my research. I also thank all the people who worked on Bosphorus Database for helping me to visualize my studies.

Thanks to all my colleagues in HAVELSAN, especially Kamil Arslankoz, Aydın Akkaya and Dilay Doğan for their support.

I would like to thank Melis Şerefoğlu, Buğsu Övünç, Sevcan Yeşiltaş and Ceren Tanış for being good friends; and the thanks extend to Doruk Cinoğlu and Birce Dilge Taşkın for being great mates at the library.

Special thanks go to my sister Aslı and my parents Kürşat and Havva Erdoğdu for their enduring support; and to my dear love Adem Onur Miskbay for always being there when needed.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

1D : One Dimensional

2D : Two Dimensional

3D : Three Dimensional

AFM : Average Face Model

AMA : Abstract Muscle Action

AU : Action Unit

CG : Computer Graphics

CGI : Computer Generated Imagery

DB : Database

FACS : Facial Action Coding System

FAP : Face Animation Parameter

GUI : Graphical User Interface

NURBS : Non-uniform rational basis spline

RBF : Radial Basis Function

TPS : Thin Plate Spline

CHAPTER 1

INTRODUCTION

Facial animation regards the process of presenting a speech and an emotional state of a character graphically on a computer system seemingly with real persons. Obtaining realistic facial animations is a very difficult and challenging task due to the various points listed below.

At first, in our daily life, every one of us sees many people, and knows how human faces should look. Because of that it is very easy to recognize abnormal facial expressions. Thus, the smallest nuances in the shape, texture or motion instantly alert the observer, resulting in the loss of realism of the animation. This is one of the main reasons which makes the creation of a virtual face that looks like a real person very challenging.

Another point, that makes facial animation elaborate, is the complication of the human face anatomy. Human face involves many layers of different kinds of material, which compose of skin, muscles, and bones. At the face, there are plenty of muscles, which are fully synchronized among themselves and look natural with the other facial parts. Control of all of these muscles in a synchronized way not only needs many computational power, but also very complicated.

Another important point that affects the complication of facial animation is the variety of faces. Every single person has unique facial characteristics, due to distinct muscle and bone dimensions, as well as particular ways of characteristic

expressions. Because of this reason, morphable animation methods has a special place in the facial animation area.

1.1. Application Areas

Extensive research has been done on the 3D face modeling for almost 40 years. With the progress of computer graphics technology, it has been seen that several application areas exist (Figure 1-1). A short summary of the applications for the virtual faces will help exploring the details of facial animation.

Medicine: Virtual head models are used mostly on preoperative planning of a surgery in the medical field. The correctness of the head model is critical for successful predictions of the effects of an operation. Detailed and volumetric patient data is obtained for a realistic model, but one data set from one type of volume scanners is not enough for whole head model. By combining several scans, highly accurate simulation of the behavior of facial skin, the bone structure, and other tissues may be achieved.

Criminology: Construction, identification, and modification of facial images and models are commonly used by law enforcement. A 3D face model can be created by combining photographs or 2D images created by artists on a computer. To identify a felon, the model is modified until the result looks close enough to the mental image of the victim. When a person is missing, after some time, old photographs do not help to the process of identification and modifying the attributes of the face becomes necessary. The simulation of age changes can also be performed on a 3D model. Additionally, reconstructing a 3D face by modeling the tissue layers on top of the skull makes it possible to identify a deceased person.

Movies: Number of computer generated movies has increased with the development of computer graphics and the appreciation of the audience. The faces of actors play a critical role in animations as much as in real world movies. Realism is essential for a successful face animation. In 1988 the Pixar short “Tin Toy” was the first computer animated film to win an Oscar, where the child’s face was animated using Waters’ model [1]. Today “Avatar”, which is 40% live-

action and 60% photo-realistic computer generated imagery (CGI), broke the record for the fastest film to reach the $1 billion mark at the box-office in worldwide gross, which took only 17 days [2]. Due to the Internet Movie Database (IMDb) “During the production of the film, the actors had cameras attached to their head so that they filmed close-ups of their faces. Dots painted on their faces allowed motion-capture software to record their facial expressions, providing a framework from which the computer graphics (CG) artists worked.” [2].

Figure 1-1: Different application areas of facial modeling and animation: (above left) Medicine [3]; (above right) Movies [2]; (below) Criminology [4]

Games: Extraordinarily, life-like characters are expected to appear in computer games like films. In contrast to the films, the quality of facial animation in a real-

time game is driven by the available hardware resources, such as graphics board. Thanks to the technological progress in both hardware and software, game graphics are consistently improving, and it is likely that cinematic quality will be reached in the near future.

Interactive Systems: Today the input channels of a desktop computer are limited to a keyboard and a mouse when a lot of information needs to be exchanged. Speech reproduction and recognition technologies are searched for more natural and richer means of communication. There is no doubt that a conversational interface, which supports the audio with an animation of a talking head, can enhance the dialog. Avatars, computer-controlled characters and representations of real users, are another application area of virtual faces. They characterize in inhabited virtual environments, like collaborative work environments or social networks. Natural communication between two users through their virtual face representations has to be very quick and precise to avoid latencies and misunderstandings that would give the dialog an awkward feeling.

1.2. Motivation

Facial animation methods can be classified into three main groups, which are (i) image-based [5, 6], (ii) geometry-based [1, 7-11], and (iii) performance based [12-14] animations. Geometric manipulation methods include parameterization [7, 8], interpolation (key-framing) [7], pseudo muscle-based animation [9, 15] and physics-based animation [1, 10, 11]. Image manipulation methods can be divided into morphing [5] and blendshaping [6]. Performance driven animation uses both geometry-based and image-based approaches to animate the face.

Interpolation [7] defines flowing movement of a vertex within two specific positions over a normalized time interval so that a face shape can be converted to another one. This approach requires extensive amount of computations and a large number of key frames. The approach is also inflexible, since the ranges of generated expressions are limited by the key frames and it is difficult to generalize the work on each different face mesh [7].

Direct parameterizations [7, 8] specify any possible facial motion with a synthesis of separate limit values. Although it overcomes few of the disadvantages and limitations of plain interpolations, parameterization has its own problems. The greatest problem of parameterization is the unusual collision between limits, which makes the facial appearance seem artificial [16]. Moreover parameter set belongs to a fixed face structure, so it is difficult to generalize the approach over different face meshes [7, 8].

Pseudo muscles [9, 15] deform the facial mesh by simulating real muscle contractions, ignoring the complicated underlying anatomy. This method generates greater outcomes than parameterization and interpolation, but it is usually unsuccessful at modeling wrinkles on the facial surface [9, 15].

Physics-based methods [1, 10, 11] are consisting of reproducing muscles according to the real muscles in our body. These approaches fall into three subcategories: mass spring systems [11], vector representations [1] and layered spring meshes [10, 17]. Physics-based approaches need detailed anatomical knowledge of face, which is a complex assembly of bones, muscles, blood vessels, fatty tissue, connective tissue and skin.

In image-based animation approaches [5, 6], a face model is built using the set of sample images took from the real person. The advantage of image-based approaches is simplicity. If a photograph or video footage is used as input, the output is naturally photo-realistic and no modeling of the different parts of a head is required. On the other hand, a change of posture that requires appearance of features invisible in the input, such as wrinkles, cannot be done without further modeling. Furthermore, realistic embedding into a 3D environment is not possible.

Performance driven animation [12-14] is the newest technique which can be implemented using both image and geometry-based methods. The motions are captured directly from a performer's face and then transferred to the virtual head model. By this way awkward key poses, one of the weak points of interpolation, are avoided. Precise marking of feature points is essential to provide an

unchanging and high-quality animation. Therefore performance driven techniques are mostly data-driven as a result of capturing numerous facial markers.

Transferring an existing expression from one face to another without destroying the animation performance is as important as performing successful animations on a face. There have been many methods for expression transfer, which is conceptually regarded as performance driven facial animation. For example, Noh and Neumann [14] clones expressions by transfer vertex motion vectors. 3D facial motion capture (mocap) data is widely used in these techniques [18-20], to directly animate 3D face models.

In this study, performance based animation is chosen for implementation, since nothing is more realistic than the real mimics acted by truly existing humans. Whenever these real mimics are correctly recreated, the outcomes are very impressive. Since a method was developed using the existing data of Bosphorus Database, capturing new data by tracking feature points was not needed. Therefore the main disadvantage of performance based animation approaches, being largely data-driven, is avoided. Instead of facial motion capture data, generic control points are transferred to the face models and used to animate them.

Since the facial skin is a smooth and flexible surface, splines have been used in several facial animation methods [4, 21-23], thanks to its continuity and smoothness properties. Till now, thin plate splines (TPS) have never been reported to be used to deform the facial surface and to create a realistic expression by expression mapping. Because of these advantages TPS was used to animate the face in this study. TPS brings in smoothness and flexibility to the animation system while conserving the advantages of performance based methods. Another reason of using TPS in animation is that TPS approach is superior by giving user the chance of taking bulges and wrinkles into account, while other performance based approaches [12, 13] usually fail to create wrinkles without an additional wrinkle creation approach.

1.3. Thesis Outline

In Chapter 1, the thesis is introduced in general with the application areas of the facial animation and the motivation of this work.

Previous studies on facial animation are summarized in Chapter 2. Special attention is given on the geometry-based, image-based, and performance based animation methods. Spline applications in facial animation and expression coding systems are also analyzed.

The anatomical structure of the human face and how the muscles create the facial expressions are examined in Chapter 3.

Mathematical methods used for this study, namely, radial basis functions and thin plate splines, are explained in Chapter 4.

In Chapter 5, whole implemented animation system is explained step by step. After preprocessing phases are presented, the process of marking muscle and wrinkle control points is explained. The chapter continues with the description of transferring control points using the facial landmarks and TPS as the morphing function. As the concluding step, animation with TPS is discussed with the sample results.

In Chapter 6, experimental results are examined. Finally, conclusions drawn from this study and the future work is stated in Chapter 7.

CHAPTER 2

RELATED WORK

2.1. Introduction

Plenty of people work on facial animation through many different methods since Parke's work (Figure 2-1) on 1972 [7] which found 250 polygons and 400 vertices sufficient to achieve a realistic face and had computer graphics problems like view angle and shading. As the years went by, facial animation techniques have been changed due to improvements in related areas, especially on computer technology.

Previous works on facial animation will be examined in this chapter. Especially geometry-based, image-based and performance based animation methods will be highlighted. In addition, spline applications in facial animation and expression coding systems will be analyzed.

Figure 2-1: A phase of the face (left) and a data photograph (right) from the first work on facial modeling and animation [7]

2.2. Face Recognition, Face Modeling and Facial Animation

Three remarkable areas of facial analysis, which are face recognition, face modeling and facial animation, attach together with strong bonds. Not only they benefit from each other's outputs, but they are also applied for the same application areas, such as criminology and medicine.

Facial recognition algorithms identify faces by obtaining landmarks from an image of the subject's face. The algorithm may examine the relative position, size or shape of the different face parts, such as eyes and nose. Later these features are used to search for other images with matching features.

Face modeling algorithms generate realistic human face models from digital images or from videos. Constructed models may have additional features like being textured.

Facial animation algorithms generate and animate images of the human face in 2D or 3D. Considering its subject and output type, it is so related to face recognition and modeling areas.

Same algorithms may be used for each area, like normalizing and then compressing a face data. Also they can work with the same source, such as a face database. Although these areas have an interconnected structure, face recognition and face modeling is out of scope for this thesis. But it must be kept in mind that the inputs and outputs of this work can always be a part of a system, which is used for face recognition or modeling.

2.3. Facial Animation Techniques

Due to Noh and Neumann [24], there are two major approaches to facial animation: image and geometry manipulation (Figure 2-2). In image-based animation, a face model is built using the set of sample images took from the real person [25]. Geometric modeling is based on deformations of the 3D shape of a human face. It is hard to classify the facial animation techniques further this classification, since there are no strict lines between the approaches. Frequently a

recent approach uses techniques presented earlier, or different approaches are adapted to separate facial divisions. Nevertheless geometric manipulation methods include parameterization, key-framing (interpolation), physics-based modeling, and pseudo muscle-based modeling; that image manipulation methods can be divided into morphing and blendshaping and that performance driven animation uses both image-based and geometry-based approaches to animate the face.

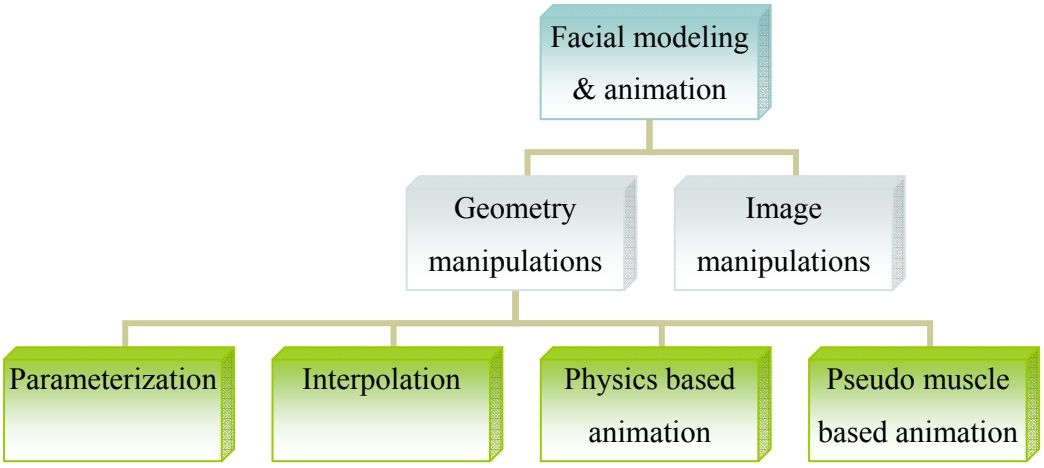


Figure 2-2: Classification of facial animation methods [24]

Manual interpolation and direct parameterization techniques belong to past times, when simplest graphics methods, like smooth shading, need to be performed without automatic tools. The pseudo muscle-based methods acts like a link between previous methods and physics-based techniques.

2.3.1. Geometry-based Animation

2.3.1.1. Interpolation

Interpolation defines flowing movement of a vertex within two specific positions over a normalized time interval [26] (Figure 2-3). If this vertex is a piece of a polygonal face, an illusion of face deformation is created. Leading work of Parke [7] presented an example of interpolation representing the skin by a little number of polygons. Nowadays higher number of vertices is needed for realism,

therefore the amount of computations is extensive and the data set required for interpolation is large. In addition the approach is inflexible, since the ranges of generated expressions are limited by the key frames and it is difficult to generalize the work on each different face mesh.

Figure 2-3: Linear Interpolation is performed on blendshapes: (left) Neutral pose, (middle) interpolated shape and (right) "A" mouth shape [27]

According to Deng and Noh [27] "Interpolation is the most intuitive and commonly used technique in facial animation". They state that "It continues to be used in projects such as Stuart Little, Star Wars, and Lord of the Rings and was adopted in many commercial animation software packages such as Maya and 3D Studio Max."

2.3.1.2. Direct Parameterization

Direct parameterizations specify any possible facial motion with a synthesis of separate limit values [7, 8]. The animation fundamentals of parameterization depend on interpolation with a smaller set of parameters. By this approach, recording the transformations as movements of the control points minimizes the data storage requirements.

Although it overcomes few of the disadvantages and limitations of plain interpolations, parameterization has its own problems. The greatest problem of

parameterization is the disagreement of parameters, which causes the expression to seem to be artificial. Moreover the set of parameters is bound to a certain facial topology. In order to animate a different face, the set of parameters needs to be rewritten, so it is difficult to generalize the approach over different face meshes.

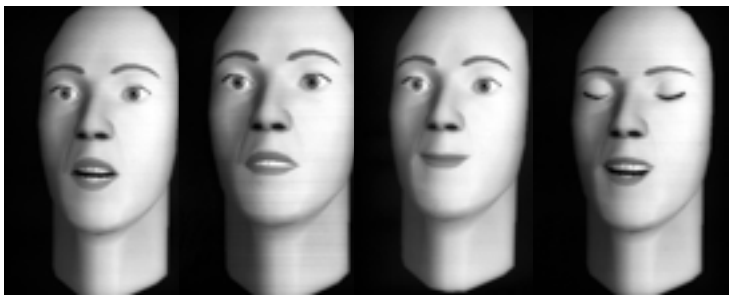


Figure 2-4: Parameterized model of Parke [7, 28]

Parke [7] (Figure 2-4) used expression parameters to control the animation, including eyelid opening, eyebrow position, eye direction, jaw rotation, upper lip position and mouth articulation. Cohen and Massaro [8] used parameters controlling tongue length, angle, width and thickness for coarticulation in realistic speech synthesis.

2.3.1.3. Pseudo Muscle-based Animation

In pseudo muscle-based animation, muscle forces are simulated using geometric deformations in the form of splines, wires or free form deformations [29]. Pseudo muscles deform the facial mesh by simulating real muscle contractions, ignoring the complicated underlying anatomy. Deformation usually occurs locally only on the surface. This technique produces better results than both interpolation and direct parameterization, but it is usually unsuccessful at modeling wrinkles on the facial surface.

The baby “Billy” in “Tin Toy” (Figure 2-5), movie of Pixar, was created by a team at Pixar lead by William Reeves [15]. He was first constructed with triangular Bezier patches but it suffered from wrinkling problems. Then instead

of Bezier patches Bicubic Catmull-Rom spline patches were used. Wrinkling problems were reduced but not completely eliminated. In spite of these problems, the film won Academy Award for Animated Short Film in 1988.

Another example is the Facial Action Control Editor (Figure 2-5) created by Waite [9]. Waite proposed to use bicubic B-spline patches with Action Units to model and animate a virtual face. The surface created by B-spline patches was smooth as a result of the C^2 continuity property of the B-spline patches. Openings for the eyes, nostrils and mouth were created with a technique called geometric trimming. Since the model included only 16 x 12 control points, it was not a very detailed model without eyelids and eyes. Some of the expressive face masks of Waite are seen in Figure 2-5.

Nahas, Huitric and Saintourens [30] presented a method using B-splines to model head data obtained with a 3D scanner. They moved the control points to effect the distortion of the face. Since adding details to one area of a continuous spline surface causes a dramatic increase in the number of patches needed in the surface, at the model of Nahas et al. the areas with high detail lack of enough amount of curvature and the model has no nostrils.

Wang et al. [22, 23] presented a system called Langwidere (Figure 2-5) that integrated simulated muscles with hierarchical spline models, which reduced the number of unnecessary control points. Wrinkles and various facial expressions could be created by muscles accompanying hierarchical spline surfaces. They showed that smoothness and flexibility could be achieved using bicubic B-splines, which is a hard job for the traditional polygonal models. But when a deformation is required to be finer than the patch resolution, the difficulty of using B-splines for complex surfaces becomes clear. Although Langwidere is very realistic, it is difficult to animate as the control points on the patches are not on the surface.

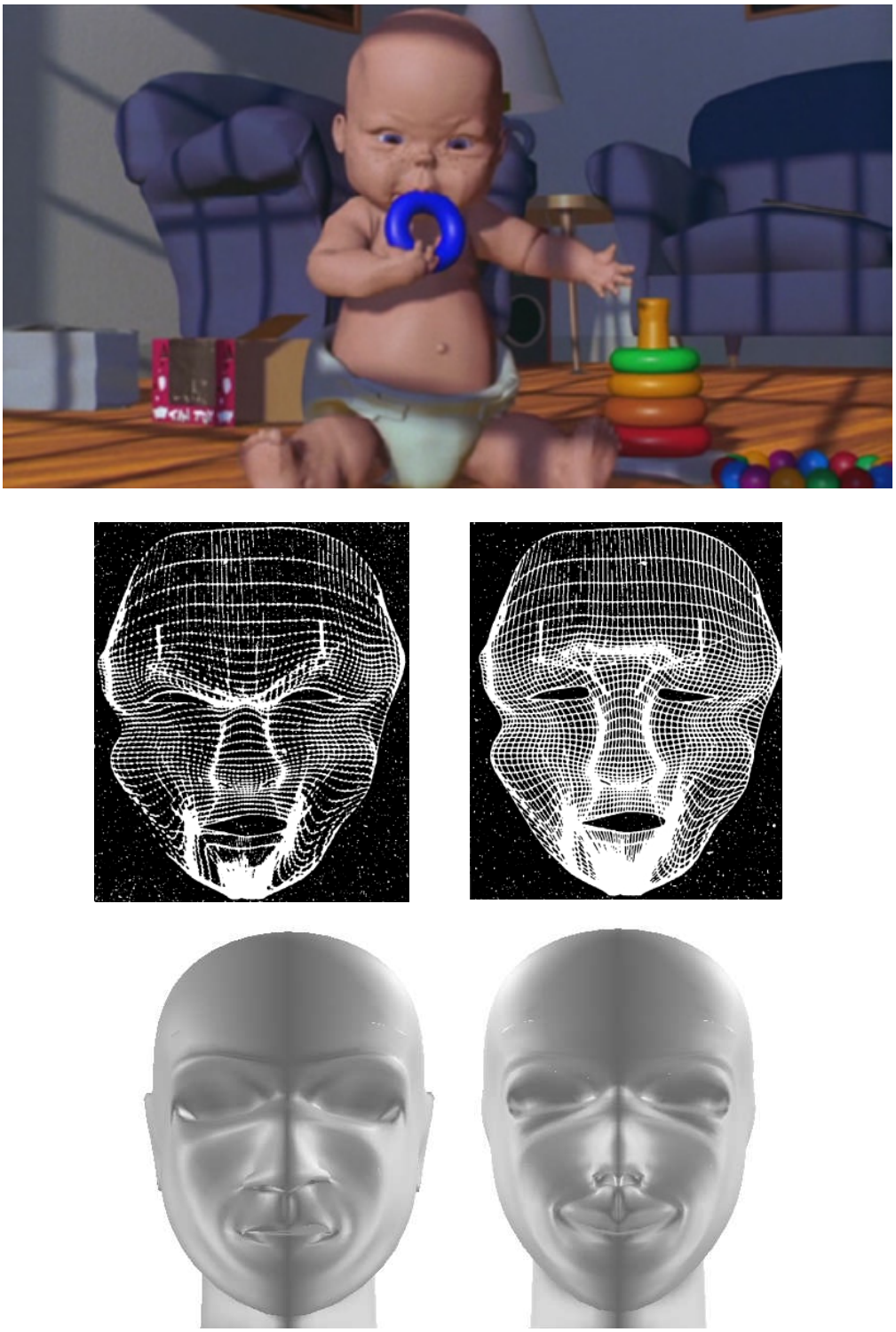


Figure 2-5: (above) Billy from “Tin Toy”; (middle) Expressive face masks of Waite [9]; (below) Langwidere of Wang et al. [31, 32]

2.3.1.4. Physics-based Animation

Physics-based methods are characterized by simulating muscles and muscle actions consistent with the actual muscles in the human body. According to Noh and Neumann [24], these approaches fall into three subcategories: mass spring systems, vector representations and layered spring meshes [11]. They defined these subcategories as the following: "Mass-spring methods propagate muscle forces in an elastic spring mesh that models skin deformation. The vector approach deforms a facial mesh using motion fields in delineated regions of influence [1]. A layered spring mesh extends a mass spring structure into three connected mesh layers to model anatomical facial behavior more detailed [10]."

The system proposed by Platt and Badler [11] is the first mass-spring approach in facial animation, which models the skin as a network of springs in three-layer structure with no thickness. Their facial model consists of three levels: bone, skin and muscles consisting of several fibers that are represented by elastic arcs.

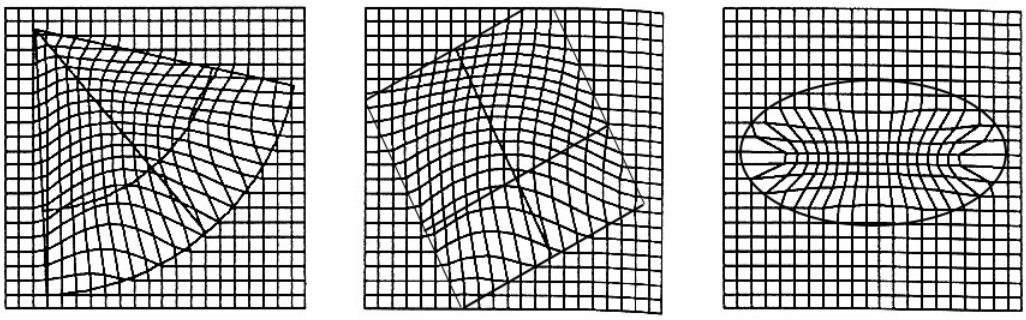


Figure 2-6: Waters' muscle model contractions: (left) linear muscle; (middle) sheet muscle and (right) sphincter muscle

A muscle model with vector representation was proposed by Waters [1]. This study defines three different muscle types by the nature of their actions, namely linear, sheet and sphincter as seen in Figure 2-6. Due to this model a muscle definition includes the vector field direction, an origin, and an insertion point. Waters animates human emotions such as anger, fear, surprise, disgust, joy, and happiness using vector based muscles implementing the Facial Action Coding

System (FACS). Most physics-based models today are still built using Waters' basic principles.

Terzopoulos and Waters [10] presented a realistic face model which combines an anatomically based facial muscle models with a physically based human tissue model. Their three-layers of deformable mesh correspond to skin, fatty tissue and muscle tied to bone. Elastic spring elements connect each mesh node and each layer.

Lee et al. [17] developed a biomechanical facial skin model based on earlier work [10]. The face model consists of three components: a biological tissue layer with nonlinear deformation properties, a muscle layer knit together under the skin, and an impenetrable skull structure beneath the muscle layer.

In the architecture of the agent called Obie (Figure 2-7), described by Bui [33], action units are used to generate emotional facial expressions on a polygonal face mesh and a B-spline surface of the lips. Bui extend Waters' muscle model [1] to generate bulges and wrinkles and to improve the combination of multiple muscle actions. He proposed simulating parallelism by calculating the resultant displacement internally, then applying it to the vertex.

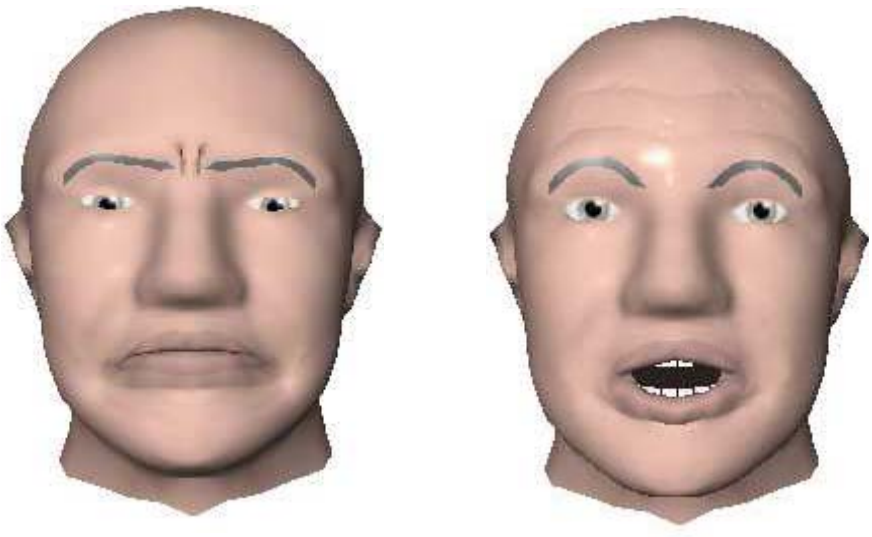


Figure 2-7: Expressive faces created with facial animation system of Bui [33]

2.3.2. Image-based Animation

In image-based animation approaches, a face model is built using the set of sample images took from the real person. According to Noh and Neumann [24], "Image-based methods include image morphing between photographic images, texture manipulations, image blending and vascular expressions."

