

**MUSCLE-BASED FACIAL ANIMATION USING BLENDSHAPES IN  
SUPERPOSITION**

A Thesis

by

**ANDREW PATRICK SMITH**

Submitted to the Office of Graduate Studies of  
Texas A&M University  
in partial fulfillment of the requirements for the degree of

**MASTER OF SCIENCE**

December 2006

Major Subject: Visualization Sciences

**MUSCLE-BASED FACIAL ANIMATION USING BLENDSHAPES IN  
SUPERPOSITION**

A Thesis

by

ANDREW PATRICK SMITH

Submitted to the Office of Graduate Studies of  
Texas A&M University  
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee,  
Committee Members,  
Head of Department,

Frederic Parke  
Ricardo Gutierrez-Osuna  
Richard Davison  
Mark Clayton

December 2006

Major Subject: Visualization Sciences

## ABSTRACT

Muscle-Based Facial Animation Using Blendshapes in Superposition. (December 2006)

Andrew Patrick Smith, B.S., Texas A&M University

Chair of Advisory Committee: Dr. Frederic Parke

The blendshape is an effective tool in computer facial animation, enabling representation of muscle actions. Limitations exist, however, in the level of realism attainable under conventional use of blendshapes as non-intersecting deformations. Using the principle of superposition, it is possible to create a facial model with overlapping blendshapes and achieve more realistic performance. When blendshapes overlap, the region of intersection is in superposition and usually exhibits undesired surface interference. In such cases we use a corrective blendshape to remove the interference automatically. The result is an animatable facial model implemented in Maya which represents the effects of muscle action superposition. Performance created with our model of a known human subject is compared to 3D scan reference data and video reference data of that person. Test animation is compared to video reference footage. The test animation seems to mimic the effects of actual muscle action superposition accurately.

## ACKNOWLEDGEMENTS

I would especially like to thank Dr. Fred Parke, my committee chair, for his guidance and effort through this process. Dr. Parke's persistent mentoring and friendship were a priceless resource. Thanks also go to members of my thesis committee, Dr. Ricardo Gutierrez-Osuna and Professor Richard Davison. Their participation and feedback was very helpful.

Thanks go to Hiroki Itokazu for his advise regarding superposition. Lance Williams also provided helpful advice early in our process.

I would like to thank my mother, Valery Smith, my father, Gordon Smith, and my brothers, Aaron Smith and Benjamin Smith. This would not have been possible without their love and support. Thanks also go to my grandparents.

## KEY DEFINITIONS

(a) blendshape - In computer graphics, a tool which allows shape interpolation.

(b) superposition - For vector fields, the *principle of superposition* states that the net displacement at a given place and time caused by two or more waves traversing the same space is the vector sum of the displacements which would have been produced by the individual waves separately. If the resultant sum is greater than either individual wave, the event occurring when the waves meet is called constructive interference, and amplitude at that point is increased. When the resultant sum is less than either displacement, then destructive interference occurs, and overall amplitude decreases.

(c) realistic – The theory or practice of fidelity in art and literature to nature or to real life and to accurate representation without idealization.

## TABLE OF CONTENTS

		Page
ABSTRACT.....		iii
ACKNOWLEDGEMENTS.....		iv
KEY DEFINITIONS.....		v
TABLE OF CONTENTS.....		vi
LIST OF FIGURES.....		viii
 CHAPTER		
I	INTRODUCTION.....	1
II	PREVIOUS WORK.....	3
	2.1 Computer Facial Models.....	3
	2.2 Facial Animation Control.....	5
III	GOALS AND OBJECTIVES.....	7
	3.1 Objectives.....	7
IV	METHODOLOGY.....	9
	4.1 Studying the Anatomy of the Human Head and Face.....	9
	4.2 Development of an Expression Taxonomy.....	12
	4.3 Development of a Polygon Model Mesh .....	13
	4.4 Collection of Reference Data.....	13
	4.5 Superposition of Overlapping Blendshapes.....	14
	4.6 Demonstration of the Developed Model With Several Facial Animations .....	16
V	IMPLEMENTATION.....	18
	5.1 Implemented Muscle Actions.....	18
	5.2 3D Facial Scans.....	19
	5.3 The Animatable Model Polygon Topology.....	20
	5.4 Video Reference Data.....	21
	5.5 Muscle Action Blendshapes.....	24

CHAPTER	Page
5.6 Corrective Blendshapes.....	25
5.7 Three or More Muscles in Superposition.....	30
5.8 Hierarchy of Corrective Blendshapes.....	32
5.9 Corrective Blendshapes for the Other Facial Muscles.....	33
5.10 Supported Muscle Action Combinations.....	34
5.11 Test Animation.....	34
VI EVALUATION AND SUMMARY.....	36
VII FUTURE WORK.....	38
REFERENCES.....	40
APPENDIX.....	42
VITA.....	43