An early work of morphing between photographic images belongs to Beier and Neely [5]. They demonstrated 2D morphing between two images with manually specified corresponding features. Reality, with this approach, requires extensive manual interaction for color balancing, correspondence selection, and tuning of the warp and dissolve parameters. Also details in facial expressions are not included in such a warped 2D image.

To overcome the limitations of 2D morphs, Pighin et al. [6] combined 2D morphing techniques with 3D transformations between different facial expressions of a geometric model to automatically create photorealistic textured 3D facial models. They animated key facial expressions with 3D geometric interpolation while image morphing was performed between corresponding texture maps.

The advantage of image-based approaches is simplicity. If a photograph or video footage is used as input, the output is naturally photo-realistic and no modeling of the different parts of a head is required. On the other hand, a change of posture that requires appearance of features invisible in the input, such as wrinkles, cannot be done without further modeling. Additionally, realistic embedding into a 3D environment is not possible.

2.3.3. Performance Based Animation

Performance driven animation is the newest technique which can be implemented using both image and geometry-based methods. A real performer observes the animations and creates with his/her motions and expressions. Real time video processing allows interactive animations, capturing the motions directly from the performer's face and then transferring to the virtual head model.

By this way awkward key poses, one of the weak points of interpolation, are avoided.

Motion data can be used to directly generate facial animation or to infer AUs of the FACS in generating facial expressions. Precise marking of feature points is essential to provide an unchanging and high-quality animation. Therefore these methods are largely data-driven as a result of capturing numerous facial markers.

Williams [12] presented the first work that acquires the expressions of real faces and applies them to computer-generated faces. He synthesized expressions by changing the 2D texture coordinates using the differences between static images.

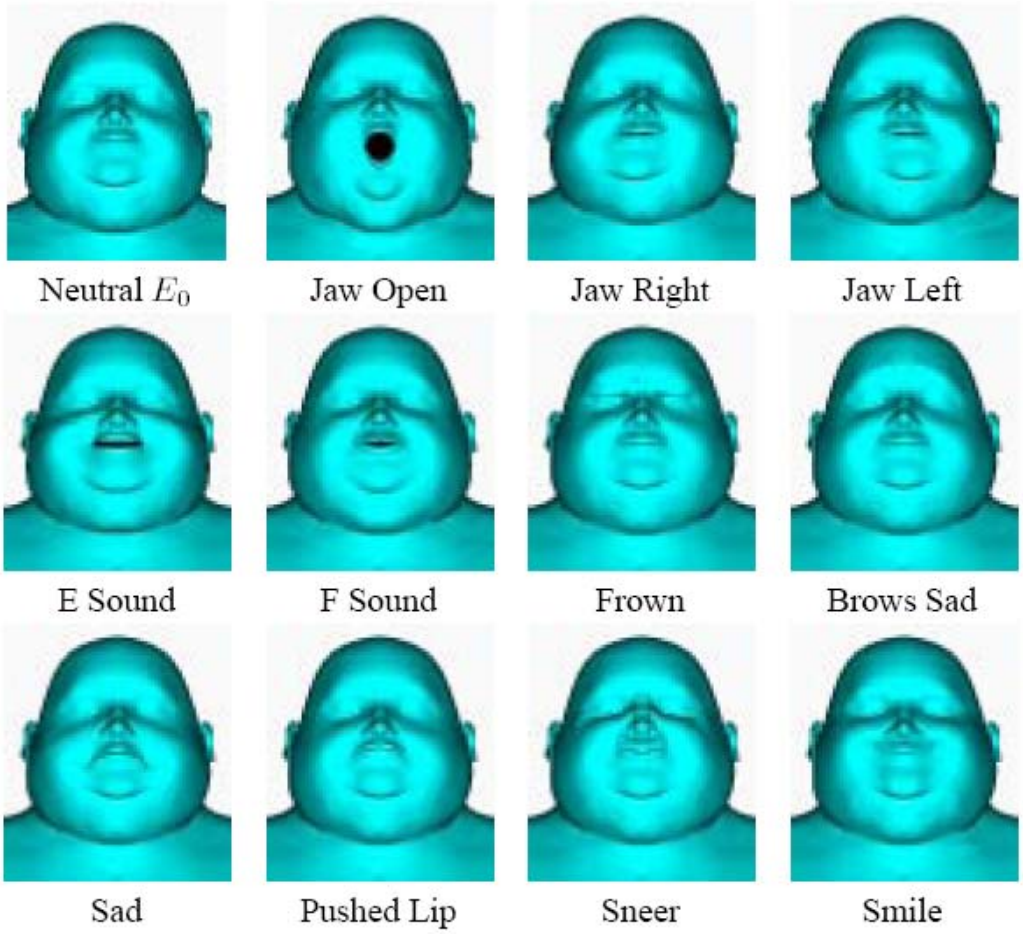


Figure 2-8: A series of facial expressions and the neutral expression used to produce real time animation by Kouadio et al. [34]

Guenter et al. [13] captured both the 3D geometry and color information for human facial expressions from a video stream, using 182 markers. Afterwards they reconstructed photorealistic, 3D animations of the captured expressions. Kouadio et al. [34] use a series of 3D facial expressions of a synthetic model and blending between them to produce real time animation (Figure 2-8).

Transferring an existing facial motion to different 3D face models can be conceptually regarded as performance driven facial animation. According to Deng and Neumann [27], "Automatically transferring facial motions from an existing model to a new model can significantly save painstaking and model-specific animation tuning for the new face model. The source facial motions can have various formats, including 2D video faces, 3D facial motion capture data, and animated face meshes, while the target models typically are a static 3D face mesh or a blendshape face model."

Noh and Neumann [14] proposed an "expression cloning" technique to transfer vertex motion vectors from a 3D source face model to a 3D target model having different geometric proportions and mesh structure (vertex number and connectivity). Thus cloned expression animations preserved the relative motions, dynamics, and character of the original facial animations. Their basic idea was to construct vertex motion mappings between models through the radial basis functions (RBF) morphing.

Sumner and Popović [35] proposed a general "deformation transfer" framework that automatically transfers geometric deformations between two triangle meshes through a correspondence map between the triangles of the source and those of the target. Deformation transfer can be directly applied to transfer facial motions from one source face mesh to a target face mesh, as seen in Figure 2-9.

Choe and Ko [18] transferred facial expressions to target blendshape face models by editing captured performances. They used the actuation of expression muscles to control facial expressions by analyzing weights through an optimization procedure.

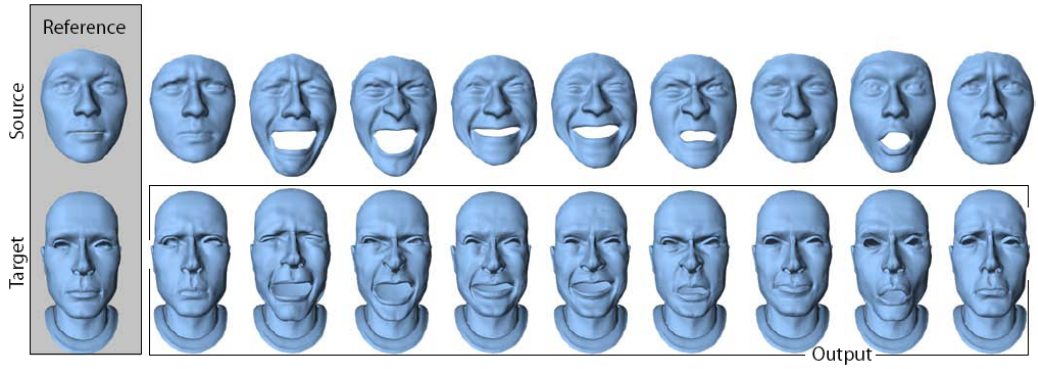


Figure 2-9: Scanned facial expressions cloned onto a digital character by Sumner and Popović [35]

Sifakis et al. [19] first created an anatomically accurate face model composed of facial musculature, passive tissue and underlying skeleton structure, using volumetric data acquired from a living male subject. Then they automatically determined accurate muscle activations from motion capture marker data.

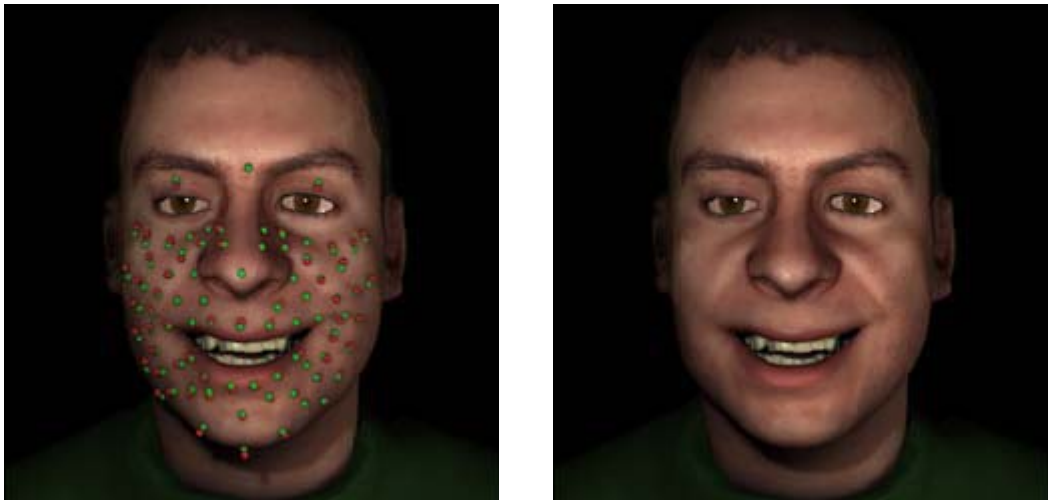


Figure 2-10: Expressions estimated from motion capture data, along with both the captured and the simulated markers by Sifakis et al. [19]

Deng et al. [20] proposed a semi-automatic technique to directly animate 3D blendshape face models by mapping 3D facial motion capture data to pre-designed 3D blendshape face spaces. In their approach, Radial Basis Function

(RBF) networks were trained to map a new motion capture frame to its corresponding blendshape weights, based on chosen training pairs between mocap frames and blendshape weights.

2.4. Spline Applications in Facial Animation

Spline methods have been used in facial animation for different purposes like texture image generation, face reconstruction and spline pseudo muscle modeling.

Spline pseudo muscle-based modeling is mentioned above in Section 2.3.1.3. As described before, pseudo muscles deform the facial mesh by simulating real muscle contractions, ignoring the complicated underlying anatomy. William Reeves [15] first constructed the baby "Billy" in "Tin Toy" movie of Pixar with triangular Bezier patches. Then instead of Bezier patches Bicubic Catmull-Rom spline patches were used. Both models suffered from wrinkling problems. Waite [9] proposed to use bicubic B-spline patches with Action Units to model and animate a virtual face. Nahas, Huitric and Saintourens [30] presented a method using B-splines to model head data obtained with a 3D scanner. They moved the control points to effect the distortion of the face. Wang et al. [22, 23] presented a system called Langwidere that integrated simulated muscles with Bicubic B-splines models, which reduced the number of unnecessary control points.

If the input photographs have been taken under uncontrolled illumination, the skin color might differ noticeably between the images. Texture images from different views are combined by using a multiresolution spline method [36, 37]. By this method visual boundaries are removed and a texture image mosaic is reached.

A face reconstruction system which fits a generic hierarchical B-spline head model to a skull mesh was proposed by Archer [21]. The prototype generated a generic hierarchical B-spline surface around a 3D scan of the skull. Sparse data points representing tissue thickness were first placed at landmarks about the scan. The generic surface was subsequently automatically placed in order to smoothly interpolate the data points.

2.5. Expression Coding Systems

Expression coding systems provide a database of muscle actions for all possible facial expressions. With this database the animator can compile a script of desired expressions through time and let the system animate them. The most popular expression coding systems are Facial Action Coding System [38, 39] and MPEG-4 models [40].

Ostermann [40] standardized more than 70 model-independent animation parameters defining low-level actions like move left mouth corner, meanwhile high-level parameters like facial expressions. He claimed that "MPEG-4 specifies a set of face animation parameters (FAPs), each corresponding to a particular facial action deforming a face model in its neutral state. The FAP value for a particular FAP indicates the magnitude of the corresponding action, e.g., a big versus a small smile. A particular facial action sequence is generated by deforming the face model in its neutral state according to the specified FAP values for the corresponding time instant. Then the model is rendered onto the screen." Feature points used to define FAPs are seen in Figure 2-11.

Facial Animation Coding System (FACS) will be discussed thoroughly in section 3.5.

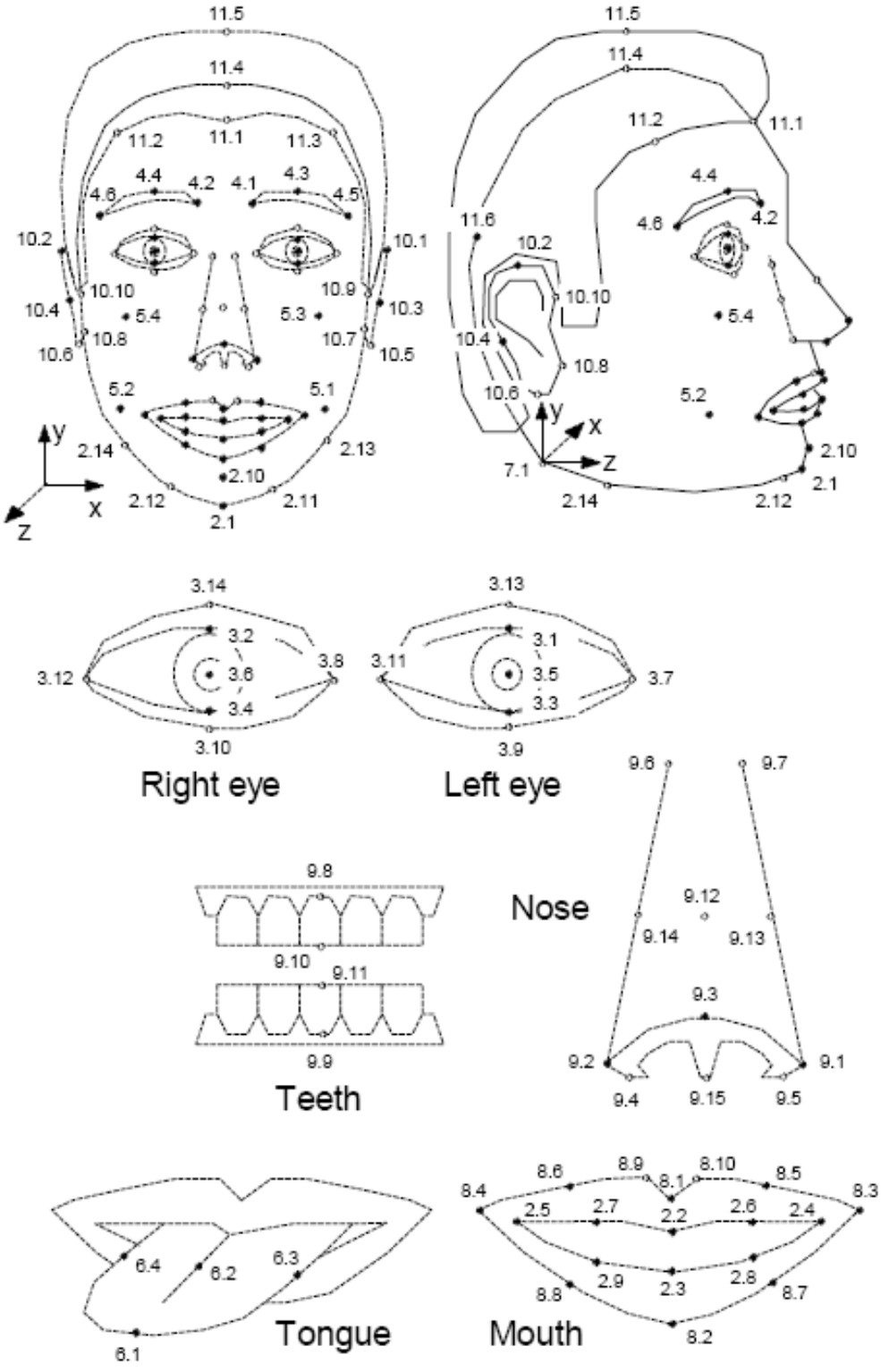


Figure 2-11: Feature points used to define FAPs [40]

CHAPTER 3

ANATOMY OF THE HUMAN FACE

3.1. Introduction

Without understanding the anatomical structure of the human face and how muscles create the expressions, we cannot model the face. Lacking this information about the underlying structure, it would be inevitable to create 3D faces that look ambiguous and incorrect.

Although this chapter examines anatomy of the human face, it does not pretend to be a full study of it. Related references [20-23] should serve as more appropriate guides for the reader desiring an in-depth study of anatomy.

3.2. The Skull

Since the layers of muscle and fat are stretched relatively thinly over the skull, its bone structure is more observable than that of any other part of the body. Figure 3-1 illustrates side and front views of the skull marked with some of the more observable bones.

The human face has one main skeletal movable component, the mandible (jawbone). The rest of the skull bones are rigidly held together by immovable joints. The skull can be divided into two sections: the cranium, which encloses the brain, consists of eight bones, and the skeleton of the face, consists of fourteen bones.

1. frontal bone
2. supraorbital ridge
3. occipital bone
4. parietal bone
5. temporal bone
6. sphenoid bone
7. mastoid process
8. ear hole
9. zygoma
10. zygomatic bone
11. maxilla
12. nasal bone
13. lacrimal bone
14. ethmoid bone
15. mandible
16. ramus
17. temporomandibular joint

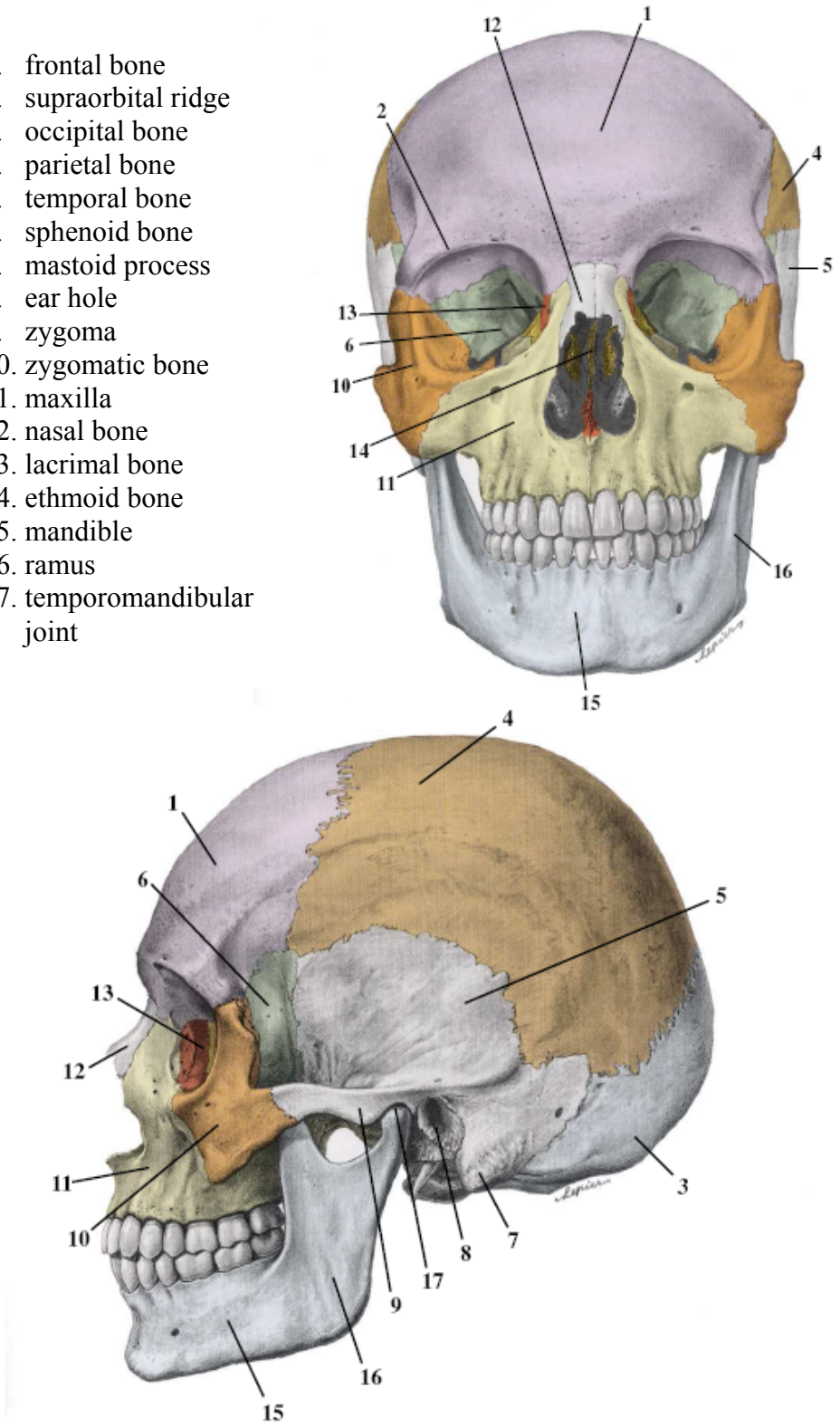


Figure 3-1: Side and front views of the skull [41]

Cranium is responsible for protecting the brain and it is formed by occipital bone, two parietal bones, frontal bone, two temporal bones, sphenoidal bone and ethmoidal bone. The frontal bone is located at the anterior of the cranium. It forms the brows as well as the protective curve over the eyes.

Facial skeleton, which rests in the frontal region of the head and serves as a base for the face, is formed by two nasal bones; two maxilla bones, or mustache bones; two lacrimal bones; two zygomatic bones, or cheekbones; two palatines; two inferior nasal conchas; vomer and mandible.

3.3. Facial Skin

Skin plays a vital role in human communication by means of facial expression, texture, color and scent. To create realistic human face models, skin should be studied carefully. Functions of skin include protection against external environmental influences, the regulation of body temperature and storage of water, fat and blood.

As shown in Figure 3-2, the skin is structured into three layers: epidermis, dermis and hypodermis. Together, the topmost layers epidermis and dermis are called the cutis, the underlying hypodermis is also known as subcutis.

Epidermis is composed of protein keratin, and gives protection against acids and alkali. It is capable of binding water, giving this layer smoothness and elasticity. Dermis is the second layer consisting of loose connective tissue with capillaries, elastic fibers, reticular fibers and some collagen. It provides strong resistance to tearing forces. Hypodermis includes adipose tissue which is mostly formed by fat cells. The fat increases the tension of the skin and isolates the body against thermal influences. The hypodermis is connected to the fascia, a layer of tissue that connects the muscle surface to the skin [42, 43].

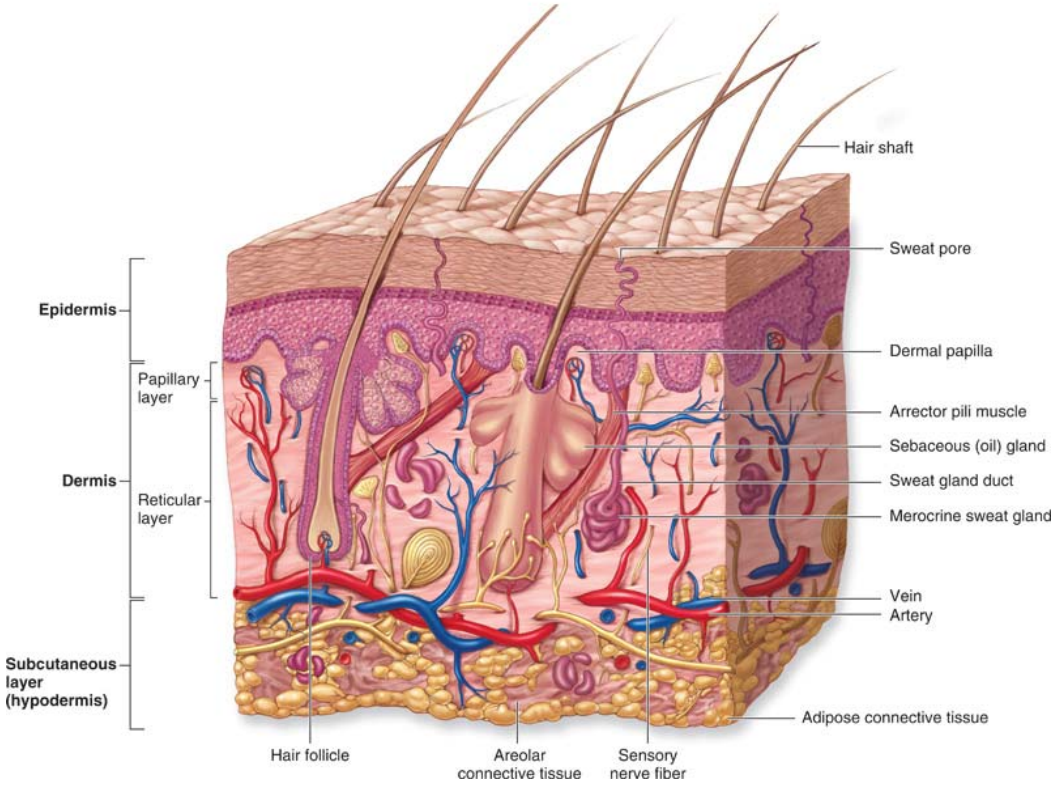


Figure 3-2: Cross section of human skin [44]

3.4. Facial Muscles

Muscles are the soft tissue, which are responsible for the motion of the human body. Motion is supplied by the contractions and the relaxations of the muscles. Facial muscles belong to skeletal muscles (the other muscle types are smooth and heart muscle). The components of a single muscle are shown in Figure 3-3.A muscle cell has cylindrical shape and each cell contains numerous cylindrical structures, called myofibrils. The myofibrils consist of a repeating pattern of sarcomeres. The combined effort of many sarcomeres leads to muscle contraction and causes visible deformation of the skin.

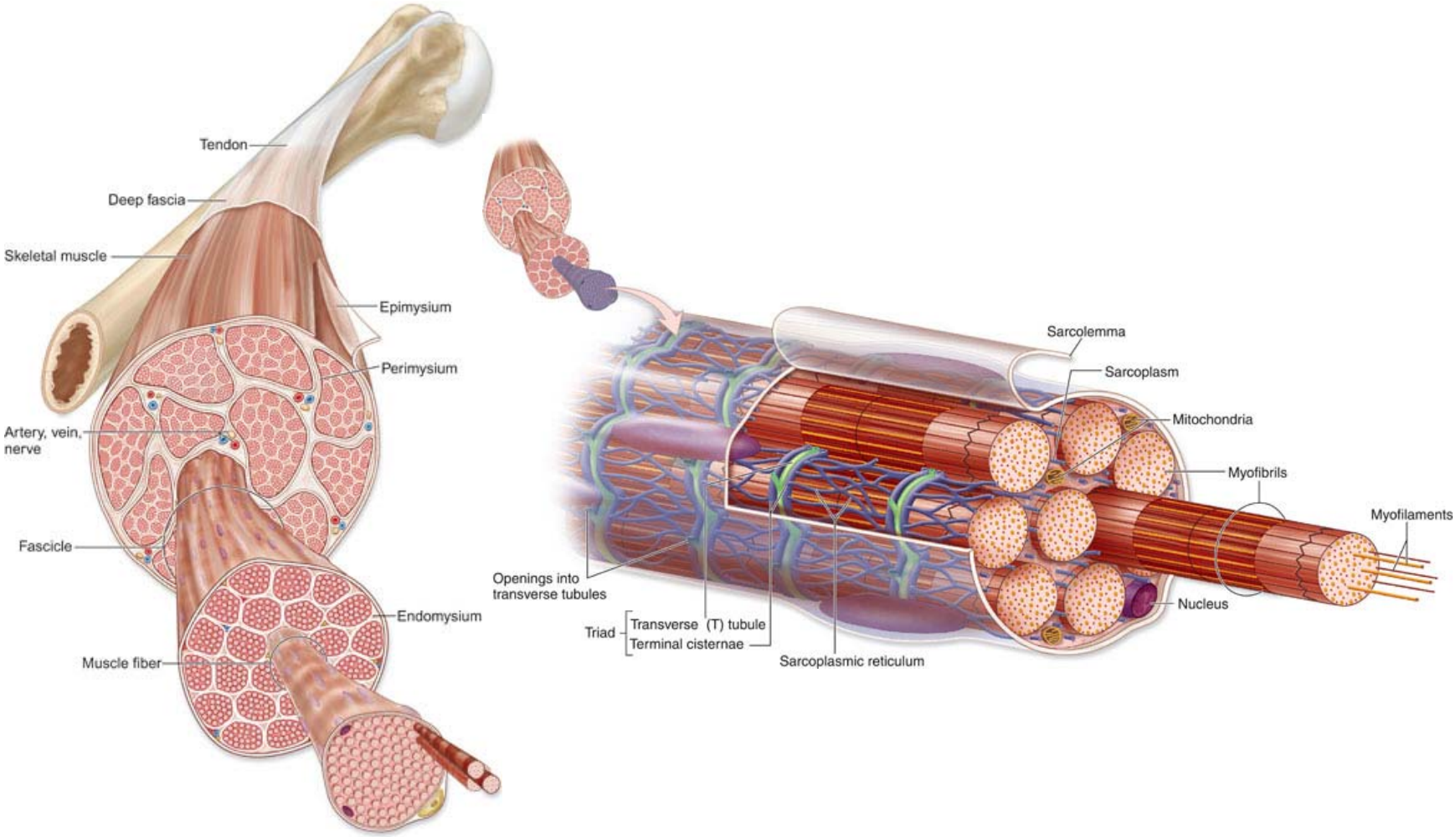


Figure 3-3: Structure of a muscle fiber [44]

Most muscles in the face originate at the bone, while their surface inserts into skin, fascia (connective tissue), cartilage (gristly elastic tissue), or fibers from other muscles. When contracting, a muscle pulls the skin towards its fixed origin. In contrast, a few facial muscles have the distinction of not being attached to bone. Either they originate in or insert into ligaments or skin or they connect to other muscles [43].

Those interested in facial animation should take some time to study these muscles and the manner in which they affect facial expressions. In the anatomy literature, the muscles of expression are categorized by their region of influence: the scalp, the eyelid, the nose, the mouth and the neck. Table 3-1 identifies the most important muscles responsible for facial expressions.

Figure 3-4 and Figure 3-5 show frontal and side views of the facial musculature. Most muscles in the face are mirrored on either side. Where the muscle descriptions at Table 3-1 use the singular, equivalent behavior for the corresponding muscles on both sides is implied.

The great importance of facial muscles in nonverbal communication becomes especially clear when they are paralyzed, as in some stroke victims. All muscles listed in Table 3-1 are innervated by cranial nerve VII, the facial nerve.

Table 3-1: Muscles of the Head for Facial Expressions [42, 43, 45, 46]

Muscle Name	Origin (O) and Insertion (I)	Action
Muscles of the Scalp		
epicranius (occipitofrontalis) (*epi* = over; *cran* = skull)	Bipartite muscle consisting of the frontalis and occipitalis connected by a cranial aponeurosis, the galea aponeurotica; the alternate of these two muscles pull scalp forward and backward.	
▪ frontalis (*front* = forehead)	O – galea aponeurotica I – skin of eyebrows and root of nose	With aponeurosis fixed, raises the eyebrows (as in surprise); wrinkles forehead skin horizontally
▪ occipitalis (*occipito* = base of skull)	O – occipital and temporal bones I – galea aponeurotica	Fixes aponeurosis and pulls scalp posteriorly
temporoparietalis (*temporo* = temple; *parietal* = walls of a cavity)	O – fascia temporalis (fascia above the ear) I – galea aponeurotica	Tightens the scalp in the temporal area. With epicranius wrinkles the forehead, widens the eyes and raises the ears
auricularis (*auricular* = ear)	O – fascia of the scalp I – areas near the earlobe	Wiggle the ears
Muscles of the Eyelid		
orbicularis oculi (*orb* = circular; *ocul* = eye)	O – frontal and maxillary bones and ligaments around orbit I – tissue of eyelid	Protects eyes from intense light and injury; various parts can be activated individually; produces blinking, squinting, and draws eyebrows inferiorly
corrugator supercilii (*corrugo* = wrinkle; *supercilium* = eyebrow)	O – arch of frontal bone above nasal bone I – skin of eyebrow	Draws eyebrows together and inferiorly; wrinkles skin of forehead vertically (as in frowning)
depressor supercilii (*depressor* = depresses)	O – root of the nose I – inner corner of the eyebrow	Lowers the inner corner of the eyebrow; forms horizontal wrinkles at the root of the nose
levator palpebrae superioris (*leva* = raise; *palpebra* = eyelid; *superior* = above, over)	O – roof of the orbit I – surface of the upper eyelid	Raises the upper eyelid; opens the eyes wide
Muscles of the Nose		
levator labii superior aleque nasi (*labi* = lip; *ala* = wing; *nasus* = nose)	O – upper frontal process of the maxilla I – skin of the lateral part of the nostril and upper lip	Dilates the nostril (wing of the nose) and elevates the upper lip

Table 3-1 (continued): Muscles of the Head for Facial Expressions

Muscle Name	Origin (O) and Insertion (I)	Action
nasalis	Consists of two parts, transverse and alar: compressor naris and dilator naris.	
▪ compressor naris	O – upper jaw near the canine tooth I – nasal cartilage on the bridge of the nose	Pulls the skin towards the back of the nose, elevating slightly the later part of its wings
▪ dilator naris	O – maxilla above the lateral incisor I – wing of the nose	Compresses and expands the wings of the nose; descends and depresses the later part of the base of the nose narrowing the wings of the nose
depressor septi (*septum* = dividing wall of the nose)	O – incisive fossa of the maxilla I – nasal septum and back part of the dilator naris	Depresses septum and draws the wing of the nose downward
procerus (*proser* – extend)	O – skin between and above the eyebrows I – lateral cartilages of the nose and internal edge of its bones	Draws the skin between the eyebrows downward and causes lateral wrinkles sinking slightly the head of the eyebrow
Muscles of the Mouth		
zygomaticus – major and minor (*zygomatic* = cheekbone)	O – zygomatic bone I – skin and muscle at corner of mouth	Raises lateral corners of mouth upward (smiling muscle)
risorius (*risor* = laughter)	O – lateral fascia associated with masseter muscle I – skin at the angle of mouth	Draws corner of lip laterally; tenses lip; synergist of zygomaticus
depressor labii inferioris (*infer* = below)	O – body of mandible lateral to its midline I – skin and muscle of upper lip	Draws lower lip inferiorly (as in a pout)
levator labii superioris	O – zygomatic bone and infraorbital margin of maxilla I – skin and muscle of upper lip	Opens lips; raises and furrow the upper lip
depressor anguli oris (triangularis) (*angul* = angle, corner; *or* = mouth)	O – body of mandible below incisors I – skin and muscle at angle of mouth below insertion of zygomaticus	Zygomaticus antagonist; draws corners of mouth downward and laterally (as in a “tragedy mask” grimace)

Table 3-1 (continued): Muscles of the Head for Facial Expressions

Muscle Name	Origin (O) and Insertion (I)	Action
levator anguli oris (caninus)	O – canine fossa of maxilla I – orbicularis oris and skin at angle of mouth	Raises the upper lip leaving at sight the canine tooth and takes backwards slightly the corners of the mouth
orbicularis oris	O – arises indirectly from maxilla and mandible; fibers blended with fibers of other facial muscles associated with the lips I – encircles mouth; inserts into muscle and skin at angles of mouth	Closes lips; purses and protrudes lips; kissing and whistling muscle
mentalis (transversus menti) (*ment* = chin)	O – mandible below incisors I – skin of chin	Protudes lower lip; wrinkles chin
buccinator (*bucc* = cheek or "trumpeter")	O – molar region of maxilla and mandible I – orbicularis oris	Draws corners of mouth laterally; compresses cheek (as in whistling and sucking); holds food between teeth during chewing
Muscle of the Neck		
platysma (*platy* = broad, flat)	O – fascia of chest (over pectoral muscles and deltoid) I – lower margin of mandible, and skin and muscle at the corner of mouth	Helps depress mandible; pulls lower lip back and down, i.e., produces downward sag of mouth; tenses skin of neck (e.g. during shaving)

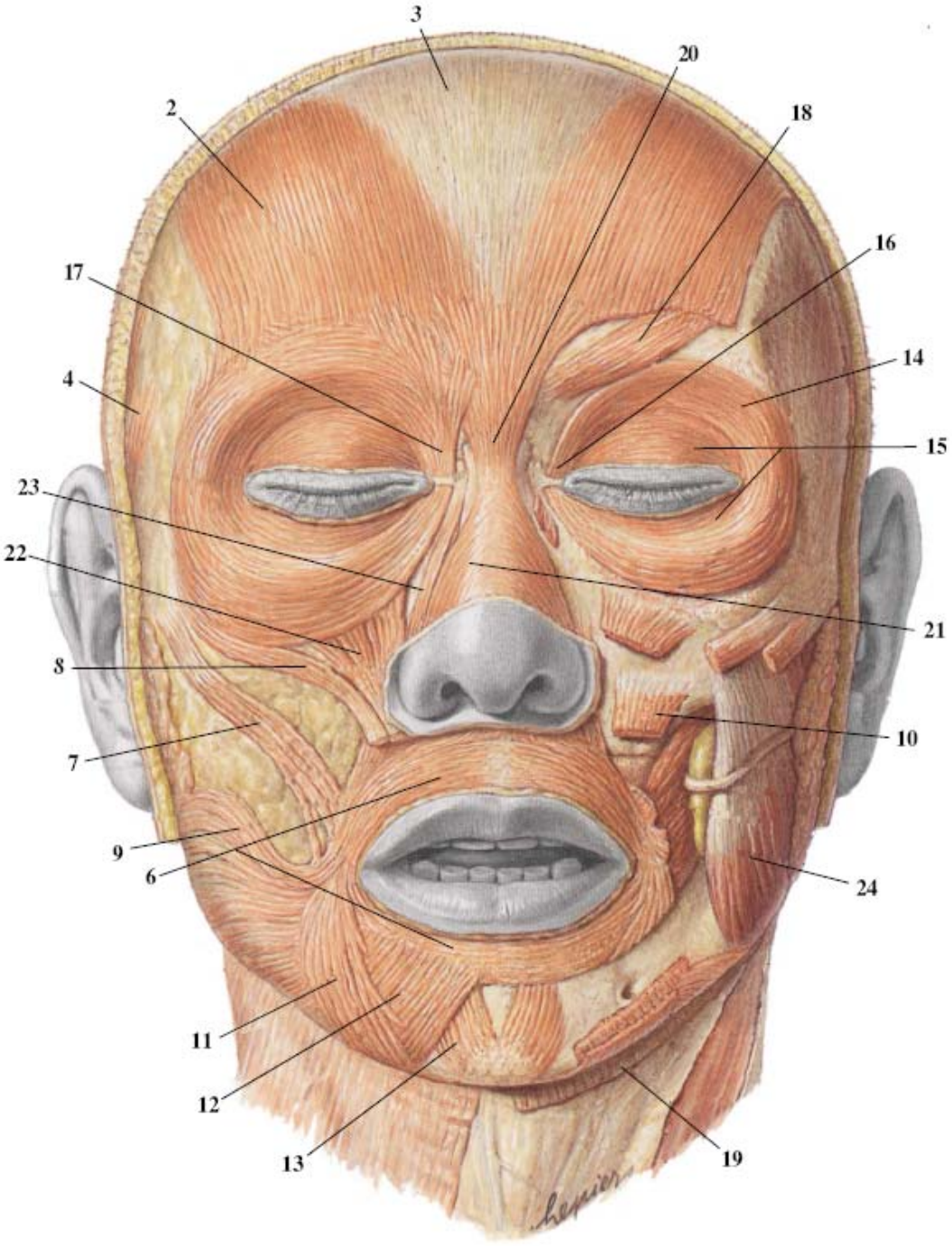
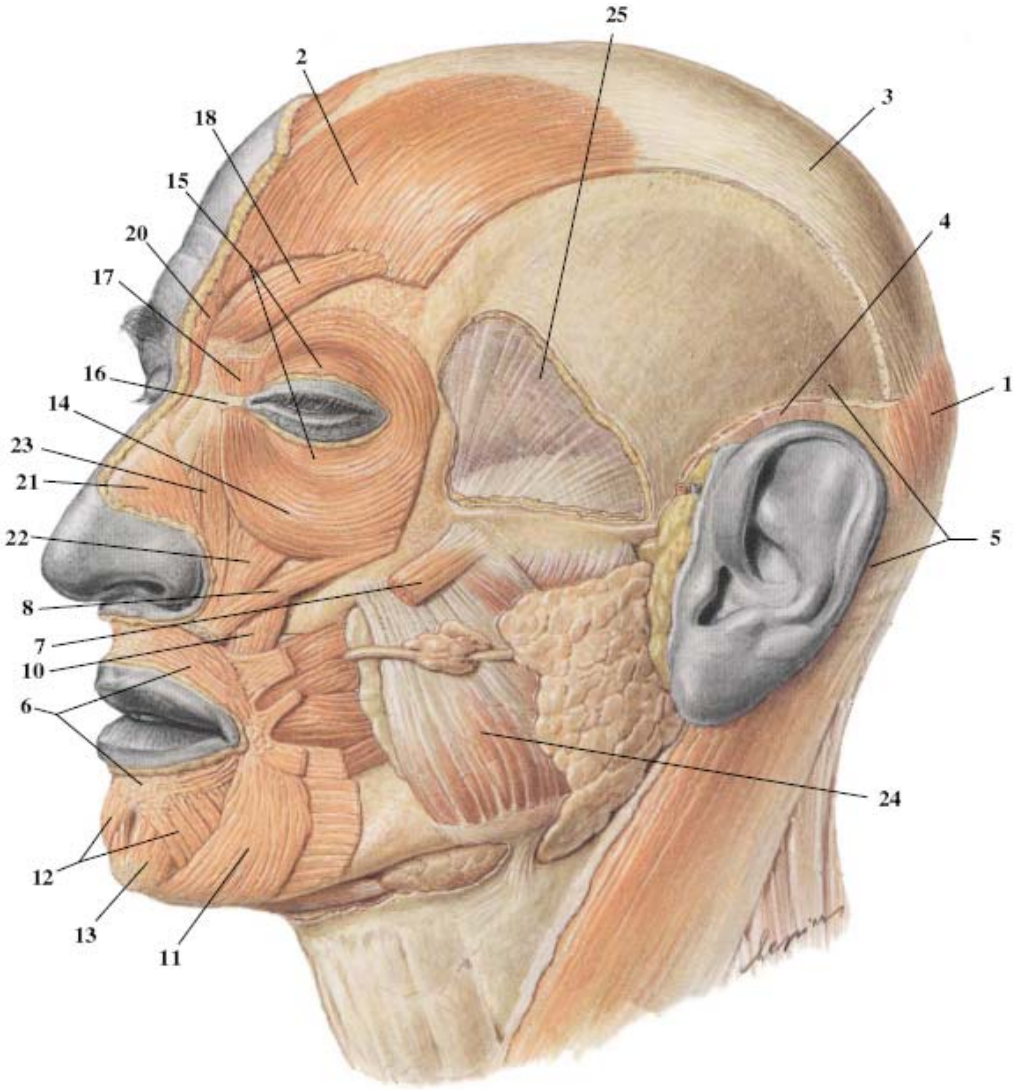


Figure 3-4: Front view of facial muscles [41]



scalp
1. occipitalis
2. frontalis
3. cranial aponeurosis
4. temporoparietalis
5. auricularis

mouth
6. orbicularis oris
7. zygomaticus major
8. zygomaticus minor
9. risorius
10. levator anguli oris
11. triangularis
12. depressor labii inferioris
13. mentalis

eyelid
14. orbicularis oculi, p. orbitalis
15. orbicularis oculi, p. palpebralis
16. levator palpebrae superioris
17. depressor supercilii
18. corrugator supercilii

neck
19. platysma

nose
20. procerus
21. nasalis
22. levator labii superioris
23. levator labii superior aleque nasi

muscles of mastication
24. masseter
25. temporalis

Figure 3-5: Side view of facial muscles [41]

3.5. Expressions and Action Units

Facial Action Coding System (FACS), the important work of the psychologists Ekman and Friesden [38], recreates facial expressions into a set of basic facial movements. These movements of individual muscles, or small groups of muscles, are described as Action Units (AU). Sample AUs are presented in Table 3-2. An example of diagrammed AUs is seen in Figure 3-6.

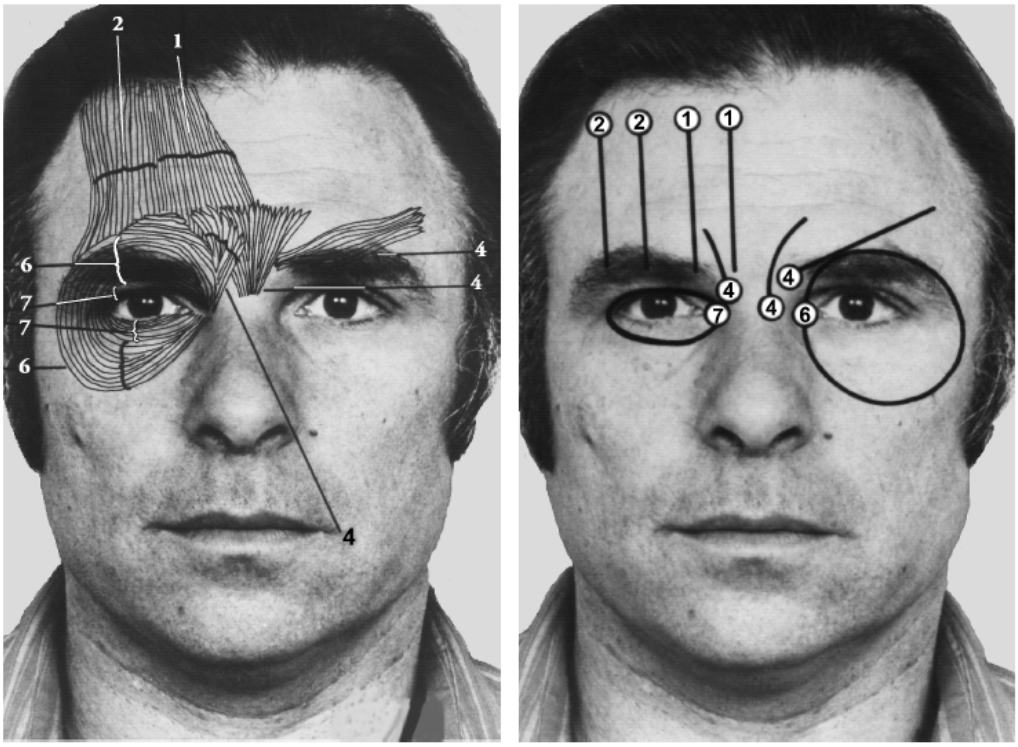


Figure 3-6: Five Action Units of the upper face [47]

Six basic emotions are described by Ekman: Anger, Fear, Surprise, Disgust, Happiness and Sadness. Each of these uses multiple combinations of Action Units. For instance, combining the AU6 (cheek raiser) and AU12 (lip corner puller) creates a happy expression. Examples of basic expressions formed by the AUs are given in Table 3-3.

A facial model can be built by parameterizing according to the motions of the AUs. Then a facial animation system can be created with a set of variables that

are in one-to-one match with the AUs. Also several abnormal expressions can be generated by joining the AUs in different ways.

Table 3-2: Single Action Units (AU) [33, 39]

AU Number	FACS Name	Muscular Basis
1	Inner Brow Raiser	frontalis, pars medialis
2	Outer Brow Raiser	frontalis, pars lateralis
4	Brow Lowerer	corrugator supercilii; depressor supercilii
5	Upper Lid Raiser	levator palpebrae superioris
6	Cheek Raiser	orbicularis oculi, pars orbitalis
7	Lid Tightener	orbicularis oculi, pars palebralis
8	Lips Toward Each Other	orbicularis oris
9	Nose Wrinkler	levator labii superior aleque nasi
10	Upper Lip Raiser	levator labii superioris
11	Nasolabial Furrow Deepener	zygomatic minor
12	Lip Corner Puller	zygomatic major
13	Cheek Puffer	caninus
14	Dimpler	buccinator
15	Lip Corner Depressor	triangularis
16	Lower Lip Depressor	depressor labii inferioris
17	Chin Raiser	mentalis
18	Lip Puckerer	incisivii labii superioris; incisivii labii inferioris
20	Lip Stretcher	risorius
22	Lip Funneler	orbicularis oris
23	Lip Tightener	orbicularis oris
24	Lip Pressor	orbicularis oris
25	Lips Part	depressor labii inferioris, or relaxation of mentalis or orbicularis oris
26	Jaw Drop	masseter; temporal and internal pterygoid relaxed
27	Mouth Stretch	pterygoids; digastric
28	Lip Suck	orbicularis oris
38	Nostril Dilator	nasalis, pars alaris
39	Nostril Compressor	nasalis, pars transversa and depressor septi nasi
41	Lid Droop	relaxation of levator palpebrae superioris
42	Slit	orbicularis oculi
43	Eyes Closed	relaxation of levator palpebrae superioris

Table 3-2 (continue): Single Action Units (AU)

AU Number	FACS Name	Muscular Basis
44	Squint	orbicularis oculi, pars palpebralis
45	Blink	relaxation of levator palpebrae and contraction of orbicularis oculi, pars palpebralis
46	Wink	orbicularis oculi

Table 3-3: Example sets of Action Units for emotion predictions [39]

Emotion	Prototypes
Surprise	AU1 + AU2 + AU5 + AU26
Fear	AU1 + AU2 + AU4 + AU5 + AU20 + AU27
Happiness	AU6 + AU12
Sadness	AU1 + AU4 + AU15
Disgust	AU9 + AU17
Anger	AU4 + AU5 + AU7 + AU23 + AU25

CHAPTER 4

RADIAL BASIS FUNCTIONS AND THIN PLATE SPLINES

4.1. Introduction

Thin Plate Splines, or TPS in short form, are used in this study for morphing from one face to another face, and also warping from the neutral face to the face with an expression. Marked control points are transferred from the source to the target face using the facial landmarks and TPS as morphing function. Then neutral target face is animated with control points by TPS as warping function.

Before going deeply to the whole facial animation method, the general class of TPS, namely radial basis functions, and their application areas in animation will be explained in this chapter. TPS and their mathematical form in 3D will also be described in detail.

4.2. Radial Basis Functions

Radial Basis Function (RBF), first proposed by Duchon in 1977 [48], is a solution to the scattered data interpolation problem, where N point samples are wanted to interpolate or extrapolate.

At the scattered data interpolation problem, a function $S(x)$ is reconstructed by given N samples (x_i, f_i), such that $S(x_i) = f_i$, where x_i are the centers. Reconstructed function is denoted as $s(x)$. There are infinite solutions to the problem, but also we have specific constraints:

- $s(x)$ should be continuous over the entire domain,

- A “smooth” surface is wanted.

RBF is a function of the form

$$s(x) = p(x) + \sum_{i=1}^{N} \lambda_i \phi(\|x - x_i\|) \tag{4.1}$$

where:

- $s(x)$ is the *RBF*,
- $p(x)$ is a *low degree polynomial*, typically linear or quadratic,
- $\phi(r)$ is the *basic function* ($r = \|x - x_i\|$),
- $\|x\|$ is the *Euclidean norm*,
- the λ_i’s are the RBF *coefficients*,
- the x_i’s are the RBF *centers*.

The RBF consists of a weighted sum of a radially symmetric basic function $\phi(r)$ located at the centers x_i and a low degree polynomial $p(x)$. Given a set of N points x_i and values f_i, the process of finding an interpolating RBF is called fitting, such that:

$$s(x_i) = f_i, \qquad i = 1, 2, \ldots, N \tag{4.2}$$

The fitted RBF is defined by the λ_i, the coefficients of the basic function in the summation, together with the coefficients of the polynomial term $p(x)$.

If we let $\{p_1, \ldots, p_l\}$ be a monomial basis for polynomials of the degree of $p(x)$ and $c = \{c_1, \ldots, c_l\}$ be the coefficients that give $p(x)$ in terms of this basis, then the interpolation conditions (4.2), that $s(x_i) = f_i$, can be rewritten in matrix form as a linear system,

(4.3) $$\begin{pmatrix} A & P \\ P^T & 0 \end{pmatrix}\begin{pmatrix} \lambda \\ c \end{pmatrix}=\begin{pmatrix} f \\ 0 \end{pmatrix},$$

where

$A_{i,j} = \phi(\|x_i - x_j\|), \qquad i, j = 1,\ldots, N,$

$P_{i,j} = p_i(x_i), \qquad i = 1,\ldots, N, \quad j = 1,\ldots, l.$

Solving the linear system (Equation (4.3)) determines λ and c, and hence $s(x)$.

Best known and commonly used types of radial basis functions are:

- Multiquadric: $\phi(r) = \sqrt{r^2 + \beta^2}$ for some $\beta > 0$,
- Gaussian: $\phi(r) = \exp(-\beta r^2)$ for some $\beta > 0$,
- Polyharmonics:
  - 2D: $\phi(r) = r^{2n} \log(r)$
  - 3D: $\phi(r) = r^{2n-1}$
- Thin Plate Spline (TPS): $\phi(r) = r^2 \ln r^2$

For more detailed knowledge on RBF, related works [49-51] should be studied.