## LIST OF FIGURES

FIGURE		Page
1	Frontal view of the skull.....	9
2	Frontal view of facial muscles.....	10
3	Distribution of tension lines on the face.....	11
4	Principal muscles around the mouth.....	11
5	Blendshape superposition.....	15
6	3D scan data.....	19
7	The complete set of 3D scan data.....	22
8	First the left half of the face was modeled, using the neutral scan as a 'template.' .....	22
9	Frontal view of the animatable face topology.....	23
10	Frontal view of the facial topology rendered as subdivision surfaces with a Blinn shader applied.....	23
11	Side view of face.....	24
12	Video reference data of our human reference subject, David Morris.....	24
13	Upper facial muscle actions.....	26
14	Lower facial muscle actions.....	27
15	Applying a corrective blendshape for a pair of muscle action blendshapes: left inner brow raiser, right inner brow raiser.....	29
16	Three muscle action blendshapes in superposition.....	31



FIGURE	Page
17 Four muscle actions in superposition: outer brow raiser left, inner brow raiser left, inner brow raiser right, and outer brow raiser right.....	31
18 Animation was patterned after video reference data.....	35

## CHAPTER I

### INTRODUCTION

The human face is often a focus of interest in artistic works. Creating believable depictions of the face is a common goal in various media. Facial animation has been a specific area of research in computer graphics over the past thirty years. Most agree that we have yet to see 3D facial animation that is indistinguishable from actual human performance. It is currently more expensive to attempt to animate a human face realistically than it is to hire a human actor. However, there are situations where *realistic* (see definition (c)) human computer facial animation is needed [Hyneman et al. 2005; Williams 2005]. The best approach to creating convincing facial animation is to base the model and animation on the performance of a human actor.

Parameterized facial models are one approach to computer facial animation [Parke 1982]. These models describe facial expression as a set of input parameters. Muscle-based facial models use parameter sets which are patterned after human facial muscles [Waters 1987]. Shape interpolation is a common tool used in parametric models to create the appearance of muscle movement [Parke and Waters 1996]. A specific shape interpolation can be referred to as a *blendshape* (see definition (a)). Blendshapes can be modeled or sculpted using 3D animation packages to achieve desired shape transitions [Osipa 2003].

Blendshapes are typically isolated deformations on a virtual surface. The facial surface is usually divided into sections, one for each blendshape. In conventional facial modeling, blendshapes cannot intersect without creating unwanted surface interference in regions of shared vertices. These models are typically limited in their range of expression because muscle actions cannot overlap.

Recently, modelers have explored the possibility of using blendshapes in *superposition* (see definition (b)) to create more accurate human facial expressions [Hyneman et al. 2005; Williams 2005]. The human face is composed of many muscles in superposition over bone. It is possible to create more convincing facial animation if the effect of muscle action superposition is mimicked in the computer model. Accounting for superposition can mean an added level of complexity in the modeling process, but the end result is more natural facial movement. Also, this approach can facilitate a relatively straightforward animation control system.

## CHAPTER II

### PREVIOUS WORK

The first published studies of facial expression were conducted by John Bulwer in 1648 and 1649 in London [Haber and Terzopoulos 2004]. G. B. Duchenne published *The Mechanism of Human Facial Expression* in Paris [1862]. Duchenne demonstrated, by applying electrical stimuli to individual muscles on the face, that facial expressions were a result of specific muscle contractions in the face. He documented which muscle contractions generated specific facial expressions. This study indirectly demonstrated the concept of muscle-based facial animation in that muscle actions were being controlled by another individual to compose expressions. In 1872, *Expression of the Emotions in Man and Animals* was written by Charles Darwin [Ekman 1973]. In this study, Darwin attempted to define principles to explain why certain facial expressions occur with particular emotions.

#### 2.1 Computer Facial Models

The first 3D computer animation models of the human face were created by Parke in 1972 [Parke 1972]. Subsequent facial models have been developed over the years. Facial models utilize techniques from a variety of approaches. Some techniques have built on Parke's original work, while other models use different animation techniques such as facial rigs which use various kinds of shape deformers, forward

kinematics, or dynamic simulations [Noh and Neumann 1999; Haber and Terzopoulos 2004].

The blendshape was a key component of Parke's original model [Parke 1972]. Shape interpolation was done using linear interpolation of polygon vertex positions [Marschner 1999; Joshi et al. 2003]. The blendshape later became a staple of most 3D animation software packages. Parke later developed a parametric model for facial animation [Parke 1982]. As previously mentioned, a parametric model uses a set of input parameters to control the structure and expression of the face. A parameter may control the shape of a facial feature by manipulating one or more blendshapes.