4.3. Applications of RBF in Computer Graphics

The scattered data interpolation problem occurs in many areas, like mesh repair; surface reconstruction; range scanning for geographic surveys or medical data; field visualization (2D and 3D); image warping, morphing, registration and artificial intelligence. As a result RBF is used in several different areas of computer graphics.

In 1995, an approach was presented by Savchenko et al. [52] for the reconstruction of geometric models and surfaces from given sets of points using

basic RBF, where the algorithm can be implemented in tomography applications, image processing, animation and CAD for bodies with complex surfaces.

Turk and O'Brien proposed a new method to create smooth implicit surfaces in 1999 [53]. They created a 3D implicit function by using a variational scattered data interpolation approach. Interactive modeling (3D sculpting) and shape transformation are some applications that they have achieved with using implicit functions [54]. In Figure 4-1 a 3D shape transformation sequence is shown using the variational implicit technique.

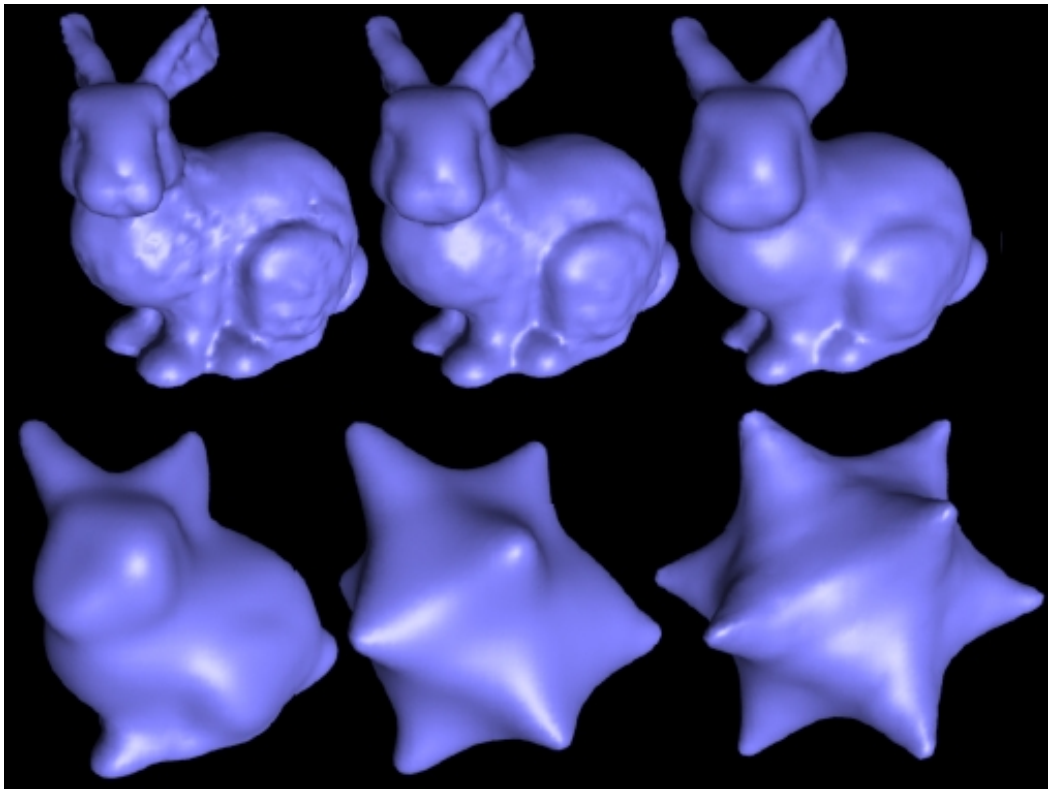


Figure 4-1: A 3D shape transformation sequence [54]

Carr and Fright employed RBF to smoothly interpolate incomplete surfaces obtained from 3D medical graphics across defect regions [55]. They matched RBFs to depth-maps of the skull's surface, derived from CT data, and repaired the defects, usually holes, in the skull with cranial implants, which is called

carnioplasty as a procedure in medicine. Figure 4-2 shows an example of fitting a surface to a cranial surface.

In 2001 Carr et al. [56] used polyharmonic RBFs for smoothing manifold surfaces from point-cloud data and for mesh repair, or in other ways, hole-filling. Fast methods for fitting and evaluating RBFs let them to model large data sets, consisting of millions of surface points, by a single RBF. They presented that the RBF representation has advantages for mesh simplification and remeshing applications. Examples of smoothing and mesh repair are given in Figure 4-3 and Figure 4-4.

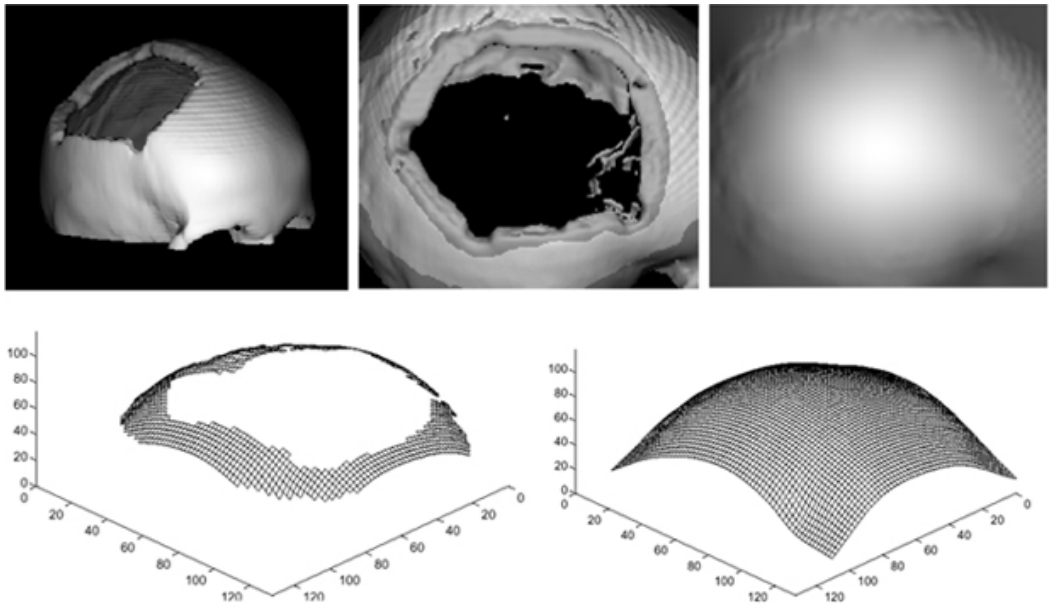


Figure 4-2: Example of fitting a surface to a cranial surface: (above) Rendered views of full CT data set, detail of defect region and surface fitted to depth-map (below) Depth-maps corresponding to defect and fitted surface [55]

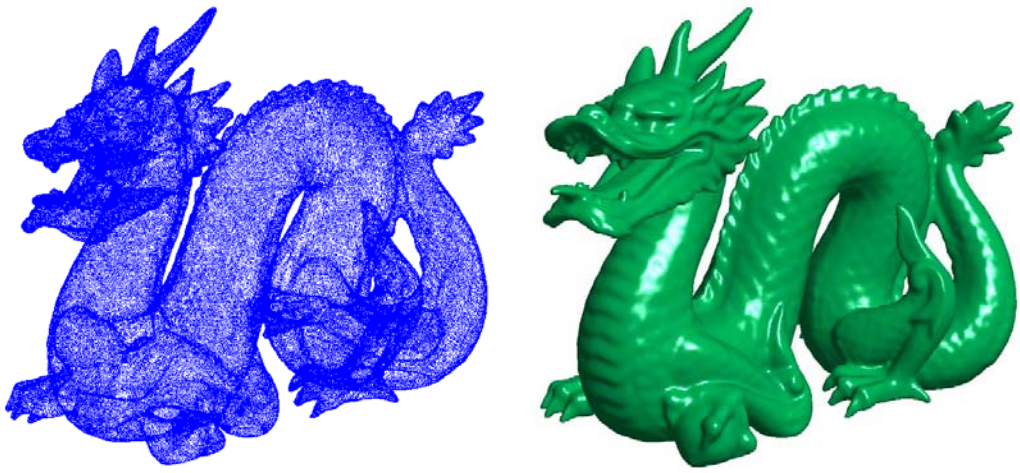


Figure 4-3: Fitting a RBF to a 438,000 point-cloud [56]

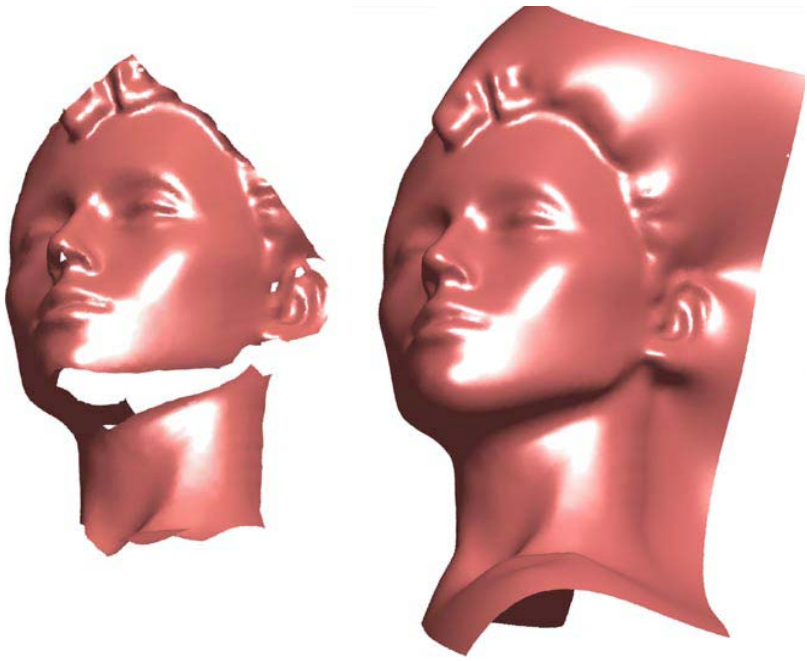


Figure 4-4: Automatic mesh repair using the biharmonic RBF [56]

4.4. Thin Plate Splines

Thin plate splines (TPS) were introduced by Duchon with the first proposition of radial basis functions [48]. Then in 1989 Bookstein analyzed TPS and used it for transformation and warping [57]. The name thin plate spline refers to a physical analogy involving the bending of a thin sheet of metal. In physical setting, the deflection is in the direction orthogonal to the plane. In coordinate transformation, one interprets the lifting of the plates as a displacement of the coordinate within the plane.

TPS is the generalization of the natural cubic splines in 1D. The purpose is to describe the deformation specified by finitely many point-correspondences in an irregular spacing.

4.4.1. Basis function $U(r)$

TPS is the fundamental solution to the biharmonic equation, and has the form

(4.4) $$U(r) = r^2 \ln r^2$$

where r is the distance $\sqrt{x^2 + y^2}$ from the Cartesian origin. The special function $z(x, y) = -U(r)$ is viewed in (Figure 4-5) from above. The X is at $(0,0,0)$; the remaining zeros of the function are on the indicated circle, where $r = 1$. The maximum of the surface is achieved along a circle of radius $1/\sqrt{e}$.

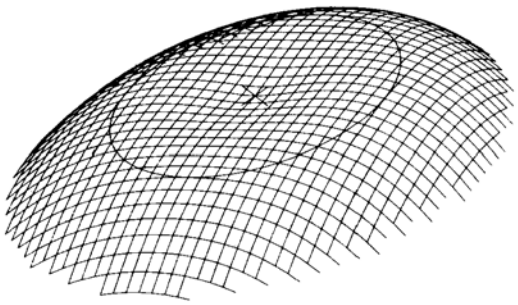


Figure 4-5: Graph of Function $U(r)$ [57]

Given a set of data points, a weighted combination of $U(r)$ centered at each data points gives the interpolation function that passes through the points exactly. Linear combination of $U(r)$ functions minimizes the "bending energy". Bending energy is defined as follows

(4.5) $$I(f) = \iint_{R^2} (f_{xx}^2 + 2 f_{xy}^2 + f_{yy}^2) dxdy$$

4.4.2. Algebra of the Thin-Plate Spline for Arbitrary Sets of Landmarks

Let $P_1 = (x_1, y_1)$ ... $P_n = (x_n, y_n)$ are n landmark points in the ordinary Euclidean plane according to any convenient Cartesian coordinate system. $r_{ij} = \left| P_i - P_j \right|$ is the distance between points i and j. Define matrices

(4.6) $$K = \begin{bmatrix} 0 & U(r_{12}) & \dots & U(r_{1n}) \\ U(r_{21}) & 0 & \dots & U(r_{2n}) \\ \dots & \dots & \dots & \dots \\ U(r_{n1}) & U(r_{n2}) & \dots & 0 \end{bmatrix}, n \times n$$

(4.7) $$P = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \dots & \dots & \dots \\ 1 & x_n & y_n \end{bmatrix}, 3 \times n$$

(4.8) $$L = \begin{bmatrix} K & P \\ P^T & O \end{bmatrix}, (n+3) \times (n+3)$$

where O is a $3x3$ matrix of zeros.

Let $V = (v_1, \dots, v_n)$ is any n-vector and $Y = (V \mid 0 \quad 0 \quad 0)^T$ is a column vector of length $n+3$. Define the vector $W = (w_1, \dots, w_n)$ and the coefficients a_1, a_x, a_y by the equation

(4.9) $$L^{-1}Y = (W \mid a_1 \quad a_x \quad a_y)^T$$

Use the elements of $L^{-1}Y$ to define the function $f(x,y)$:

$$f(x,y) = a_1 + a_x x + a_y y + \sum_{i=1}^{n} w_i U(|P_i - (x,y)|) \tag{4.10}$$

First 3 terms of Equation (4.10) describes global affine transform and the rest of the terms describes (nonglobal) non-linear transformation.

Then the following three propositions hold:

- $f(x_i, y_i) = v_i$ for all i (interpolation property)
- The function f minimizes the bending energy, since U(r) is the fundamental solution of biharmonic equation

  $$I(f) = \iint_{R^2} (f_{xx}^2 + 2f_{xy}^2 + f_{yy}^2) dxdy$$

- The value of I_f (bending energy) is proportional to

  $$WKW^T = V(L_n^{-1} K L_n^{-1}) V^T,$$

  where L_n^{-1} is the upper left $n \times n$ sub block of L^{-1}.

Take the points (x_i, y_i) to be landmarks and V to be the $n \times 2$ matrix

$$V = \begin{bmatrix} x_1' & x_2' & \dots & x_n' \\ y_1' & y_2' & \dots & y_n' \end{bmatrix}$$

where each (x_i', y_i') is the landmark homologous to (x_i, y_i) in another copy of $\mathbb{R}^2$. The resulting function $f(x,y)$ maps each point (x_i, y_i) to its homolog (x_i', y_i') and has the minimum bending energy among all such mapping functions.

4.4.3. TPS in 3D

Let $P_1 = (x_1, y_1, z_1)$ ... $P_n = (x_n, y_n, z_n)$ are n landmark points in the 3D Euclidean space according to any convenient Cartesian coordinate system. $r_{ij} = \left\| P_i - P_j \right\|$ is the Euclidean distance between points i and j in 3D. Define matrices

$$K = \begin{bmatrix} 0 & U(r_{12}) & \dots & U(r_{1n}) \\ U(r_{21}) & 0 & \dots & U(r_{2n}) \\ \dots & \dots & \dots & \dots \\ U(r_{n1}) & U(r_{n2}) & \dots & 0 \end{bmatrix}, n \times n \tag{4.11}$$

$$P = \begin{bmatrix} 1 & x_1 & y_1 & z_1 \\ 1 & x_2 & y_2 & z_2 \\ \dots & \dots & \dots & \dots \\ 1 & x_n & y_n & z_n \end{bmatrix}, 4 \times n \tag{4.12}$$

$$L = \begin{bmatrix} K & P \\ P^T & O \end{bmatrix}, (n+4) \times (n+4) \tag{4.13}$$

where O is a 4×4 matrix of zeros.

Take the points (x_i, y_i, z_i) to be landmarks and each (x_i', y_i', z_i') is the landmark homologous to (x_i, y_i, z_i) in another copy of $\Re^3$. Define matrices

$$V = \begin{bmatrix} x_1' & x_2' & \dots & x_n' \\ y_1' & y_2' & \dots & y_n' \\ z_1' & z_2' & \dots & z_n' \end{bmatrix}, n \times 3 \tag{4.14}$$

$$Y = \begin{bmatrix} V \mid O^T \end{bmatrix}, (n+4) \times 3 \tag{4.15}$$

The resulting function $f(x, y, z)$ maps each point (x_i, y_i, z_i) to its homolog (x_i', y_i', z_i') and has the minimum bending energy among all such mapping functions. Define matrices

$$L^{-1}Y = \begin{bmatrix} W \mid A \end{bmatrix} \tag{4.16}$$

(4.17) $$W = \begin{bmatrix} w_{11} & \dots & w_{1n} \\ w_{21} & \dots & w_{2n} \\ w_{31} & \dots & w_{3n} \end{bmatrix}, n \times 3$$

(4.18) $$A = \begin{bmatrix} a_{11} & a_{1x} & a_{1y} & a_{1z} \\ a_{21} & a_{2x} & a_{2y} & a_{2z} \\ a_{31} & a_{3x} & a_{3y} & a_{2z} \end{bmatrix}, 4 \times 3$$

Use the elements of W and A matrices to define the function $f(x, y, z)$:

(4.19) $$f(x, y, z) = \begin{bmatrix} 1 \\ x \\ y \\ z \end{bmatrix} A + \begin{bmatrix} U(\|P_1 - (x, y, z)\|) \\ \vdots \\ U(\|P_n - (x, y, z)\|) \end{bmatrix} W$$

In Figure 4-6, a warping example using TPS is given for a plain surface patch which is controlled by 5 points. After the control points are moved in 3D, the final positions of all surface points are interpolated by TPS.

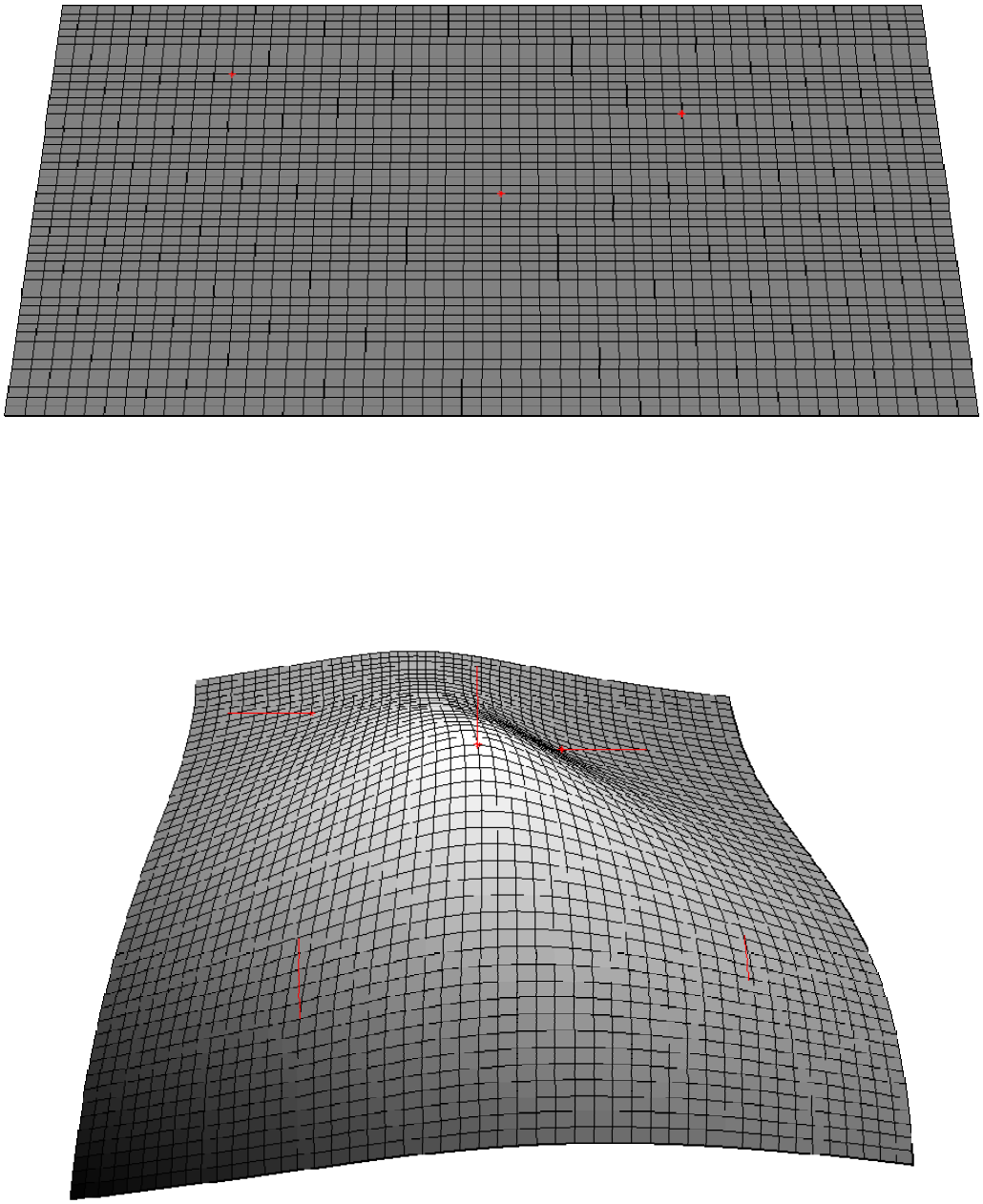


Figure 4-6: A TPS surface with 5 control points in 3D

4.5. Applications of TPS in Facial Animation

As mentioned in Section 2.4, spline methods have been used in facial animation for different purposes like texture image generation, face reconstruction and spline pseudo muscle modeling. Similar to other spline and RBF methods, TPS is also used in many researches about facial animation.

Kähler et al. used a TPS interpolation for reconstruction of faces from skull data [4]. They fitted an anatomy-based virtual head model, including skin and muscles, to a scanned skull using statistical data on skull-tissue relationships.

Guo et al. [58] introduced a semi-automatic deformation alignment method using TPS, to generate a 3D morphable face model from 3D face data. They replaced the optical flow algorithm in [59] with TPS method to perform the alignment for 3D data.

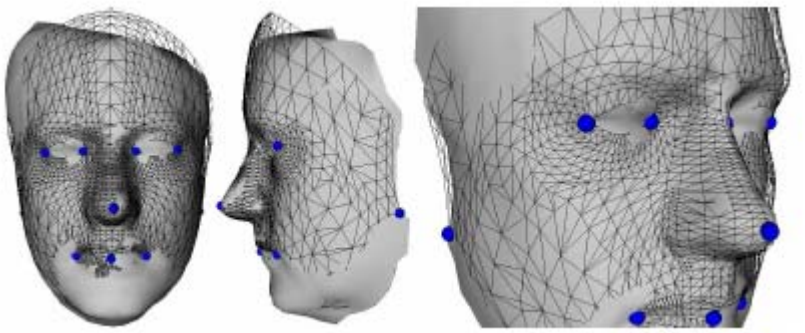


Figure 4-7: Face warping process using 10 landmarks by Orvalho et al. [22]

In 2006 Orvalho et al. [22] presented a facial deformation system that helps the character setup process and gives artists the possibility to manipulate models as if they were using a puppet. They found the correspondence of the main attributes of a generic rig, transferred them to different 3D face models and automatically generate a sophisticated facial rig based on an anatomical structure, using TPS as the warping function (Figure 4-7).

Another recent work on facial animation with TPS belongs to Aina [23]. An incremental, landmark-based process was presented by Aina for fitting a generic skull to any given face model, using TPS, as seen in Figure 4-8.

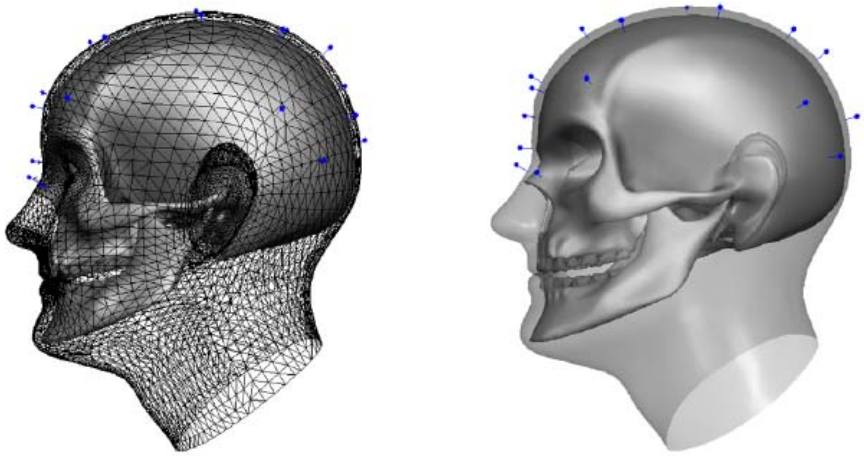


Figure 4-8: Fitting the generic skull to the European head model by TPS [23]

CHAPTER 5

FACIAL ANIMATION

5.1. Introduction

After examining previous works on facial animation, understanding the anatomical structure of the human face and how the muscles create the expressions, and recognizing radial basis functions and TPS, our aim is to explore the facial animation method in depth.

In this thesis, performance based animation was chosen for implementation, since nothing is more realistic than the real mimics acted by truly existing human beings. Whenever these real mimics are recreated in a right way, the outcomes are very impressive. Since we developed our method using the existing data of Bosphorus Database, capturing new data by tracking feature points was not needed. Therefore the main disadvantage of the performance based animation approaches, being largely data-driven, was avoided.

Till now, thin plate splines (TPS) have never been reported to be used to deform the facial surface and to create realistic mimics by expression mapping. In this study, while conserving the advantages of performance based methods, TPS, which brings in smoothness and flexibility to the animation system, is used in animating the face. Another reason of using TPS in animation is that TPS approach is superior with respect to other approaches since it gives the user the chance of taking bulges and wrinkles into account, while other performance based approaches [12, 13] usually fail to create wrinkles without an additional wrinkle creation approach.

5.2. System Overview

In this thesis, smooth and realistic face models were animated by morphing emotions and facial expressions from one face to another with TPS. Neutral face models were animated and compared with the actual expressive face models where all face models were obtained from Bosphorus Database. This database includes 105 subjects in various poses, expressions and occlusion conditions.

Whole system can be summarized in three phases: (i) preprocessing, (ii) modeling, and (iii) animation, the flowchart of which is given in Figure 5-1:

(i) Preprocessing Phase:

- Face models were preprocessed for pose and size normalization, using the average face model of the database as the reference data. The average face model was constructed for automatic 3D face recognition by Salah et al. [60].