Platt and Badler developed a muscle-based facial model at the University of Pennsylvania [1981]. Additional studies on muscle-based techniques were conducted by Waters in 1987, Thalmann, et al in 1988, and Waters and Terzopoulos in 1990. Before these muscle-based models, facial models dealt only with surface shape of the skin. The underlying motivation of dynamic properties of the face had previously not received attention. In Waters' model [Waters 1987], muscles are defined by muscle vectors, which specify the effect of muscle contraction on the shape of the skin surface.

Williams published a technique for capturing expressions from a human actor and mapping the expressions to a computer model [1990]. Williams later presented results of research conducted by Disney's Human Face Project [Haber and Terzopoulos 2004; Hyneman et al. 2005; Williams 2005]. A primary goal of the Human Face Project was to create facial animations which were indistinguishable from the performance of a well-known human actor. Williams outlined a technique for modeling facial expressions

using superposition of muscle blendshapes. These muscle blendshapes were modeled in superposition to match 3D scans of a reference actor. The details of this approach have not been published.

Hiroki Itokazu developed the technique of blendshape superposition used in the Human Face Project [Williams 2005; Itokazu 2005]. His approach involved correcting regions of surface interference resulting from blendshape superposition to achieve desired shapes by matching 3D scan data of the face. A ‘corrective’ blendshape could be created which contained the vector difference between the surface interference and the modeled target shape. This ‘difference’ blendshape could be used to correct the surface interference.

## **2.2 Facial Animation Control**

Systems have been developed to describe facial expression. In 1969 Hjortsjo developed the Mimic Language for describing facial expression [1969]. Hjortsjo’s system attempted to describe facial expression in terms of facial features, gestures, and postures. In the Mimic language, facial expression is the result of a combination of static structural elements of the face and dynamic elements, such as variation in form and appearance. Static aspects of expressions are determined by underlying bone structure and the composition of soft tissues on the face. Dynamic aspects of the facial expression result from the interplay of soft facial features and changes in shape. The system relates dynamic elements to the mental condition and emotional state of the individual. Soft

facial features are described by the set of Mimic muscles which correspond to facial expression patterns.

A more prevalent system for describing facial expression was developed by Ekman and Friesen [1978]. The Facial Action Coding System (FACS) was designed to describe basic facial muscle movements and their effects on facial expression. Though FACS was not designed for computer animation, it is a commonly used basis for controlling expression in facial animation models.

FACS categorizes facial expression as a set of muscle actions units, which are intended to describe all performable movement of the face. Examples of actions units include: the inner brow raiser, the outer brow raiser, the lid tightener, and the lip corner puller [Ekman and Friesen 1978]. These action units cannot be divided into smaller performable units. FACS was developed with the intent of reliably describing all visually distinguishable facial movements. Such movements may involve one or more muscle contractions.

Many facial animation models use FACS or similar systems as a way of specifying muscle actions needed to achieve desired facial expressions [Haber and Terzoploulos 2004]. The animator may compose facial performances by actuating FACS based muscle action controllers of the facial model.

## **CHAPTER III**

### **GOALS AND OBJECTIVES**

Our goal has been to develop a computer model for animating human facial expressions based on blendshape superposition. We planned to implement the model using an off-the-shelf 3D animation package, Autodesk Maya. Subsequently, the intent has been to produce several short clips of facial animation with this model. Our expectation has been that these animations would closely resemble scanned expressions and video reference of a human subject.

Another goal has been that the implementation be fairly efficient. Efficiency of our model is determined by the number of blendshape controls the animator must use. A graphical interface using existing Maya blendshape tools would be helpful. The objective is a high quality face model usable by experienced modeling and animation students.

#### **3.1 Objectives**

Our first objective is to develop and implement a muscle-based facial animation model which utilizes superposition of blendshapes. The model is intended to allow portrayal of fairly realistic 3D animation of facial expression. This included the following tasks.

- 1) Develop a polygon model of a face based on 3D scans and video reference of a human model.



- 2) Model facial expressions using superimposed blendshapes based on human facial muscle actions. Develop corrective blendshapes, finding a near fit for each target expression.
- 3) Use blendshape controls to compose and animate facial expressions.
- 4) Generate several short animations to demonstrate use and effectiveness of the developed model.

Our second objective is to develop an expression control taxonomy to help facilitate animation. This taxonomy is to consist of a description of blendshapes similar to the Facial Action Coding System such that:

- 1) Each muscle blendshape target is to be named according to the corresponding facial muscle action on which it is based.
- 2) Corrective blendshapes are to be named according to the muscle blendshapes involved in regions of superposition.

Our third objective is to develop a Maya graphical interface. This interface is to be based on Maya's blendshape editor. Names of muscle blendshapes and corrective blendshape controls are to follow the developed taxonomy of facial expressions.

## CHAPTER IV

### METHODOLOGY

Before implementing our facial model we developed an overall methodology for approaching the project. Guidelines were laid out to help navigate the problems encountered in the process. This is the plan of how we intended to accomplish our goals and objectives.

#### 4.1 Studying the Anatomy of the Human Head and Face

##### 1) Anatomical data:

Familiarity with skull structure and the mechanics of facial muscle movement has aided in developing a more anatomically accurate model [Fried 1976]. The intent was to form a conceptual basis for developing muscle blendshapes. Figure 1 shows

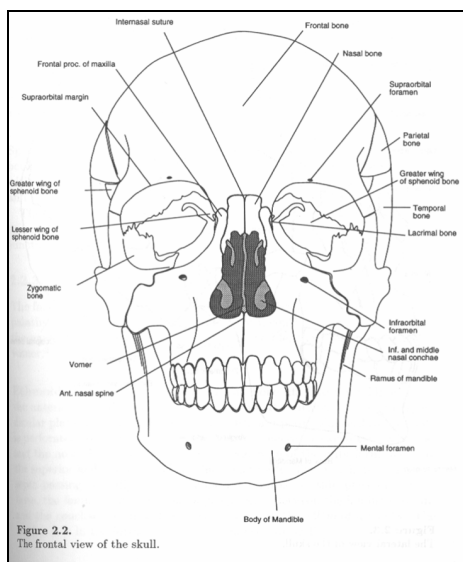


Figure 1. Frontal view of the skull.  
(from [Parke and Waters 1996], used with permission)

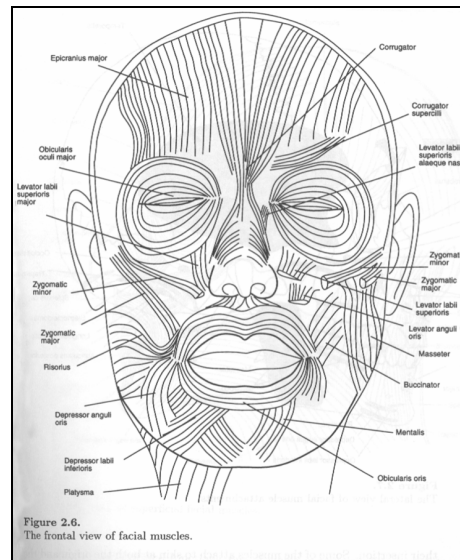


Figure 2. Frontal view of facial muscles.  
(from [Parke and Waters 1996], used with permission)

a frontal view of the human skull. Understanding this underlying structure of the face helped form a basis for modeling the face.

Figure 2 shows the facial muscles. Understanding how facial muscles are oriented beneath the skin was imperative to correctly modeling facial contours in the polygonal model. The skin has lines of tension that span the surface of the face (see Fig. 3). The polygons of the facial model need to follow a similar alignment. The way the polygons are laid out determines how the blendshapes can move. Muscle structure was studied to understand the motion in facial expressions. This knowledge proved useful when modeling muscle actions.

Figure 4 shows muscles of the mouth region. The mouth is an area of high flexibility that is greatly influenced by these surrounding muscles. Attention was given to this region to accurately capture its varied nuances from expression to expression. The

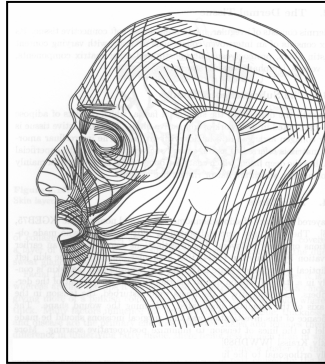


Figure 3. Distribution of tension lines on the face.  
(from [Parke and Waters 1996], used with permission)

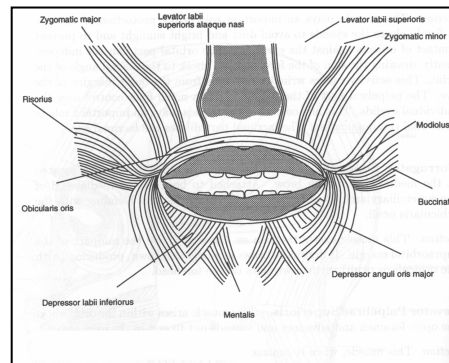


Figure 4. Principal muscles around the mouth.  
(from [Parke and Waters 1996], used with permission)

orbicularis oris, or the lip muscles, are not attached to bone, but to skin and other soft tissue. The lips are pushed and pulled by the surrounding muscles.