(ii) Modeling Phase:

- Muscle and wrinkle control points were located to the source face with neutral expression according to the human anatomy. During this process, FACS was used to determine the control points and the affected face region.
- Final positions of the control points after a facial expression were received from the expressive scan data of the source face by transferring the control points from the neutral face to the expressive face, using the facial landmarks and TPS as the morphing function.

(iii) Animation Phase:

- Control points were transferred to the target face, again using the facial landmarks and TPS as morphing function.
- Neutral target face was animated with control points by TPS as the warping function.

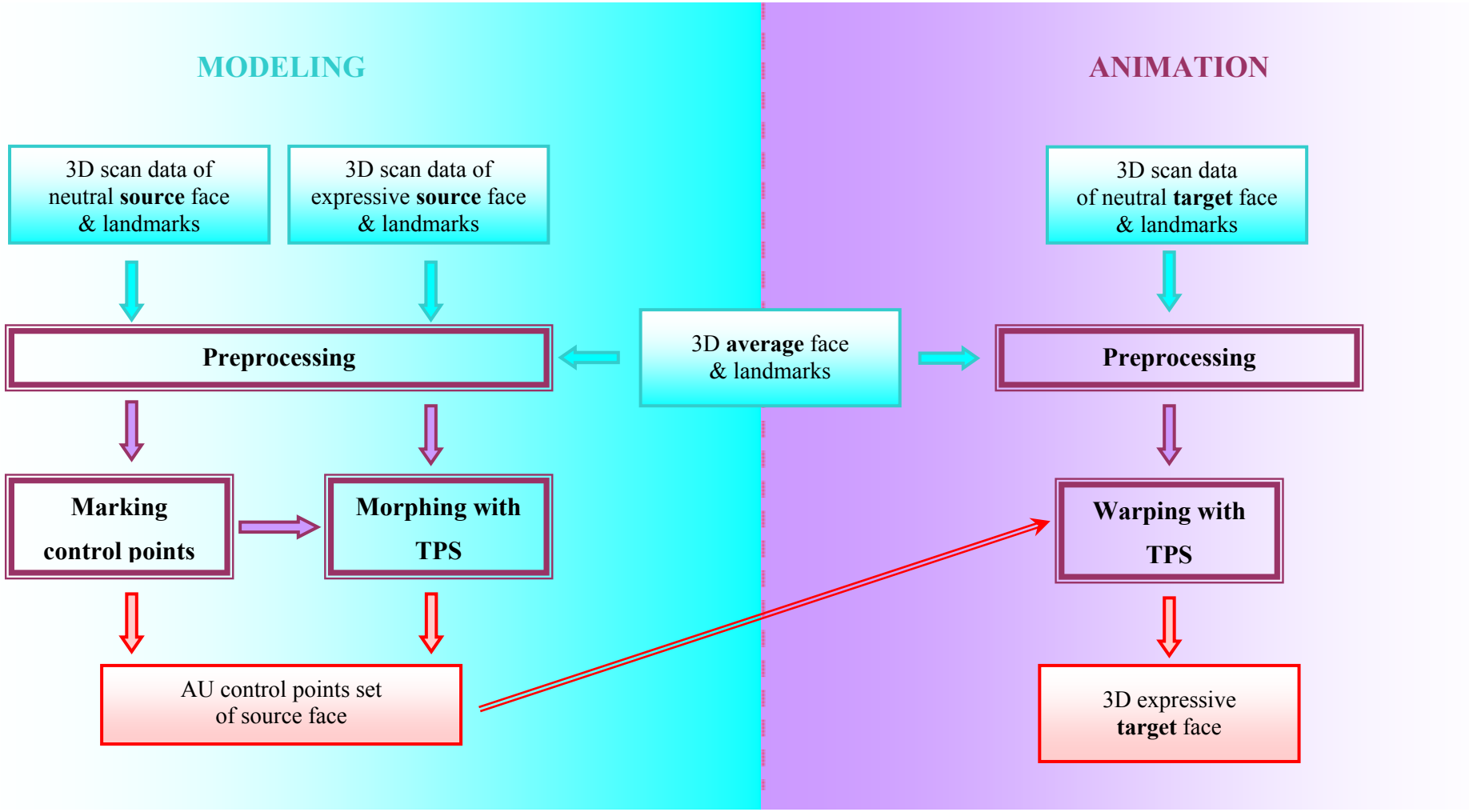


Figure 5-1: Flow chart of the system

5.3. Bosphorus Database

Bosphorus Database [61] was used for the method developed here. It is a 3D face database that includes a rich set of expressions and systematic variations of poses from 105 subjects. Since various facial expressions as well as many Action Units are included and since many actors/actresses are incorporated to obtain more realistic expressions, Bosphorus Database is very suitable for facial expression synthesis. The action units are grouped into three sets: 20 lower face AUs, 5 upper face AUs and 3 AU combinations, which can be seen in Figure 5-3.

All subjects, which were used as source and target face models, have 3D scans of various expressions with marked landmarks and texture data, which were acquired using a commercial structured-light based 3D digitizer device. Each scan was manually labeled for 24 facial landmark points, provided that they are visible in the given scan [61]. The facial landmarks that were used in this study are shown in Figure 5-2.

1-3. Outer/middle/inner left eyebrow
4-6. Inner/middle/outer right eyebrow
7-8. Outer/inner left eye corner
9-10. Inner/outer right eye corner
11-12. Left/right nose saddle
13. Left nose peak
14. Nose tip
15. Right nose peak
16. Left mouth corner
17. Upper lip outer middle
18. Right mouth corner
19. Upper lip inner middle
20. Lower lip inner middle
21. Lower lip outer middle
22. Chin middle

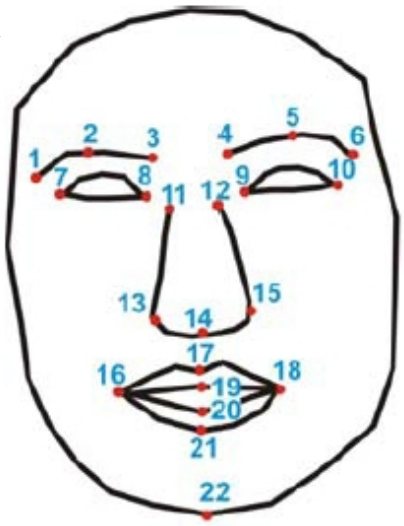


Figure 5-2: Facial landmarks [61]

At the preprocessing phase of the method, average face model (AFM) constructed for automatic 3D face recognition by Salah et al. [60] was used. To generate an AFM, they initially found a consensus landmark distribution on the

training faces. They rectified the landmarks of the consensus shape to present a fully frontal face, centered at the origin of the 3D coordinate system. Afterwards, they morphed each training face to the consensus shape by TPS. Before adding the training faces to the AFM, they cropped the faces to encompass the facial area. They sampled the depth values of the interpolated faces from a regular x-y grid. After adding all of the training faces to the AFM, depth values were simply averaged. The AFM constructed by Salah et al. can be seen in Figure 5-4.

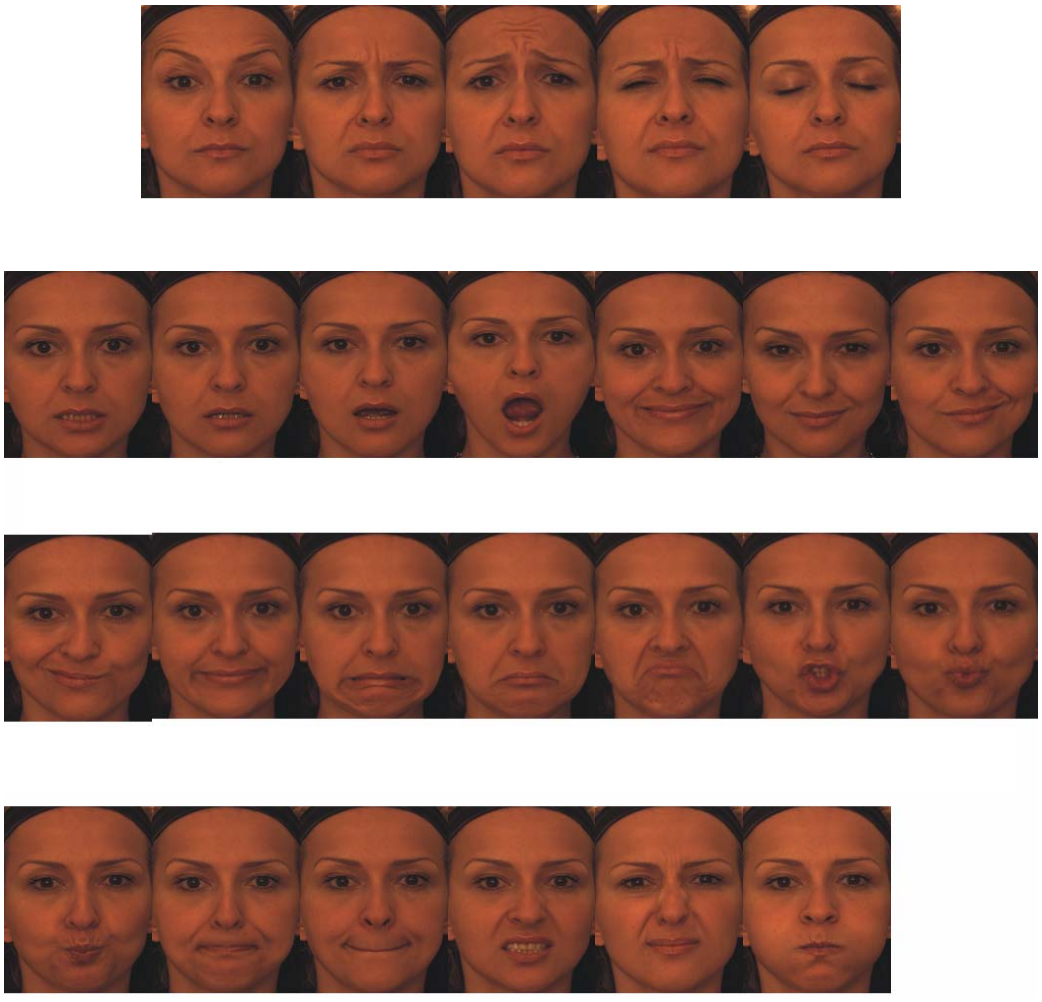


Figure 5-3: Upper face (first row) and lower face (second, third and forth rows) action units from Bosphorus Database [61]

Figure 5-4: The AFM of Bosphorus Database

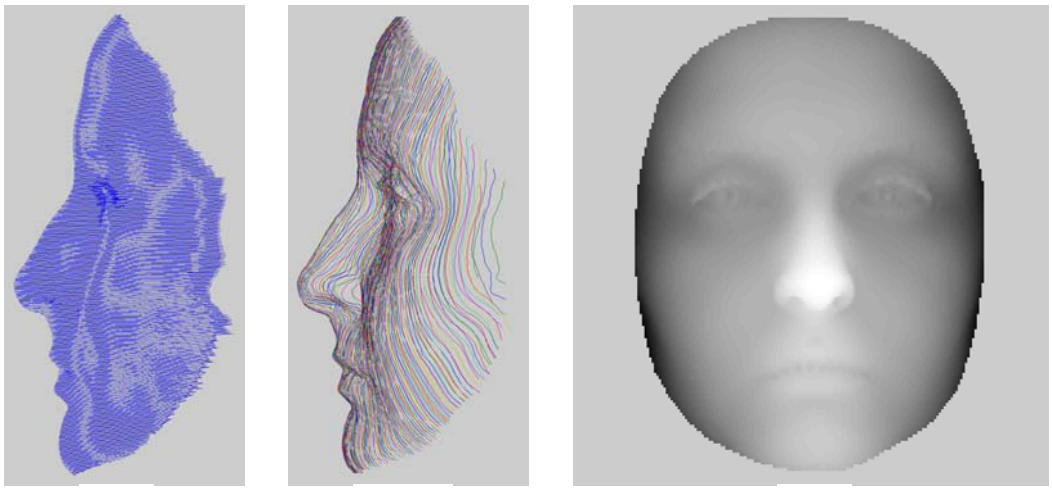
5.4. Preprocessing Phase

Before modeling and animating phases, the face models were preprocessed for pose and size normalization, using the average face model of the database as the reference data. The 3D model data was processed to improve the animation performance while keeping the high quality of the models. This preprocessing phase contains two steps. In the first step, the number of vertices and polygons in the face were reduced. A 3D face model in Bosphorus database consists of approximately 35K points in Cartesian space. For this work, background holes around faces were excluded and face point clouds were transformed into surfaces by linear interpolation using valid data points. x and y coordinates were fitted into 150x150 interpolated coordinates by `linspace` and `meshgrid` functions of MATLAB. Then z coordinate was found for these new interpolated coordinates by `griddata` function of MATLAB.

After this process, the face models contain a polygonal face mesh composed of 150x150 vertices, which are defined in u-v coordinates sequentially. The underlying rectangular grid induces four-sided patches on the surface, which defines a mesh of quadrilaterals or a quad-mesh. For the surface mesh, UV mapping, which maps the Cartesian coordinates to u-v coordinates, was

generated for the sake of further processes like marking control points on the surface and texture mapping over the surface. The steps of the first phase of the preprocessing are shown in Figure 5-5.

The landmarks of the face model, which can be found in the database, were transferred to the interpolated face by finding the nearest neighboring points to the original landmarks on the interpolated face model. For this purpose `DelaunayTri` and `nearestNeighbor` functions of MATLAB were used.



(a) (b) (c)

(d) (e)

Figure 5-5: A woman subject from Bosphorus DB: face mesh before (a) and after (b) preprocessing; facial surface without (c) and with (e) texture mapping; texture of the face (d)

In the second step, the faces were corrected around the x, y and z axes with respect to the pose and the size of the average face model (AFM) of the database [60] for normalization. At the beginning, the nose tips of faces were transformed to the same point, (0, 0, 0). Then the registered face model was transformed in 3D based on a transformation matrix. The affine transformation, which was defined between the outer eye corners and the chin middle vertices of the registered face model and the corresponding vertices of the AFM, was composed of linear transformations (rotation, scaling or shear) and a translation.

Define l_i, for $i = 1, 2, 3, 4$ are the landmarks of the face model for the left and right outer eye corners, chin middle and tip of the nose. Similarly, a_i, for $i = 1, 2, 3, 4$ are the landmarks of the AFM for the corresponding facial features. Since nose tips of all the models were transformed to (0, 0, 0), we can define 4×4 matrices as follows:

(5.1) $$A = \begin{bmatrix} a_{x,1} & a_{x,2} & a_{x,3} & 0 \\ a_{y,1} & a_{y,2} & a_{y,3} & 0 \\ a_{z,1} & a_{z,2} & a_{z,3} & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

(5.2) $$B = \begin{bmatrix} l_{x,1} & l_{x,2} & l_{x,3} & 0 \\ l_{y,1} & l_{y,2} & l_{y,3} & 0 \\ l_{z,1} & l_{z,2} & l_{z,3} & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

The transformation matrix can be computed as:

(5.3) $$T = AB^{-1}$$

Then we can transform the (x, y, z) vertex of a face by the following function:

(5.4) $$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = T \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

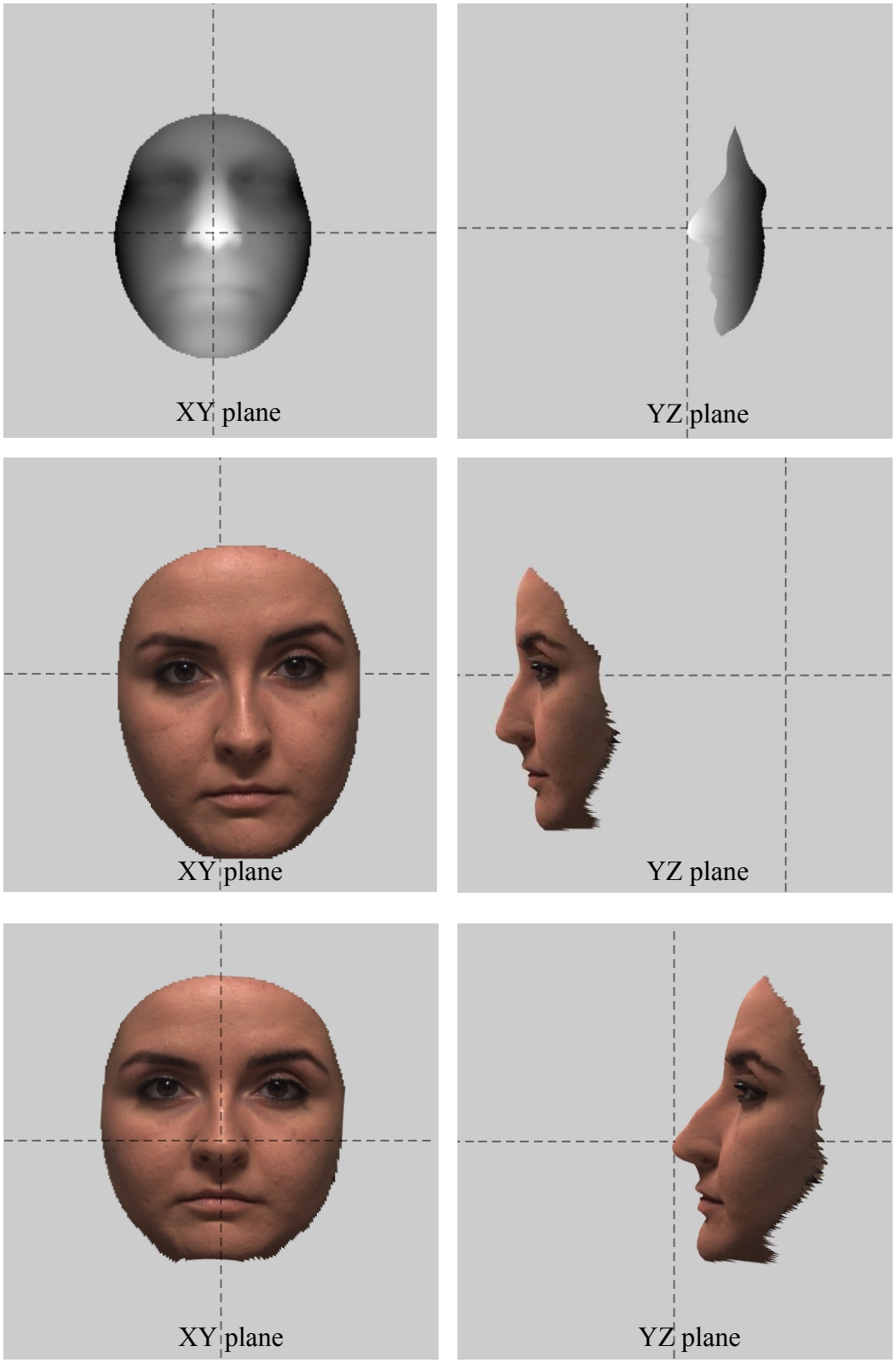


Figure 5-6: Example of normalization: average face model (top); a face model before (middle) and after (bottom) normalization phase

Preprocessing was applied to all the face models in the database, both neutral and expressive. During these preprocessing phases, landmarks of the faces were also transformed using the same transformation matrix. An example of normalization is shown in Figure 5-6.

5.5. Modeling Phase

Before animating a target face, source and target face models were modeled for further processes. After preprocessing, muscle and wrinkle control points were located to the source face with neutral expression according to the human anatomy. At the final step of the modeling phase, final positions of the control points after a facial expression were received from the expressive scan data of the source face.

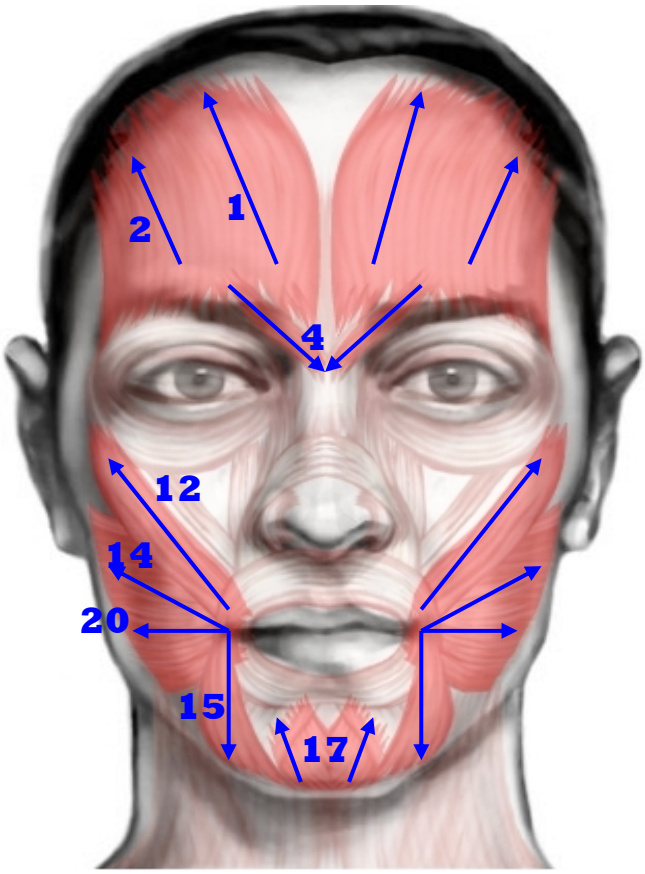
5.5.1. Marking Muscle Control Points

At the beginning of the process of marking control points, muscle control points were located manually to a selected face model considering the anatomical layout of the muscles on the face. FACS [38] associates AUs with the muscles underlying the actions covered. With the guidance of FACS figures, telling muscular anatomy and muscular action, control points were placed where the muscle fibers were anatomically located.

To find the positions of the generic muscle control points with respect to facial landmarks, several iterations were done. First muscle control points were positioned, and then source face was animated by warping with TPS, which is described in the following sections. The results were compared with the real expressive source face. When the results were not natural, the positions of the control points were changed. Working on a few face models, the positions were fixed with respect to facial landmarks and defined generally for all the face models.

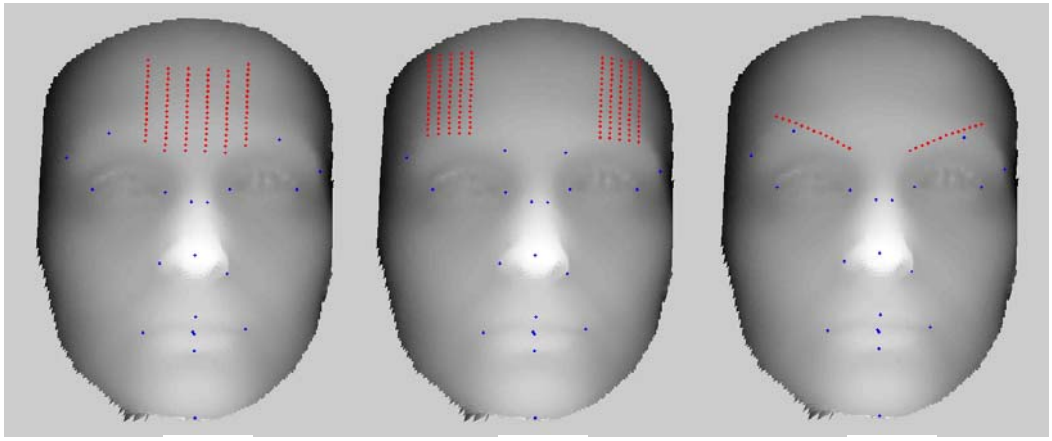
Five lower face and three upper face action units were simulated throughout this study. The simulated AUs are AU1-Inner Brow Raiser, AU2-Outer Brow Raiser, AU4- Brow Lowerer of the upper face; AU12- Lip Corner Puller, AU14-

Dimpler, AU15- Lip Corner Depressor, AU17- Chin Raiser, AU20- Lip Stretcher of the lower face. Schematic of all implemented action units and related muscles can be seen on Figure 5-7. The positions of the muscle control points for each simulated AU is also given in Figure 5-6.

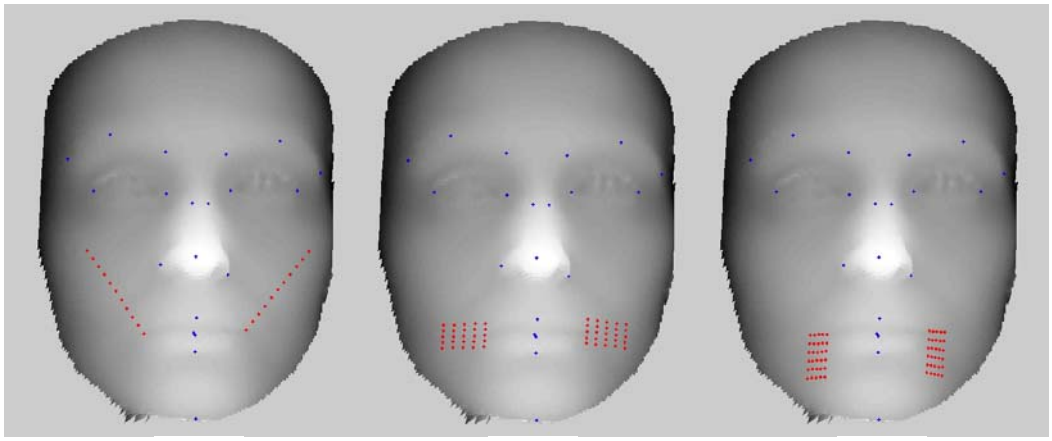


AU Number	FACS Name	Muscular Basis
1	Inner Brow Raiser	frontalis, pars medialis
2	Outer Brow Raiser	frontalis, pars lateralis
4	Brow Lowerer	corrugator supercilii; depressor supercilii
12	Lip Corner Puller	zygomatic major
14	Dimpler	buccinator
15	Lip Corner Depressor	triangularis
17	Chin Raiser	mentalis
20	Lip Stretcher	risorius

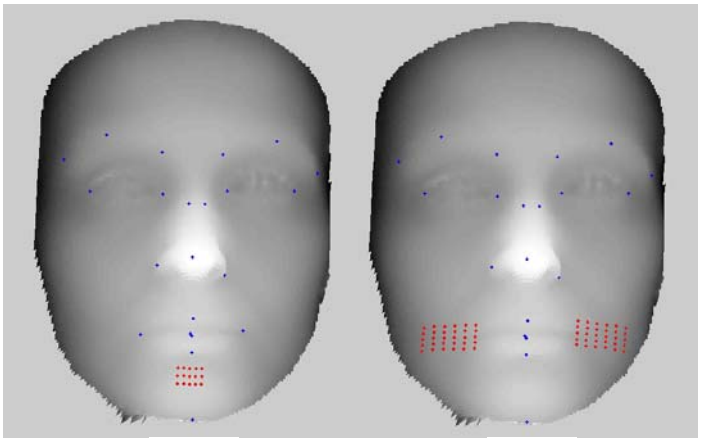
Figure 5-7: Schematic of all implemented action units and muscles [62]



(a) (b) (c)



(d) (e) (f)



(g) (h)

Figure 5-8: The positions of the muscle control points for (a) AU1 (b) AU2 (c) AU4 (d) AU12 (e) AU14 (f) AU15 (g) AU17 (h) AU20, control points and landmarks are seen in red and blue respectively

5.5.2. Marking Wrinkle Control Points

Bulges and wrinkles play an important role in the interpretation of the facial expressions. The wrinkles make the expressions more visible and they are generally used to distinguish between emotional expressions. Since only muscle control points were not enough to create wrinkles on the facial surface, additional wrinkle control points were placed on some critical regions of the face. Those critical regions were selected with respect to AU. For each AU, wrinkle control points were placed on the surface areas which mostly get wrinkled during the expression, such as the nasolabial furrow area for AU12 - Lip Corner Puller.