## 2) Artistic Basis:

Modeling facial expressions with a computer also draws on traditional artistic skills such as drawing, painting, and sculpting [Faigin 1990]. Accurate visual perception, dexterity, and aesthetics are needed to produce realistic visual work regardless of the medium.

Little artistic skill is needed to move vertices on the model to match the surface of a 3D scan. Artistic skill is helpful, however, in laying out the vertex contours so that muscle actions look realistic. Observation of video reference data aided in the modeling process. A certain level of comfort with the process of drawing or painting a portrait was useful when modeling the animatable face with a computer. Modeling utilizes similar skills of observation and hand-eye coordination.

#### **4.2 Development of an Expression Taxonomy**

Each blendshape name corresponds to the muscle action it represents. We have based our blendshapes on a subset of muscle actions from FACS. The selected subset is intended to be sufficient for demonstration of our method for handling overlapping blendshapes. Each blendshape is modeled to look as much like an actual muscle action as possible. The set of muscle actions is fairly large, including each side of the face. To have intuitive control of the blendshapes we have followed the FACS naming convention. This is intended to guide the animator when using the blendshape editor in determining which control slider activates a particular muscle action.

Each 3D scan and its associated animatable target shape is named according to its expression. Each scanned expression contained one or more muscle actions. Since the expressions did not always correspond to specific muscle actions, we named the scans according the general expression of the pose, for example: smile, sneer, and brows raised. This helped in differentiating 3D scans and copies of the polygon mesh, while maintaining a distinction from muscle actions.

### **4.3 Development of a Polygon Model Mesh**

The polygon mesh for a realistic facial model needs to be fairly detailed [Osipa 2003]. Some care must be taken when modeling to capture the detail of the face in such a way that it can be animated realistically. The polygons need to follow a pattern of contours around facial features that allow transitions between muscle actions to look realistic during motion. The polygons needed to follow a consistent orientation from one pose to the next.

The mesh must be flexible enough to capture all of the extreme target poses for a performance. The mesh was tested to verify that it could capture each expression. This was done by modeling extreme poses and making adjustments to the mesh so that it could conform to each expression. This testing phase allowed for details such as wrinkles, furrows, and dimples.

### **4.4 Collection of Reference Data**

Blendshapes were modeled in superposition to match 3D scan data as well as video reference data. Reference data allowed creation of a more realistic facial model.

#### 1) 3D Scan Data:

The animatable face model is based on 3D scans of a human reference subject made using the Cyberware 3D scanner in the Texas A&M Visualization Laboratory. A set of muscle blendshapes were modeled based on reference scanned data of the human subject. The polygon mesh was aligned with the 3D scan data and vertices were positioned on the scan surfaces to capture the likeness of the reference subject, David Morris, in each expression pose.

## 2) Video Reference Footage:

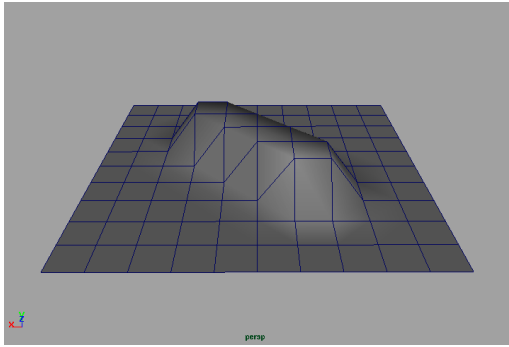
We collected video reference footage of our human subject to understand the movement of each muscle action. Since we modeled shapes in transition, the polygon mesh must accurately capture elements of motion. Video reference footage made this possible because 3D scan data captures only static aspects of the facial geometry. David Morris performed a wide range of expressions recorded on video which contained all of our intended muscle actions.

### **4.5 Superposition of Overlapping Blendshapes**

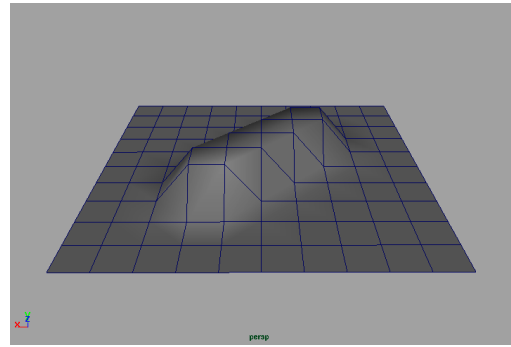
A technique of muscle-based facial animation using superposition of blendshapes is developed. This technique uses muscle action blendshapes and corrective blendshapes which are applied automatically to regions of superposition. The corrective blendshapes are connected to superimposed muscle action blendshapes using Maya's expression editor. This has involved development of a methodology for correcting surface interference resulting from overlapping blendshapes.

If two or more blendshapes overlap on a surface region, superposition occurs (see Fig. 5(a) and Fig. 5(b)). Erroneous shapes usually result (see Fig. 5(c)). The desired shape of the superposition region may look somewhat different (see Fig. 5(d)). We model a corrective blendshape for this region (see Fig. 5(f)). Another level of superposition is used to apply this correction.

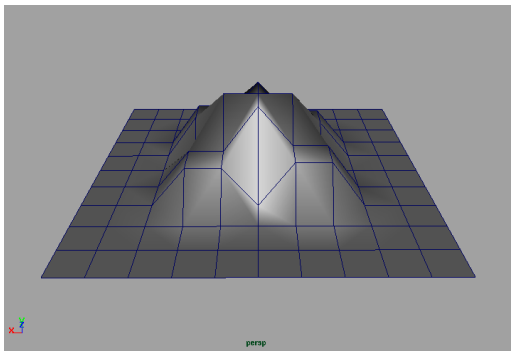
The corrective blendshape can be linked to the overlapping blendshapes using Maya's expression editor. This connection allows the correction to be actuated only as needed,



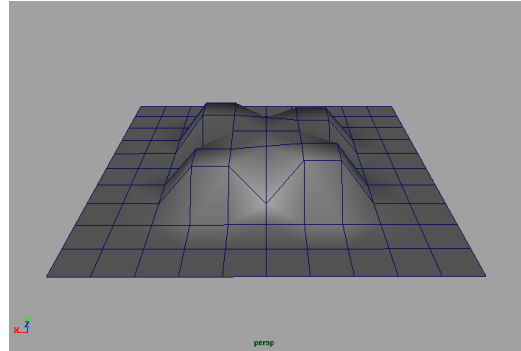
(a) Example blendshape A.



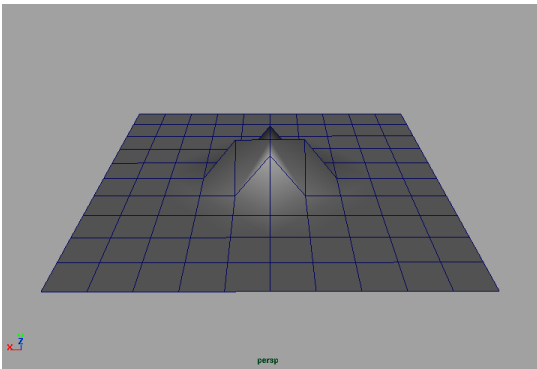
(b) Example blendshape B.



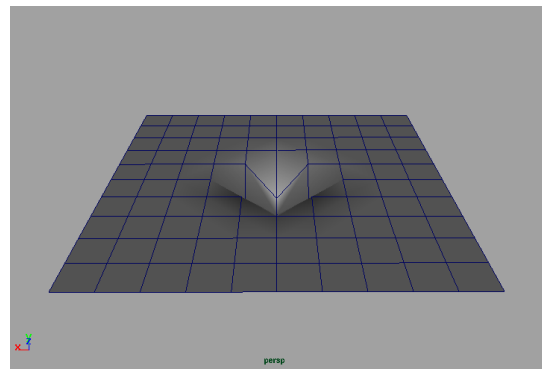
(c) Both blendshapes are applied together, creating superposition interference.



(d) Desired surface.



(e) The vector difference.



(f) Vector difference inverted.

Figure 5. Blendshape superposition.



when the overlapping blendshapes are activated simultaneously. The corrective blendshapes are always actuated in proportion to the strength of the overlapping muscle actions. This gives the appearance a seamless and fluid superposition of muscles.

We must also account for combinations of more than two muscle blendshapes. To correct three or more muscle actions in superposition using our technique, we first apply corrective blendshapes to each pair of superimposed blendshapes participating in the combination. Once pair-wise corrections are in place, we can apply three overlapping blendshapes together.

We may notice new surface interference because the specific combination of three blendshapes has not been corrected yet. We then model a corrective blendshape for this new combination by sculpting our desired target shape and extracting the difference between the corrected geometry and the surface interference.

The expression used for the three-way combination is activated in proportion to the strength of the participating two-way corrections. This expression causes the three-way correction to be applied only when a combination of three muscle blendshapes are applied. This hierarchical method can be extended for higher combinations of muscle action blendshapes.