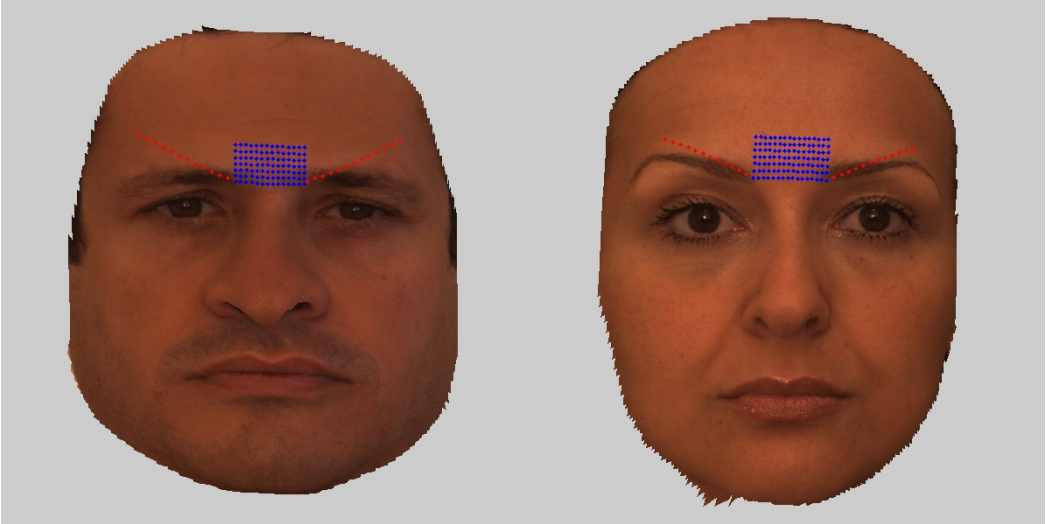


Figure 5-9: Muscle (red) and wrinkle (blue) control points for AU4 marked on male (left) and female (right) source models

With a similar approach to the position finding of the generic muscle control points, several iterations were done to find the positions of generic wrinkle control points with respect to the facial landmarks. After fixing the generic muscle control points as described in the previous section, the wrinkle control points were positioned. Then the source face was animated by warping with TPS. The results were compared with the real expressive source face. If the needed wrinkles cannot be created, the positions of the wrinkle control points

were changed. Working on a few face models, the positions were fixed with respect to the facial landmarks and defined generally for all the face models. A set of marked control points can be seen in Figure 5-9.

5.5.3. Transferring Control Points to Expressive Face

As mentioned in Chapter 4, RBF and TPS have been used to morph 3D face models and anatomical substructures in different researches [4, 14, 23, 58, 63]. In this thesis, instead of tracking motion control points during scanning human subjects, muscle and wrinkle control points positioned on the neutral source face model were transferred to the expressive source face model using the facial landmarks and TPS as morphing function. With this way, final positions of the control points after a facial expression can be received from the expressive scan data of the source face.

A small set of feature points was needed for TPS morphing function. The existing 3D landmarks in Bosphorus Database were used for this purpose. Additionally, automatic facial landmark detection algorithms can be used, which are commonly worked in the face recognition area [64, 65]; however, this is not in the scope of this study.

The neutral face and the expressive face must be aligned by means of semantic positions. The faces are oriented to look in the positive-z direction. The y-axis points through the top of the head and the x-axis points through the right ear. These preliminary conditions were obtained by the preprocessing steps, which are mentioned in Section 0.

Suppose S_n is the neutral source face which has a set of landmarks $P_i = (x_i, y_i, z_i)$ where $i = 1, 2, ..., n$ $(n = 22)$. After expression, the expressive source face S_e has the landmarks as $P_i' = (x_i', y_i', z_i')$. A continuous vector mapping function $f(x, y, x)$ from S_n to S_e is needed, such that

$$f(P_i) = P_i' \qquad \forall i = 1, 2, ..., n \ (n = 22) \tag{5.5}$$

Given such a mapping function, we can map each vertex of neutral face S_n into that of the expressive face S_e. Therefore, we can find the displacement of the control points with an expression.

The 3D TPS function given in Equation (4.19) and explained in Section 4.4.3 was used as the mapping function. K, P and L matrices, Equations (4.11), (4.10) and (4.11) respectively, were found using the landmarks $P_i = (x_i, y_i, z_i)$ of the neutral source face S_n and $r_{ij} = \|P_i - P_j\|$.

V and Y matrices, Equations (4.14) and (4.15) respectively, were found using the landmarks $P_i' = (x_i', y_i', z_i')$ of the expressive source face S_e. Using matrices L^{-1} and Y, coefficient matrices W and A can be found. The mapping function $f(x, y, z)$ S_n to S_e was defined using the elements of W and A matrices as in Equation (4.19).

The final positions of the control points were found by first applying the function $f(x, y, z)$ and then finding the nearest neighboring points to the point (x', y', z') on the expressive face model.

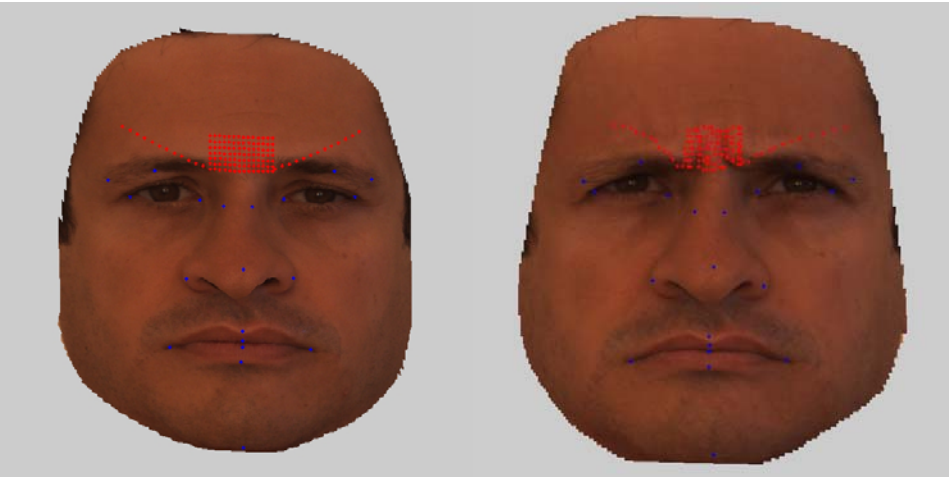


Figure 5-10: Transferring AU4 control points (red) from neutral source face (left) to expressive source face (right) using landmarks (blue)

For further processes $\vec{v_i}$, the displacement vector for each control point i, was calculated by subtracting the i^{th} control point of neutral source face S_n from i^{th} control point of expressive source face S_e:

$$\vec{v_i} = P_i' - P_i \tag{5.6}$$

Examples of transferring control points from the neutral source face to the expressive source face can be seen in Figure 5-9.

5.6. Animation Phase

After modeling source and target faces, the target face can be animated. At first, control points of the source face were transferred to the target face, by using TPS, as a similar process to the mapping control points on the expressive source face. Afterwards, neutral target face was animated using control points and TPS as warping function.

5.6.1. Transferring Control Points to Target Face

Similar to the mapping of the control points on the expressive source face, a similar approach was followed for the morphing of the control points from the source face to the target face, using the facial landmarks and TPS.

The source face and the target face must be aligned by means of semantic positions. The faces were oriented to look in the positive-z direction. The y-axis points through the top of the head and the x-axis points through the right ear. Those preliminary conditions were obtained by the preprocessing steps, which are mentioned in Section 0. The models were assumed to be at the neutral expression state when the lips were together and all the facial muscles were relaxed.

Suppose S_n is the neutral source face which has a set of landmarks $P_i = (x_i, y_i, z_i)$ where $i = 1, 2, ..., n$ $(n = 22)$. The neutral target face T_n has the

same landmarks with source face as $Q_i = (x_i', y_i', z_i')$. A continuous vector mapping function, $f(x, y, x)$ from S_n to T_n is needed, such that

(5.7) $$f(P_i) = Q_i \qquad \forall i = 1, 2, ..., n \ \ (n = 22)$$

Given such a mapping function, we can transfer each control point of source face S_n into that of the target face T_n.

The 3D TPS function given in Equation (4.19) and explained in Section 4.4.3 was used as the mapping function. K, P and L matrices, Equations (4.11), (4.10) and (4.11) respectively, were found using the landmarks $P_i = (x_i, y_i, z_i)$ of the neutral source face S_n and $r_{ij} = \left\| P_i - P_j \right\|$, the Euclidean distance between points i and j in 3D.

V and Y matrices, Equations (4.14) and (4.15) respectively, were found using the landmarks $Q_i = (x_i', y_i', z_i')$ of the neutral target face T_n. Using matrices L^{-1} and Y, coefficient matrices W and A can be found. The mapping function $f(x, y, z)$ from S_n to T_n was defined using the elements of W and A matrices as in Equation (4.19).

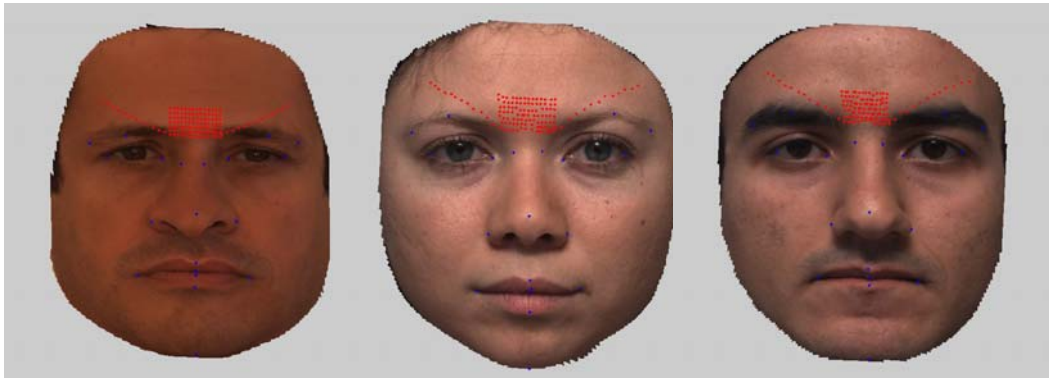


Figure 5-11: Transferring AU4 control points from a source face (left) to a female target face (middle) and a male target face (right); landmarks and expression control points are shown as blue and red respectively.

The control points on the neutral target face were found by first applying the function $f(x, y, z)$ to the control points of the source face and then finding the nearest neighboring points to the point (x', y', z') on the neutral target face model. Examples of transferring control points from a neutral source face to a neutral target face can be seen in Figure 5-11.

5.6.2. Warping Target Face with TPS

After control points were transferred from neutral source face to expressive source face and neutral target face, the control points of the neutral target face were moved and an expression was obtained by warping the face with TPS.

To improve the performance of the animation method, the face mesh was divided into the upper and lower face regions, where the AUs were presented according to FACS [38]. To be able to deform the face model by a particular motion, the algorithm has to search for all vertices on the face mesh that lie inside the influence zone of the motion. The region division was used to reduce this search area by skipping vertices that do not lie in the regions where the motion has effect on. This division also helps to prevent visual artifacts generated by the displacement of vertices in the regions that are not affected by the expressive movement. For example, without specifying the lower face region, the polygons of the mouth can be distorted by brow lowerer expression (AU4), which results in unnatural animations.

Suppose Tlf_n is the lower part of the neutral target face which has a set of control points $P_i = (x_i, y_i, z_i)$, which are found by the processes described in Section 5.6.1. After an expression, the positions of the control points on the lower part of the expressive target face Tlf_e can be found by the following function:

$$P_i' = (x_i', y_i', z_i') = P_i + (\vec{v_i} \times m \times {}^{s_t}\!/_{s_s}), \tag{5.8}$$

where $\vec{v_i}$ was calculated before in Section 5.5.3 by subtracting the i^{th} control point of neutral source face S_n from i^{th} control point of expressive source face S_e. Remember that the control points of the neutral source face were found in Section 5.5.1 and 5.5.2, and the control points of the expressive source face were found in Section 5.5.3.

Define m as the magnitude value of the expression, which has a value between 0 and 1. Define s_s and s_t as the mouth sizes of source face and target face respectively, which were calculated by the distances between the left and right mouth corners of the source and target faces.

A continuous vector mapping function $f(x, y, x)$ from Tlf_n to the lower part of the expressive target face Tlf_e is needed, such that

$$f(P_i) = Q_i \qquad \forall i = 1, 2, ..., n \tag{5.9}$$

Given such a mapping function, we can transform each vertex of the lower part of the neutral target face Tlf_n into that of the expressive target face Tlf_e.

The 3D TPS function given in Equation (4.19) and explained in Section 4.4.3 was used as the mapping function. K, P and L matrices, Equations (4.11), (4.10) and (4.11) respectively, were found using the control points $P_i = (x_i, y_i, z_i)$ of the lower part of the neutral target face Tlf_n, and $r_{ij} = \left\| P_i - P_j \right\|$, the Euclidean distance between points i and j in 3D.

V and Y matrices, Equations (4.14) and (4.15) respectively, were found using the positions of the control points $P_i' = (x_i', y_i', z_i')$ on the lower part of the expressive target face Tlf_e. Using matrices L^{-1} and Y, coefficient matrices W and A can be found. The mapping function $f(x, y, z)$ from Tlf_n to Tlf_e was defined using the elements of W and A matrices as in Equation (4.19).

The same algorithm was used for the upper part of the target face Tuf , where the only difference was the positions of the control points after the expression was found without multiplying with the mouth size ratio s_t / s_s :

$$P_i' = (x_i', y_i', z_i') = P_i + (\vec{v_i} \times m) \tag{5.10}$$

By changing the magnitude, m , different level of animations can be achieved like in Figure 5-12. If $m = 0$, the neutral face is reached after the animation. If $m = 1$, the expression level of the source face is reached. More animation results can be found in Chapter 6 and Appendix A. For the examples, face models were rendered with Gouraud shading based on MATLAB technologies.

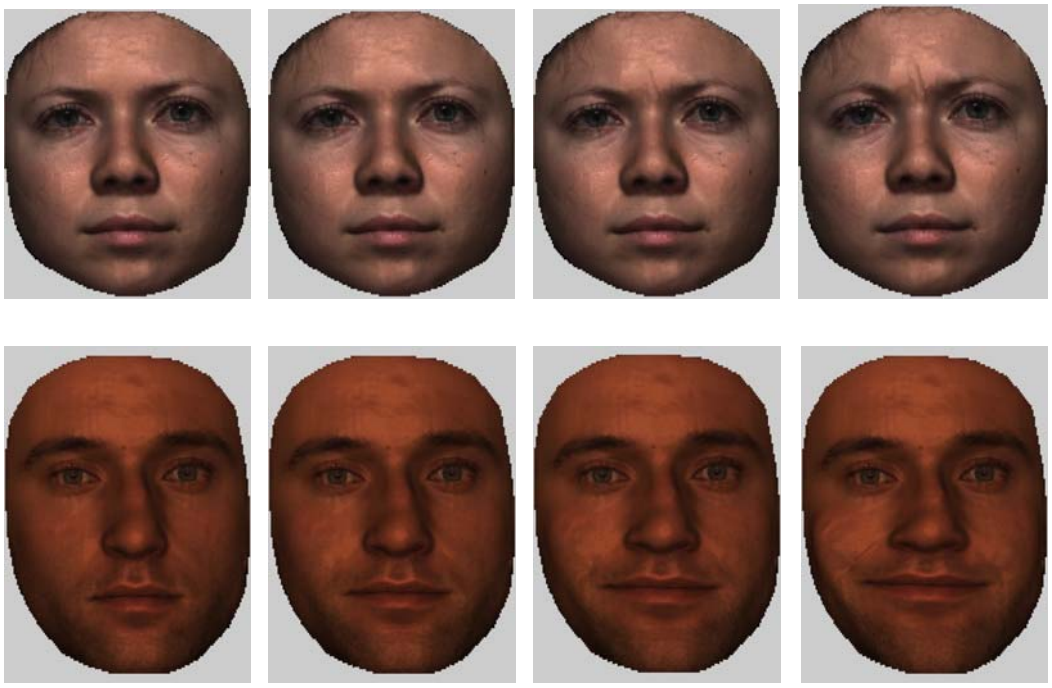


Figure 5-12: Animation of AU4 – Brow Lowerer (up) and AU12 – Lip Corner Puller (bottom) by changing magnitude m .

CHAPTER 6

EXPERIMENTAL RESULTS

In this thesis, a new three dimensional (3D) facial animation method is presented to morph emotions and facial expressions from one face to another. Using this method, realistic facial expressions were created with bulges and wrinkles. This was done by defining control points on a source face model according to the human anatomy and then transferring them to a target face, using the facial landmarks and TPS as morphing function. Target face was animated by deforming the facial surface, using control points and TPS as warping function. The result shows that our method has achieved the aim of generating and transferring natural and realistic facial expressions.

Facial expressions were created on 3D neutral face scan data of a human subject and these synthetic faces were compared to the real 3D scan data of the subject with the same facial expressions, which are available in the dataset. Examples of synthetic faces created with this technique can be seen in Figure 6-1, Figure 6-2, and Figure 6-3. More examples of animation sets are shown in Appendix A.

As seen in Figure 6-1, in the figure of AU4 (Brow Lowerer) corrugator muscles generate clearly visible wrinkles between the eyebrows when the eyebrows are frowned. The natural nasolabial bulges generated by the zygomatic muscles can be seen in Figure 6-2. The lips look smooth after distortion. Finally, the small bulges around the mouth can be seen in Figure 6-3, which makes the face more expressive. Since the faces were divided in two regions as upper and lower, the effects of the upper face action units cannot be seen in lower face and vice versa.

Figure 6-1: Animation of AU4 – Brow Lowerer: (a-b) FACS examples [47] (c) Corrugator muscle [62] (d) Neutral source face (e) Animated source face (f) Real expressive source face (g) Neutral target face (h) Animated target face (i) Real expressive target face

Figure 6-2: Animation of AU12 – Lip Corner Puller: (a-b) FACS examples [47] (c) Zygomatic muscle [62] (d) Neutral source face (e) Animated source face (f) Real expressive source face (g) Neutral target face (h) Animated target face (i) Real expressive target face

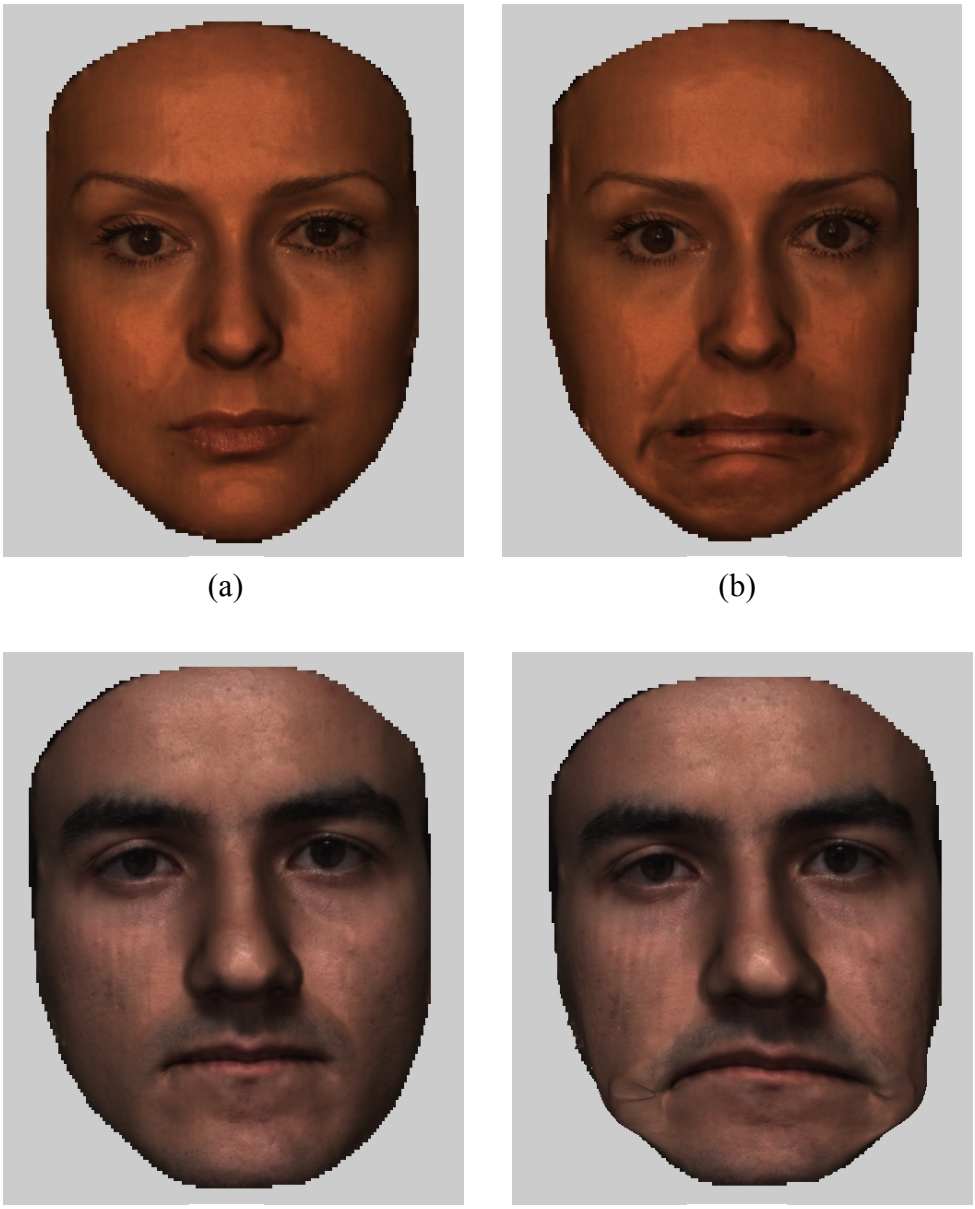
Figure 6-3: Animation of AU15 – Lip Corner Depressor: (a-b) FACS examples [47] (c) Triangularis muscle [62] (d) Neutral source face (e) Animated source face (f) Real expressive source face (g) Neutral target face (h) Animated target face (i) Real expressive target face

As mentioned before in Section 5.5.2, bulges and wrinkles play an important role in interpretation of the facial expressions. The wrinkles make the expressions more visible and they are used to distinguish between emotional expressions. By adding wrinkle control points on some critical regions of the face, we obtained more realistic faces. The small wrinkles which make the face more expressive for AU4 can be seen in Figure 6-4.

Figure 6-4: Animated expressive face without (left) and with (right) wrinkle control points

Through this study, some unnatural results were also observed. One of the main reasons was using the source faces with exaggerated expressions. An example of mapping an exaggerated expression from the source face to the target face can be seen in Figure 6-5.

If the expressive scan data of the source face for an asymmetric expression, like half-smiling, is used, such asymmetric expression can also be animated with our approach. For asymmetric expressions, the control points related to an AU were also used, but only the control points on one side of the face were involved. An example of an asymmetric animation can be seen in Figure 6-6.



(a) (b)

(c) (d)

Figure 6-5: Animation of AU20 – Lip Stretcher with an exaggerated expression: (a) Neutral source face (b) Real expressive source face (c) Neutral target face (d) Animated target face

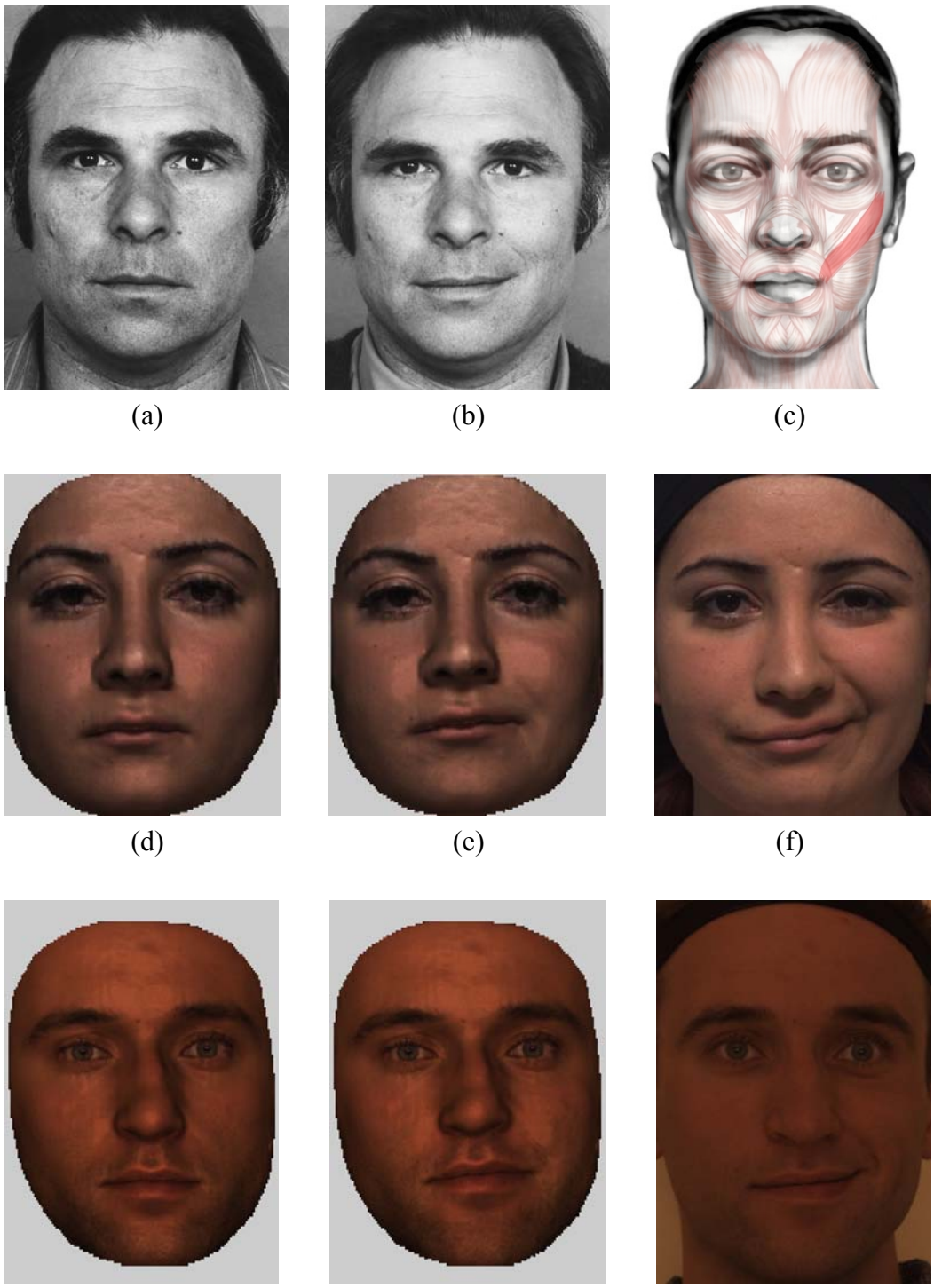


Figure 6-6: Animation of AU12 (left side only): (a-b) FACS examples [47] (c) Zygomatic muscle [62] (d) Neutral source face (e) Animated source face (f) Real expressive source face (g) Neutral target face (h) Animated target face (i) Real expressive target face

CHAPTER 7

CONCLUSION

7.1. Conclusions and Summary

As mentioned in the previous chapters, facial animation methods can be classified into three main groups as (i) image-based [5, 6], (ii) geometry-based [1, 7-11], and (iii) performance based [12-14] animations. Geometric manipulation methods include parameterization [7, 8], interpolation (key-framing) [7], pseudo muscle-based animation [9, 15], and physics-based animation [1, 10, 11]. Performance driven animation uses both image-based and geometry-based approaches to animate the face.