#### **4.6 Demonstration of the Developed Model With Several Facial Animations**

Animations demonstrating blendshape superposition were created to mimic selected clips of video reference footage. The test animation was intended to show a fairly close imitation of the video reference footage. Muscle action blendshapes were

key-framed to match video footage. The animation was intended demonstrate in detail the effectiveness of our method of correcting blendshapes in superposition.

## **CHAPTER V**

### **IMPLEMENTATION**

To create realistic computer facial animation, we based our model on an actual human reference subject. The person needed to be able to perform a range of muscle actions while being scanned with the Cyberware 3D scanner and also while being recorded with a digital video camera. We decided to use a male subject because we wanted someone with harder features whose expressions would be extreme enough to adequately demonstrate the advantages of blendshape superposition. A woman's face would likely be smoother and surface interference from overlapping blendshapes might not be as noticeable. David Morris, a student in the Visualization Laboratory at the time, volunteered as a reference subject. IRB Human Subjects approval was obtained for this protocol.

#### **5.1 Implemented Muscle Actions**

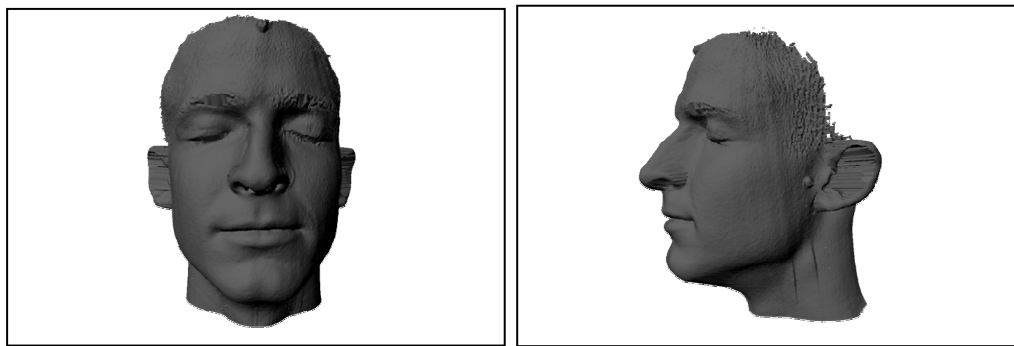
The next step in implementing the model was to determine an appropriate set of muscle actions to capture. We wanted a subset of the FACS muscle actions that would be appropriate for demonstrating the principle of superposition with muscle blendshapes. We selected 12 muscle actions. Most of these were modeled on both sides of the face. The selected muscle actions included:

- Inner Brow Raiser (both sides)
- Outer Brow Raiser (both sides)
- Brow Lowerer (both sides)
- Eyes Closed (both sides)
- Nose Wrinkler

- Upper Lip Raiser (both sides)
- Lip Corner Puller (both sides)
- Lip Stretcher (both sides)
- Lip Corner Depressor (both sides)
- Chin Raiser
- Lip Tightener
- Lip Puckerer

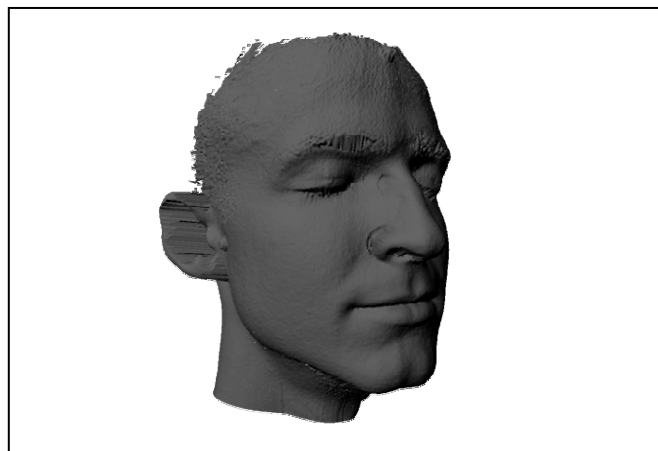
## 5.2 3D Facial Scans

Once the muscle action set was determined, the task was to gather 3D scans of our reference subject. First we scanned David in a neutral pose (see Fig. 6). This is the pose



(a) Frontal view of a neutral 3D scan pose.

(b) Side view of the neutral pose.



(c) Three-quarters view of the neutral scan pose.

Figure 6. 3D scan data.

we would use as reference to create our animatable polygon model (See Fig. 8).

After scanning the neutral pose, we scanned a group of expressions that captured our set of muscle actions. Some expressions displayed several muscle actions at once. These scans, along with video reference data, served as the basis for modeling the muscle actions of the animatable model. The 3D scans (see Fig. 7) and video reference data were also a guide for correcting surface interference from muscle blendshape superposition.