Interpolation and direct parameterization [7, 8] are old-fashion methods, which require extensive amount of computations and key frames. These approaches are also inflexible, since it is difficult to generalize the work on each different face meshes. Pseudo muscle approaches [9, 15] usually fail to model furrows, bulges, and wrinkles in the skin. Physics-based methods [1, 10, 11] need detailed anatomical knowledge of face, which is a complex assembly of bones, muscles, blood vessels, fatty tissue, connective tissue and skin. In image-based animation approaches [5, 6], a change of posture that requires appearance of features invisible in the input, such as wrinkles, cannot be done without further modeling. Additionally, realistic embedding into a 3D environment is not possible.

Performance driven animation [12-14] is the newest technique which can be implemented using both image and geometry-based methods. The motions are captured directly from a performer's face and then transferred to the virtual head

model. By this way, awkward key poses, one of the weak points of interpolation, is avoided. Accurate tracking of feature points or edges is also important to maintain a consistent and high-quality animation. Therefore performance driven techniques are largely data-driven as a result of capturing numerous facial markers.

In this thesis, a performance based animation method was proposed. Since a method was developed using the existing data of Bosphorus Database, capturing new data by tracking feature points was not needed. Therefore the main disadvantage of performance based animation approaches, being largely data-driven, was avoided. Instead of facial motion capture data, generic control points were transferred to the face models and used to animate them.

Till now, thin plate splines (TPS) have never been reported to be used to deform the facial surface and to create realistic expression by expression mapping. In this study, while conserving the advantages of performance based methods, TPS, which brings in smoothness and flexibility to the animation system, was used to animate the face. This brings one of the main advantages of our facial animation method, which is being morphable. This property cannot be achieved with interpolation and direct parameterization methods. As being morphable, the approach is easy to apply to any face model. A face model labeled with landmarks can be used as a source face model and the expressions can be transformed to another face model. Thus when we want to create animation on another face, there is no need to manually process that face model again.

Another reason of using TPS in animation is that TPS approach is superior to other approaches by giving user the chance of taking bulges and wrinkles into account, without merging the technique with any other wrinkle creation method. Pseudo muscles or parameterization approaches usually fail to create wrinkles because they ignore the underlying anatomy of the face. Also many performance based approaches [12, 13] cannot create wrinkles without an additional wrinkle creation approach. With the visibility of the wrinkles the expression looks more realistic and is easier to recognize. In addition, while other spline based methods usually compress the face model to define splines on the facial surface, we

worked on the detailed face model. Thus the synthetic faces created by our technique preserve the details derived by the 3D scanner, which results in more realistic animation.

The main disadvantage of the method is being source face oriented. Every single human has different face mimics. Since we are transferring the expressions from one face to another, characteristic of the source faces are also transferred with the algorithm. Small differences in wrinkles between the synthetic faces having same expression are formed due to different source faces, which can be observed on Figure 6-4. To dispose of this disadvantage, action units should be set on average face model, which is proposed as a future work.

Figure 7-1: AU4 – Brow Lowerer animation by using different source faces: source faces on the left, target faces on the middle and right

7.2. Future Work

The presented animation approach can be improved with some additional works, such as:

- The face mesh was divided into upper and lower face regions to improve the performance of the animation method. If these divisions were defined in more detail, then some visual artifacts, such as the change of nose with the movement of mouth, could be prevented.

- With an automatic additional landmark finding technique, success of morphing faces and transferring control points from one face to another one may be increased as done by Kähler et. al [4].

- For some of the subjects in the database, unexpected results were produced due to incorrectly positioned facial landmarks. Such results may be avoided with a graphical user interface (GUI) to correct landmarks and expression control points on the face.

- As mentioned in the conclusion, the main disadvantage of this method was being source face oriented. To dispose of this disadvantage, action units should be set on the average face model. In addition, since there would be no more calculation on each different source face, faster animation could be achieved.

In addition to these improvements, the animation system can be turned into a full "Facial Animation Software" by adding more features as listed below:

- Three upper and five lower face action units were studied in this study. Defining all the facial action units would be the first step of turning this work to a complete facial animation system. For this aim, orbicularis oris (muscle of the mouth), orbicularis oculi (muscle of the eyes) and the jaw rotation must be implemented.

- As seen in Figure 7-2, by merging muscle actions and AUs, emotions can be created. To create all the emotions, AUs must be put together in logical combinations with an additional technique, like Bui's approach [33] where the

muscle contractions in the face model was combined by simulating their parallelism.

- Since open mouth action units are not implemented, the synthetic faces created with this work do not have a tongue model. Therefore, a good tongue model would be necessary to implement open mouth action units and/or to make the articulation of the speech.

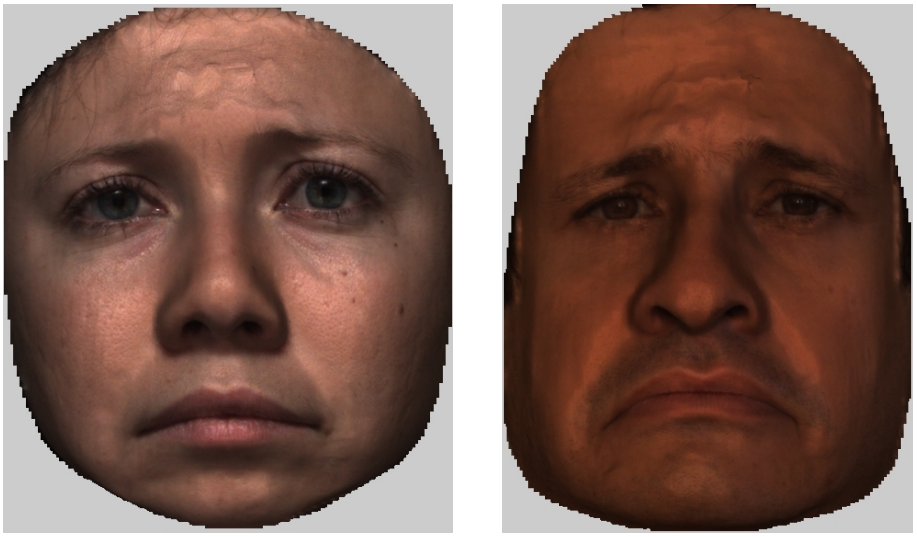


Figure 7-2: Sad emotion created by merging AU1 and AU15

REFERENCES

1. Waters, K., *A Muscle Model for Animation Three Dimensional Facial Expression.* Computer Graphics, 1987. **21**(4): p. 17-24.
2. IMDb. *Avatar (2009)*. The Internet Movie Database 01/05/2010]; Available from: http://www.imdb.com/title/tt0499549.
3. Sheppard, L.M., *Virtual Surgery Brings Back Smiles.* IEEE Computer Graphics and Applications, 2005. **25**(1): p. 6-11.
4. Kähler, K., J. Haber, and H.-P. Seidel, *Reanimating the Dead: Reconstruction of Expressive Faces from Skull Data.* ACM Transactions on Graphics, 2003. **22**(3): p. 554–561.
5. Beier, T. and S. Neely, *Feature-Based Image Metamorphosis.* Computer Graphics (Siggraph proceedings 1992). **26**: p. 35-42.
6. Pighin, F., et al. *Synthesizing Realistic Facial Expressions from Photographs*. in *25th Annual Conference on Computer Graphics and Interactive Techniques*. 1998.
7. Parke, F.I. and K. Waters, *Computer Facial Animation*. 1996.
8. Cohen, M.M. and D.W. Massaro, *Modeling Coarticulation in Synthetic Visual Speech*, in *Models and Techniques in Computer Animation*, N.M. Thalmann and D. Thalmann, Editors. 1993, Springer-Verlag: Tokyo.
9. Waite, C.T., *The Facial Action Control Editor, FACE: A Parametric Facial Expression Editor for Computer Generated Animation*. 1989, MIT.
10. Terzopoulos, D. and K. Waters, *Physically-Based Facial Modeling, Analysis, and Animation.* Journal of Visualization and Computer Animation, 1990. **1**(2): p. 73-80.
11. Platt, S.M. and N.I. Badler, *Animating Facial Expressions.* Computer Graphics, 1981. **15**(3): p. 245-252.
12. Williams, L., *Performance-Driven Facial Animation.* Computer Graphics, 1990. **24**(4): p. 235-242.
13. Guenter, B., et al. *Making Faces*. in *Computer Graphics Proceedings SIGGRAPH'98*. 1998.
14. Noh, J.-y. and U. Neumann, *Expression Cloning*, in *Computer Graphics (SIGGRAPH '01 Conf. Proc.)*. 2001: Los Angeles, CA, USA. p. 277-288.
15. Reeves, W.T., *Simple and Complex Facial Animation: Case Studies*, in *State of the Art in Facial Animation*. 1990. p. 88-106.
16. Radovan, M. and L. Pretorius. *Facial Animation in a Nutshell: Past, Present and Future*. in *SAICSIT*. 2006.
17. Lee, Y.C., D. Terzopoulos, and K. Waters. *Realistic Modeling for Facial Animation*. in *Siggraph*. 1995.

18. Choe, B., H. Lee, and H.-S. Ko, *Performance-Driven Muscle-Based Facial Animation.* The Journal of Visualization and Computer Animation, 2001. **12**(2): p. 67-79.
19. Sifakis, E., I. Neverov, and R. Fedkiw, *Automatic Determination of Facial Muscle Activations from Sparse Motion Capture Marker Data.* ACM Trans. Graph., 2005. **24**(3): p. 417-425.
20. Deng, Z., et al. *Animating Blendshape Faces by Cross-Mapping Motion Capture Data*. in *ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*. 2006.
21. Archer, K.M., *Craniofacial Reconstruction using hierarchical B-Spline Interpolation*, in *Department of Electrical and Computer Engineering*. 1997, University of British Columbia.
22. Orvalho, V.C.T., E. Zacur, and A. Susin. *Transferring Facial Expressions to Different Face Models*. in *Ibero-American Symposium on Computer Graphics - SIACG*. 2006.
23. Aina, O.O., *Generating Anatomical Substructures for Physically-based Facial Animation Part 1: A Methodology for Skull Fitting.* The Visual Computer, 2009. **25**(5-7): p. 617-625.
24. Noh, J.Y. and U. Neumann, *A Survey of Facial Modeling and Animation Techniques*, in *Technical Report*. 1998, University of Southern California. p. 99-705.
25. Pighin, F., R. Szeliski, and D.H. Salesin, *Modeling and Animating Realistic Faces from Images.* International Journal of Computer Vision, 2002. **50**(2): p. 143-169.
26. Sera, H., S. Morishima, and D. Terzopoulos, *Physics-based Muscle Model for Mouth Shape Control.* IEEE international Workshop on Robot and Human Communication, 1996: p. 207-212.
27. Deng, Z. and U. Neumann, *Computer Facial Animation: A Survey* in *Data-Driven 3D Facial Animation*. 2007, Springer London. p. 1-28.
28. Haber, J. and D. Terzopoulos, *Course Notes: Facial Modeling and Animation*, in *SIGGRAPH 2004*. 2004.
29. Kalra, P., et al., *Simulation of Facial Muscle Actions Based on Rational Free Form Deformations.* Eurographics, 1992. **11**(3): p. 59-69.
30. Nahas, M., H. Huitric, and M. Saintourens, *Animation of a B-spline Figure.* The Visual Computer, 1988. **3**(5): p. 272-276.
31. Wang, C.L.Y. and D.R. Forsey. *Langwidere: A New Facial Animation System*. in *Proceedings of Computer Animation*. 1994.
32. Wang, C.L.-Y., *Langwidere: A Hierarchical Spline Based Facial Animation System with Simulated Muscles*, in *Computer Science*. 1993, The University of Calgary. p. 109.
33. Bui, T.D., *Creating Emotions and Facial Expressions for Embodied Agents*. 2004, University of Twente. p. 223.
34. Kouadio, C., P. Poulin, and P. Lachapelle. *Real-Time Facial Animation based upon a Bank of 3D Facial Expressions*. in *Computer Animation 98*. 1998.
35. Sumner, R.W. and J. Popovic, *Deformation Transfer for Triangle Meshes.* ACM Trans. Graph., 2004. **23**(3): p. 399-405.
36. Burt, P.J. and E.H. Adelson, *A Multiresolution Spline with Application to Image Mosaics.* ACM Transactions on Graphics, 1983. **2**(4): p. 217-236.

37. Tarini, M., et al. *Texturing Faces*. in *Proceedings Graphics Interface*. 2002.
38. Ekman, P. and W.V. Friesen, *Facial Action Coding System*. 1978: Consulting Psychologists Press.
39. Ekman, P., W.V. Friesen, and J.C. Hager, *Facial Action Coding System Investigator's Guide*. 2002: Research Nexus division of Network Information Research Corporation.
40. Ostermann, J., *Animation of Synthetic Faces in MPEG-4.* IEEE Computer Animation, 1998: p. 49-55.
41. Sobotta, J., *Atlas of Human Anatomy*. 13 ed. Vol. 1. 2001: Lippincot Williams & Wilkins.
42. Gray, H., *Anatomy of the Human Body*, W.H. Lewis, Editor. 2000 (1918).
43. Marieb, E.N. and K. Hoehn, *Human Anatomy and Physiology*. 7 ed. 2007, San Francisco: Pearson Benjamin Cummings.
44. *Anatomy & Physiology Revealed*. 2009, The University of Toledo, The Mc-Graw Hill Companies.
45. Flores, V.C. *ARTNATOMY/ARTNATOMIA*. 01/05/2010]; Available from: http://www.artnatomia.net.
46. Hager, J.C. *DataFace*. 01/05/2010]; Available from: http://face-and-emotion.com.
47. Ekman, P., W.V. Friesen, and J.C. Hager, *Facial Action Coding System The Manual on CD-ROM*. 2002: Research Nexus division of Network Information Research Corporation.
48. Duchon, J., *Splines Minimizing Rotation-Invariant Semi-Norms in Sobolev Spaces*, in *Constructive Theory of Functions of Several Variables*. 1977, Springer: Berlin / Heidelberg. p. 85-100.
49. Buhmann, M.D., *Radial Basis Functions.* Acta Numerica, 2000: p. 1-38.
50. Buhmann, M.D., *Radial Basis Functions: Theory and Implementations*. Cambridge Monographs on Applied and Computational Mathematics, ed. P.G. Ciarlet, et al. 2003: Cambridge University Press.
51. Baxter, B.J.C., *The Interpolation Theory of Radial Basis Functions*, in *Trinity College*. 1992, Cambridge University.
52. Savchenko, V.V., et al., *Function Representation of Solids Reconstructed from Scattered Surface Points and Contours.* Computer Graphics Forum, 1995. **14**(4): p. 181-188.
53. Turk, G. and J.F. O'Brien, *Variational Implicit Surfaces*. 1999, Georgia Institute of Technology.
54. Turk, G. and J.F. O'Brien. *Shape Transformation Using Variational Implicit Functions*. in *ACM SIGGRAPH 99*. 1999.
55. Carr, J.C., W.R. Fright, and R.K. Beatson, *Surface Interpolation with Radial Basis Functions for Medical Imaging.* IEEE Transaction Medical Imaging, 1997. **16**(1): p. 96-107.
56. Carr, J.C., et al., eds. *Reconstruction and Representation of 3D Objects with Radial Basis Functions*. Proceedings of SIGGRAPH 2001. 2001. 67-76.
57. Bookstein, F.L., *Principal Warps: Thin-Plate Splines and the Decomposition of Deformations.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1989. **11**(6): p. 567-585.

58. Guo, H., J. Jiang, and L. Zhang, *Building a 3D Morphable Face Model by Using Thin Plate Splines for Face Reconstruction.* Sinobiometrics 2004, LNCS, 2004. **3338**: p. 258-267.
59. Blanz, V. and T. Vetter, *Face Recognition Based on Fitting a 3D Morphable Model.* IEEE Trans. on PAMI, 2003. **25**(9): p. 1063-1074.
60. Salah, A.A., N. Alyüz, and L. Akarun, *Registration of three-dimensional face scans with average face models.* Journal of Electronic Imaging, 2008. **17**(1): p. 011006.
61. Savran, A., et al., *Bosphorus Database for 3D Face Analysis*, in *The First COST 2101 Workshop on Biometrics and Identity Management (BIOID 2008)*. 2008: Roskilde University, Denmark.
62. Flores, V.C. *ARTNATOMY/ARTNATOMIA*. 01/05/2010]; Available from: http://www.artnatomia.net.
63. Kähler, K., et al. *Head Shop: Generating Animated Head Models with Anatomical Structure*. in *Proc. ACM SIGGRAPH Symposium on Comp. Anim. (SCA)*. 2002.
64. Wang, Y., C.-S. Chua, and Y.-K. Ho, *Facial Feature Detection and Face Recognition from 2D and 3D Images.* Pattern Recognition Letters, 2002. **23**: p. 1191-1202.
65. Hjelmas, E. and B.K. Low, *Face Detection: A Survey.* Computer Vision and Image Understanding, 2001. **83**: p. 236-274.

APPENDIX A

ANIMATION EXAMPLES OF ACTION UNITS

Face models were rendered with Gouraud shading based on MATLAB technologies for all example sets. The descriptions of the figure labels are:

(a) FACS example for neutral face [47]

(b) FACS example for expressive face [47]

(c) Muscle anatomy [62]

(d) 3D neutral source face

(e) 3D animated source face

(f) Real image of expressive source face

(g) 3D neutral target face – female subject

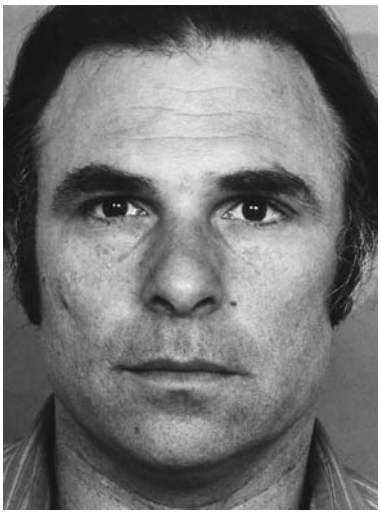
(h) 3D animated target face – female subject

(i) Real image of expressive target face – female subject

(j) 3D neutral target face – male subject

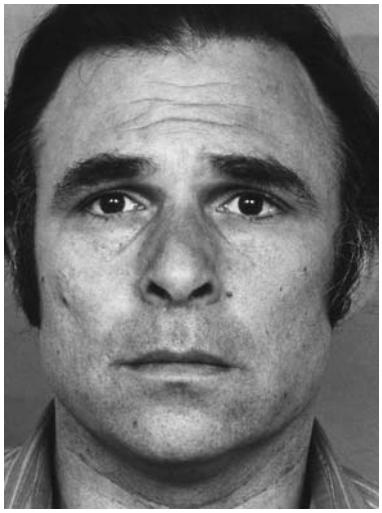
(k) 3D animated target face – male subject

(l) Real image of expressive target face – male subject



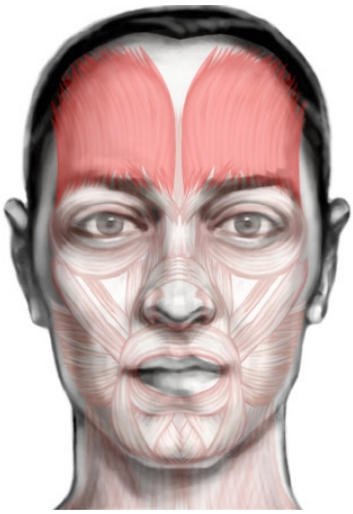
(a)

(d)



(b)

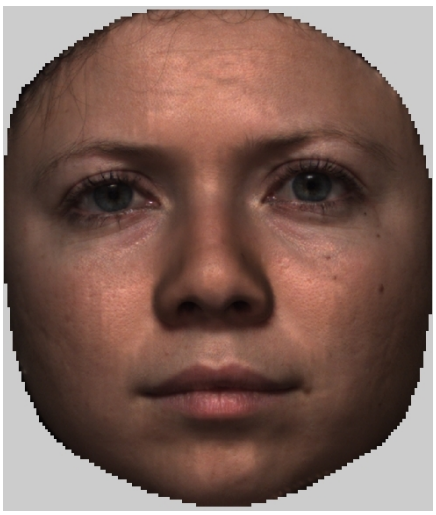
(e)



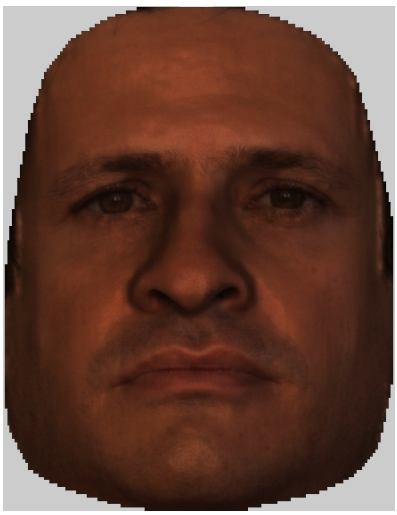
(c)

(f)

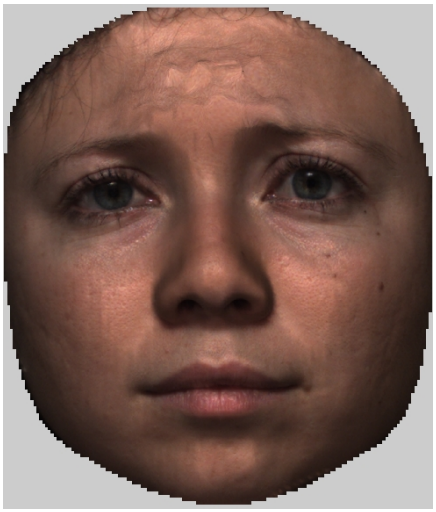
Figure A-1: Animation of AU1 – Inner Brow Raiser



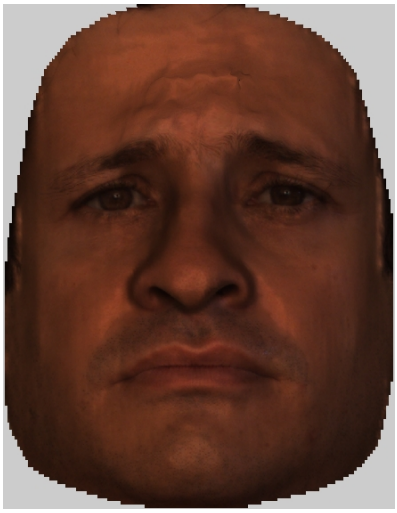
(g)



(j)



(h)

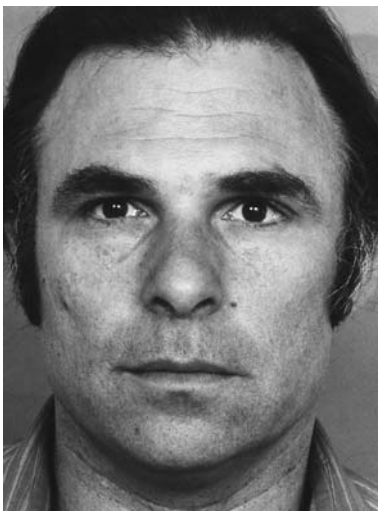


(k)

(i)

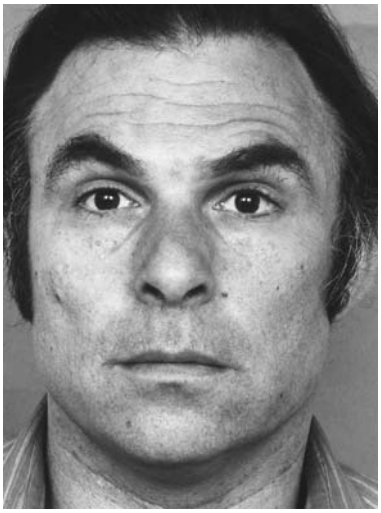
(l)

Figure A-1 (continued): Animation of AU1 – Inner Brow Raiser



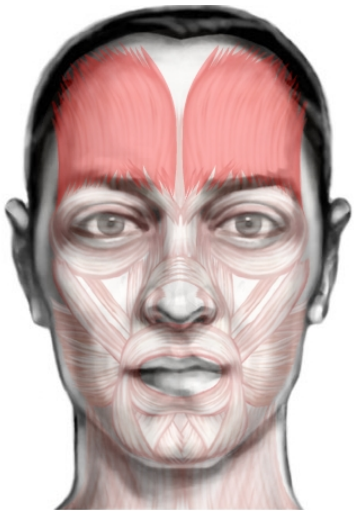
(a)

(d)

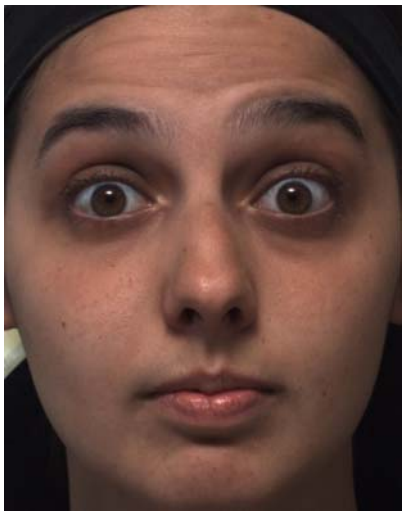


(b)

(e)



(c)



(f)

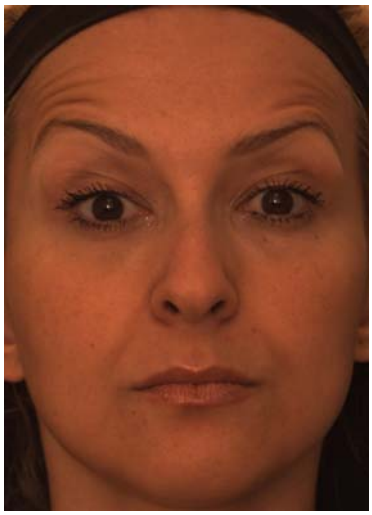
Figure A-2: Animation of AU2 – Outer Brow Raiser

(g)

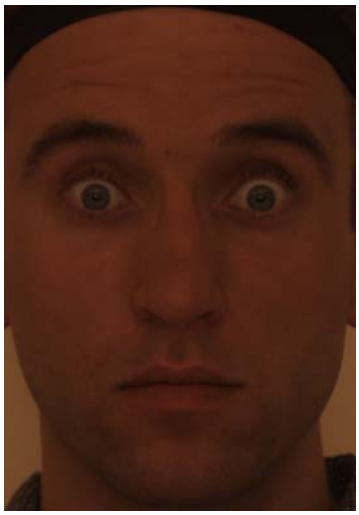
(j)

(h)

(k)

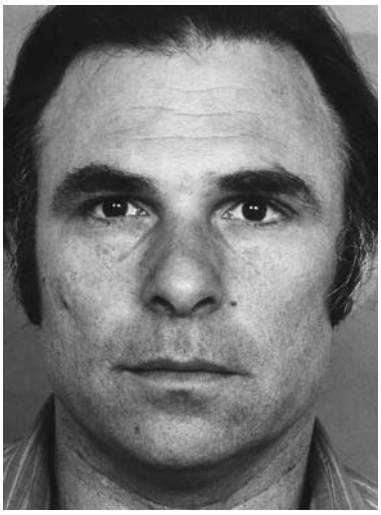


(i)

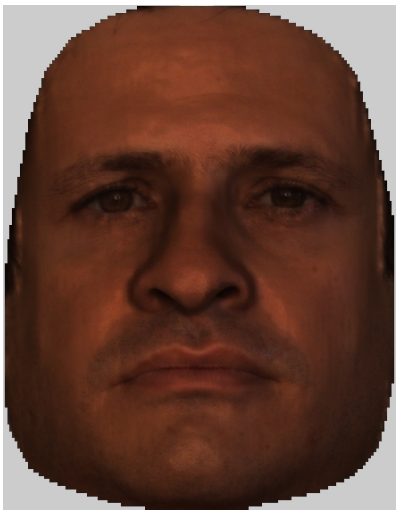


(l)

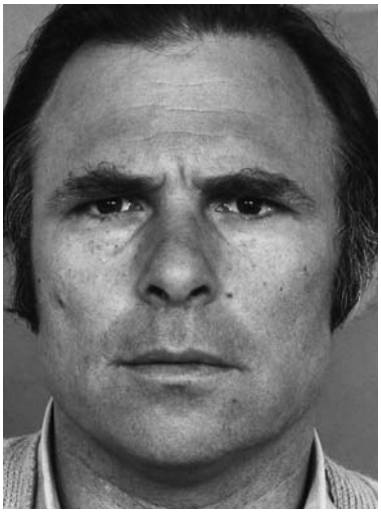
Figure A-2 (continued): Animation of AU2 – Outer Brow Raiser



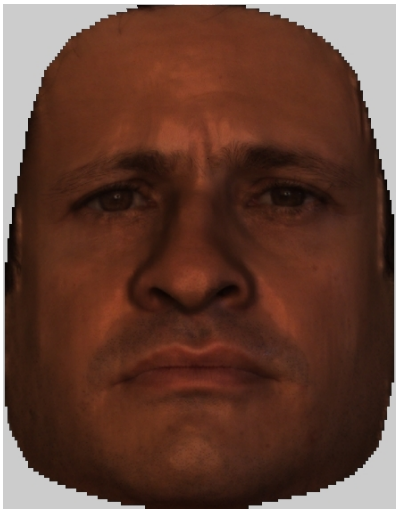
(a)



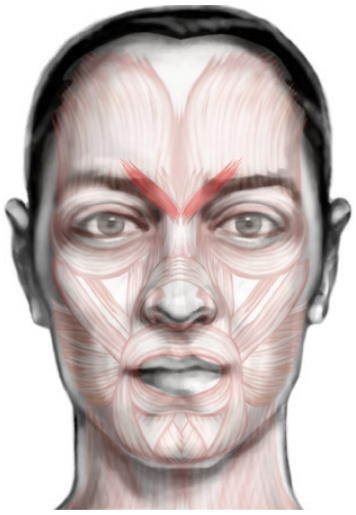
(d)



(b)



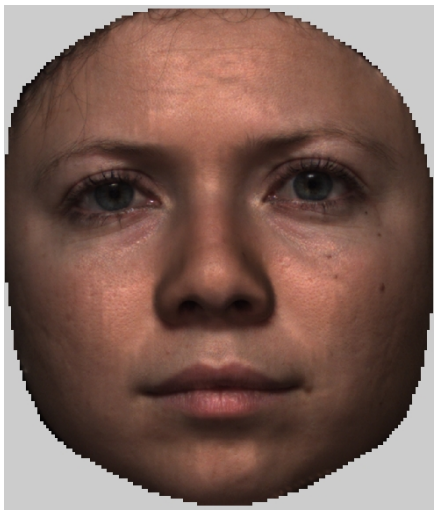
(e)



(c)

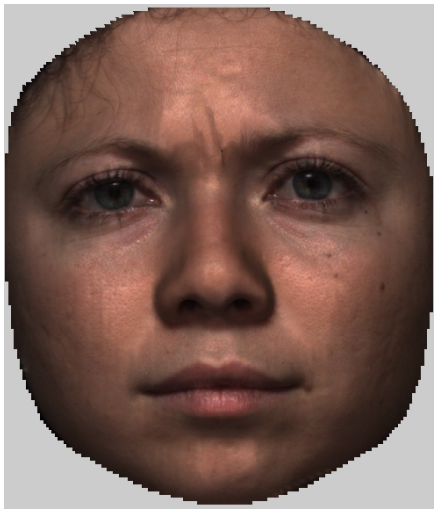
(f)

Figure A-3: Animation of AU4 – Brow Lowerer



(g)

(j)



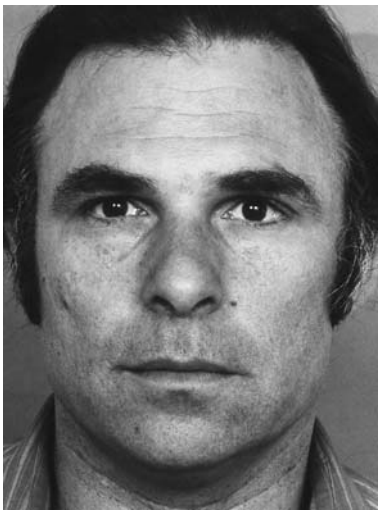
(h)

(k)

(i)

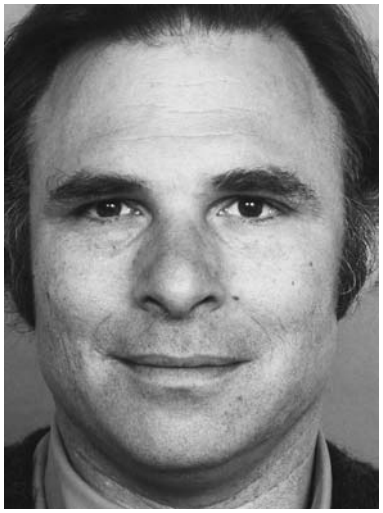
(l)

Figure A-3 (continued): Animation of AU4 – Brow Lowerer



(a)

(d)



(b)

(e)

(c)

(f)

Figure A-4: Animation of AU12 – Lip Corner Puller

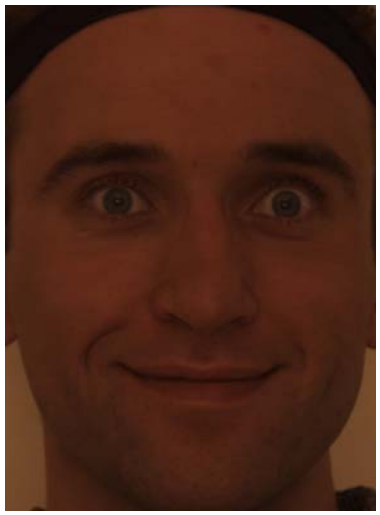
(g)

(j)

(h)

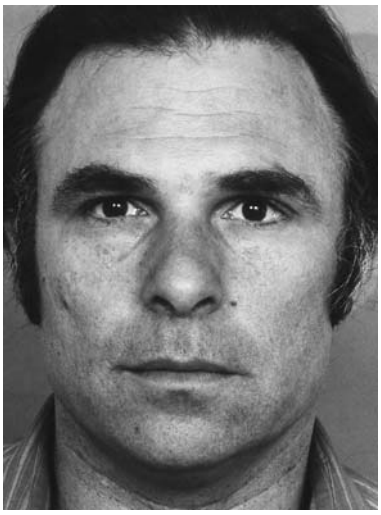
(k)

(i)



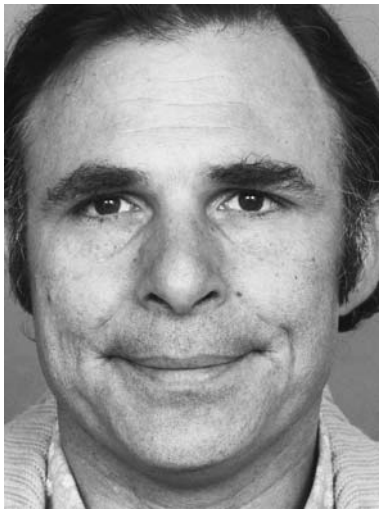
(l)

Figure A-4 (continued): Animation of AU12 – Lip Corner Puller



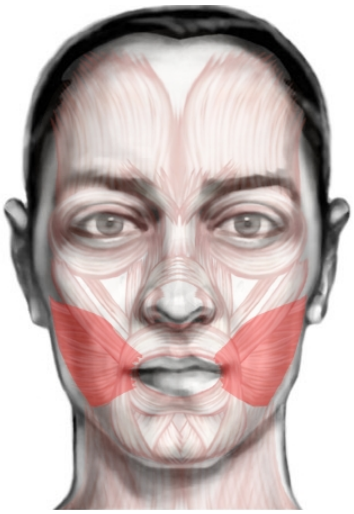
(a)

(d)



(b)

(e)



(c)

(f)

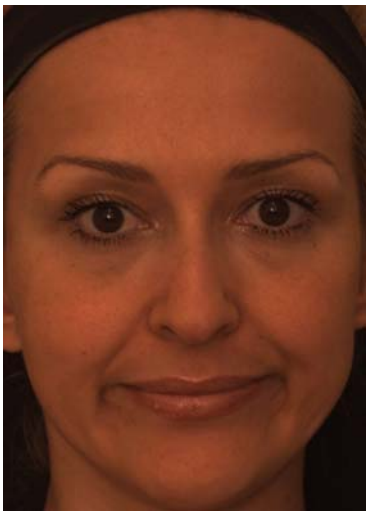
Figure A-5: Animation of AU14 – Dimpler

(g)

(j)

(h)

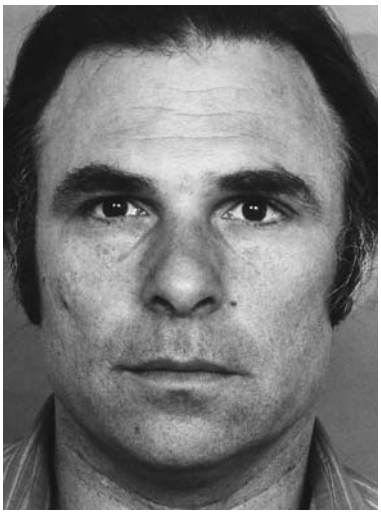
(k)



(i)

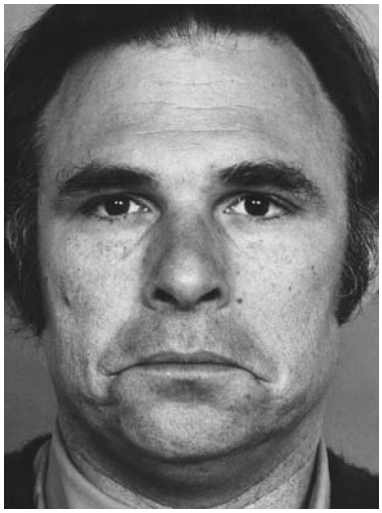
(l)

Figure A-5 (continued): Animation of AU14 – Dimpler



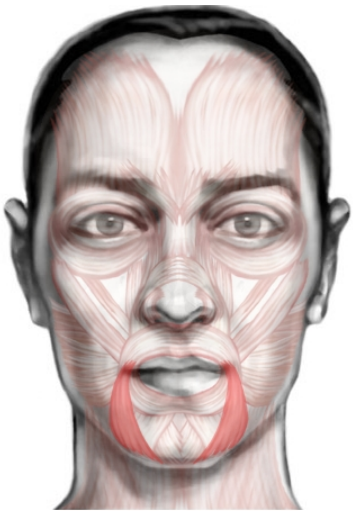
(a)

(d)

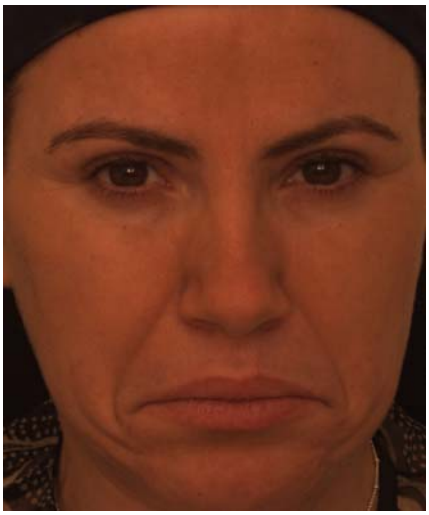


(b)

(e)

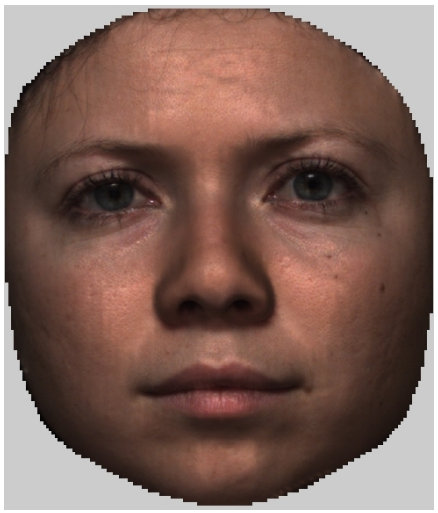


(c)

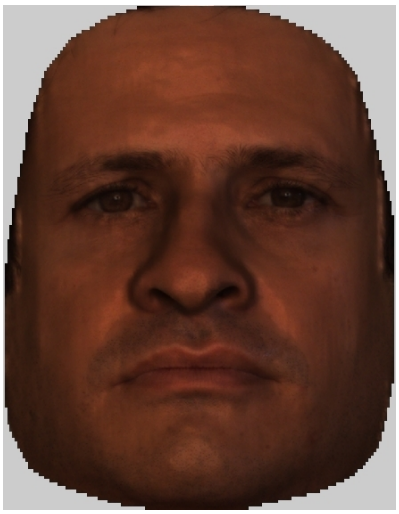


(f)

Figure A-6: Animation of AU15 – Lip Corner Depressor

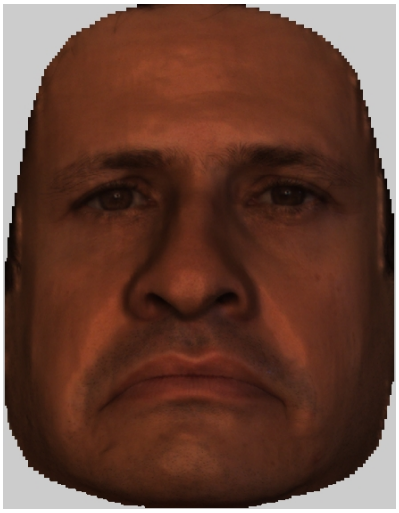


(g)



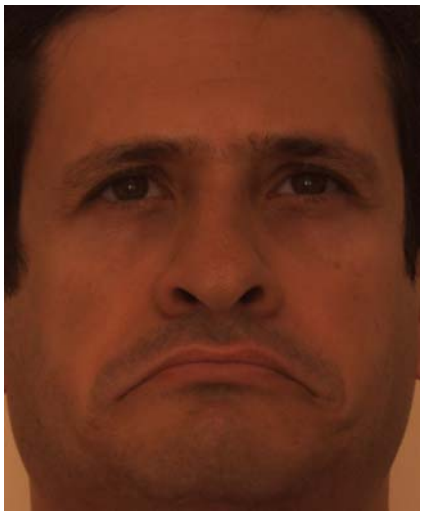
(j)

(h)



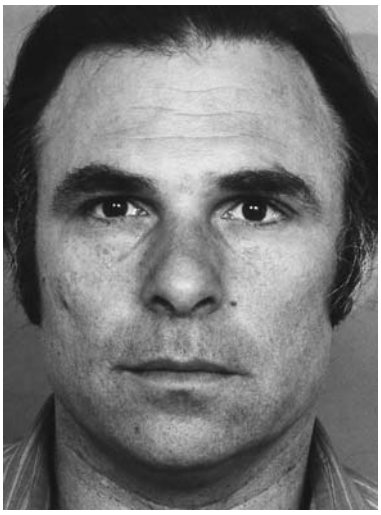
(k)

(i)



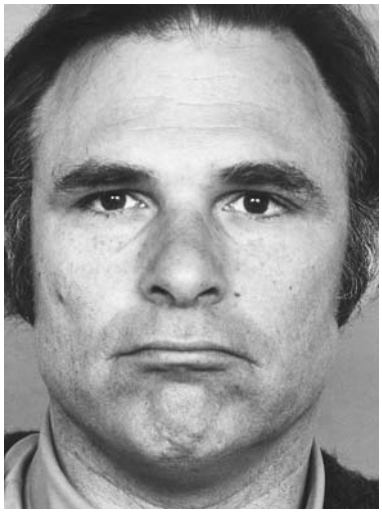
(l)

Figure A-6 (continued): Animation of AU15 – Lip Corner Depressor



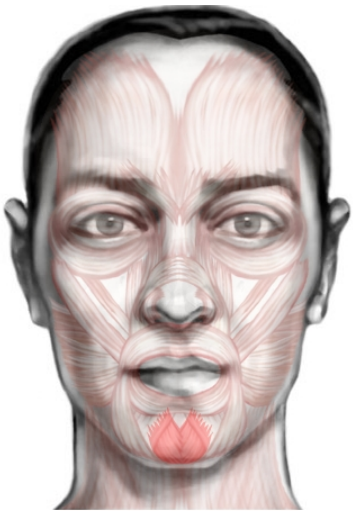
(a)

(d)



(b)

(e)

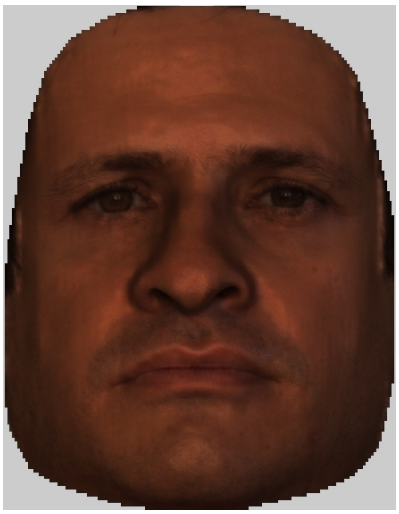


(c)

(f)

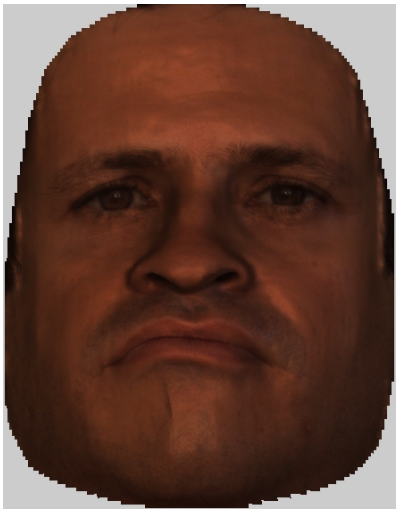
Figure A-7: Animation of AU17 – Chin Raiser

(g)

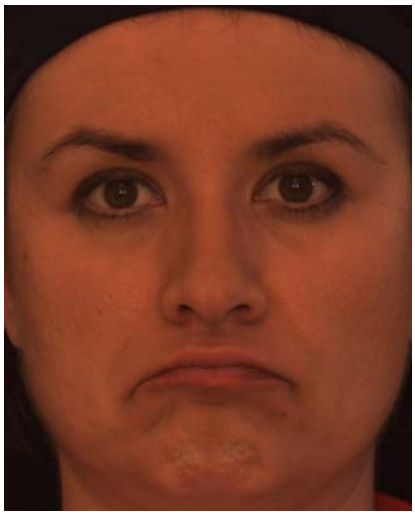


(j)

(h)



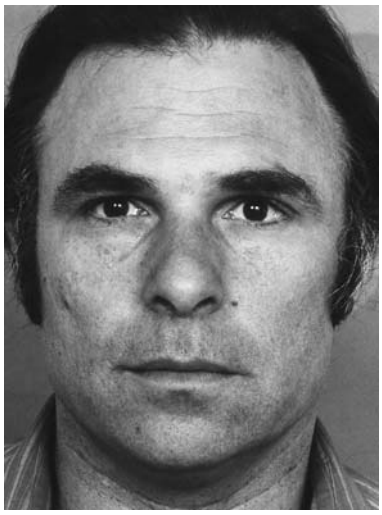
(k)



(i)

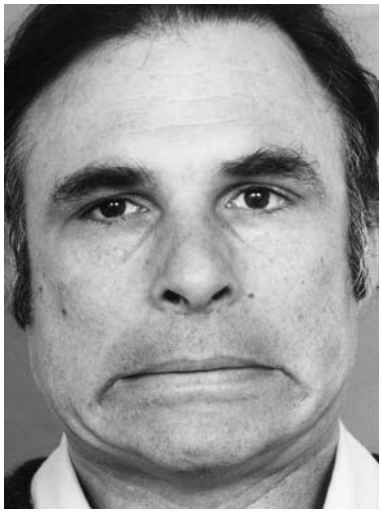
(l)

Figure A-7 (continued): Animation of AU17 – Chin Raiser



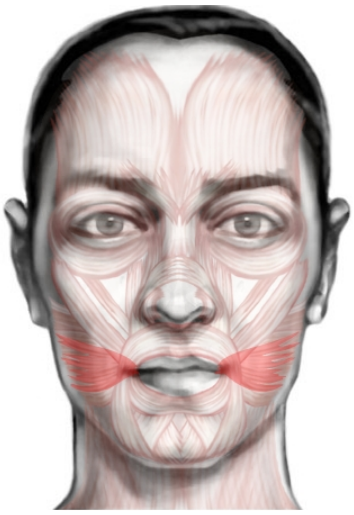
(a)

(d)



(b)

(e)



(c)

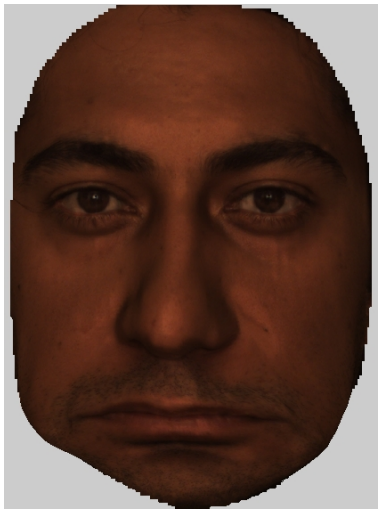
(f)

Figure A-8: Animation of AU20 – Lip Stretcher

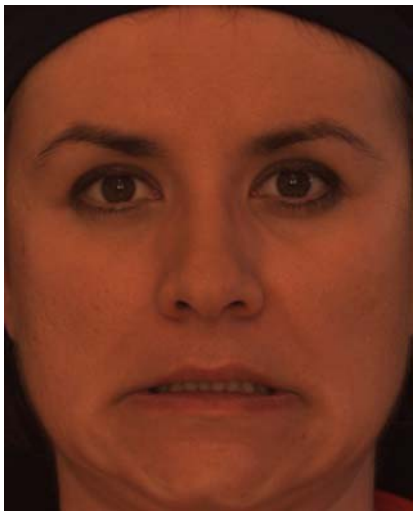
(g)

(j)

(h)



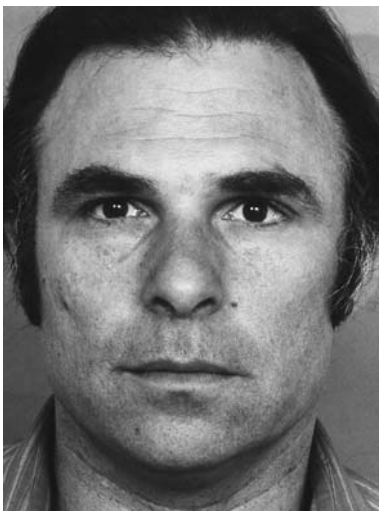
(k)



(i)

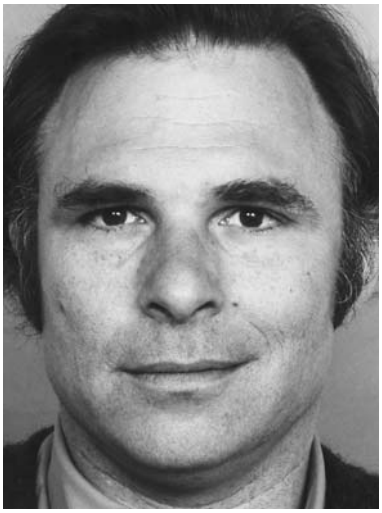
(l)

Figure A-8 (continued): Animation of AU20 – Lip Stretcher



(a)

(d)



(b)

(e)

(c)

(f)

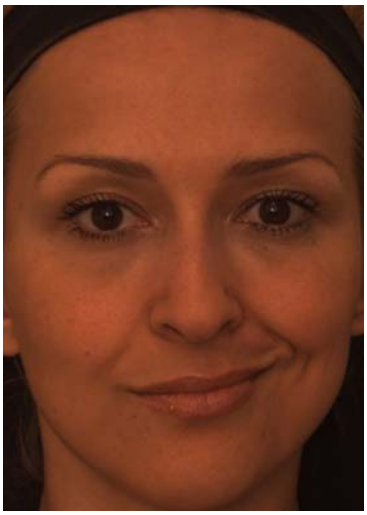
Figure A-9: Animation of AU12 – Lip Corner Puller (left side only)

(g)

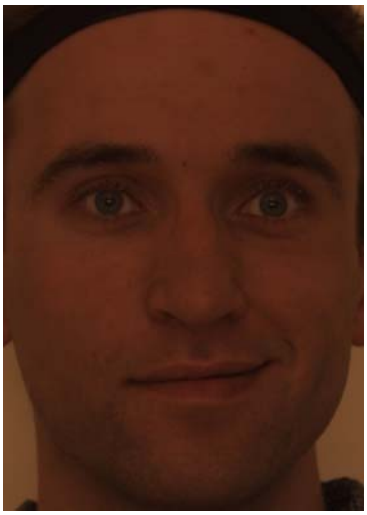
(j)

(h)

(k)



(i)



(l)

Figure A-9 (continued): Animation of AU12 – Lip Corner Puller (left side only)