Markers were used on the forehead, nose, and temple regions so that all scans could be aligned with the neutral pose. Each scan was aligned with the neutral scan because the animatable model would also be aligned with the neutral scan. The neutral scan would serve as the ‘template.’ The polygon model was fashioned to fit this data. The expression scans needed to be properly aligned so they could be used as the bases for muscle action blendshapes which would appear rooted in the neutral pose.

### **5.3 The Animatable Model Polygon Topology**

Example polygons topologies for a human face are given by Parke and Waters [1996], as well as Osipa [2003]. We used these topologies as references when laying out the polygon mesh over the neutral 3D scan data (see Fig. 8). The left half of the face was modeled first. The geometry was duplicated and mirrored to create the right side. The two halves were then attached in the middle to form a complete face (see Figs. 9 - 11). We placed the polygon vertices on the scanned surface to obtain the likeness of our human subject. In areas of high curvature, the mesh is denser than in areas of low curvature. The topology needed to be detailed and flexible enough to conform to every

target expression we intend to use. The topology must allow smooth transitions between poses. Polygon contours encircle the mouth and eye sockets. This aids animation in these regions.

The developed topology was tested to verify that it was flexible enough for all of the expressions. This was done by modeling extreme poses from the scanned data in each muscle region. If regions of the topology could not fully conform to a particular pose, the mesh was modified to accommodate that expression.

The mesh structure must accurately accommodate the muscle action shape as well as the nature of the motion from the neutral pose to the muscle action shape. The testing process was carried out for the inner and outer brow raiser muscle actions, the brow lowerer, the lip corner puller, eyes closed, and the mouth stretch. The mouth stretcher was not ultimately included in the model, but it improved the flexibility of the mesh, none the less. The mesh was modified to accommodate wrinkles, dimples, furrows, etc. The final mesh was conformable to all of our 3D scanned target expressions.

#### **5.4 Video Reference Data**

Video reference data was taken of David Morris in a wide range of expressions, including performance which captured more muscle actions than are in our target set (See Fig. 12). The purpose of this data was to provide further means of verifying the model by observing dynamic aspects of the face. The 3D scans only captured static aspects. The video also served as a reference for muscle actions which were not explicitly captured in our 3D scans. This video also later served as a means of

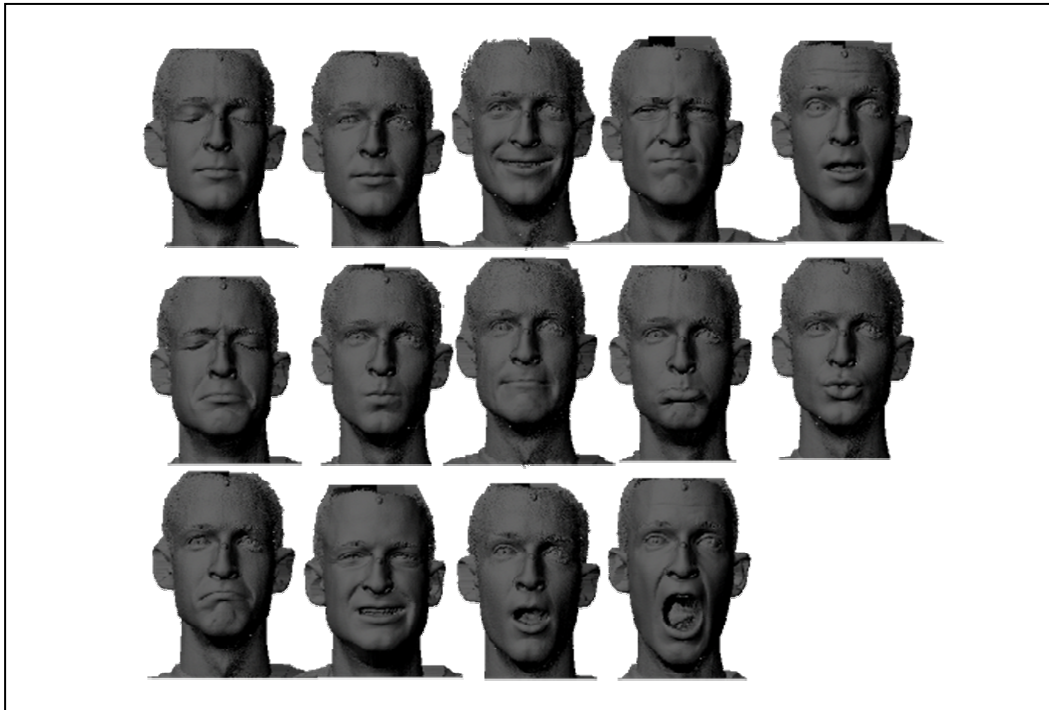


Figure 7. The complete set of 3D scan data.

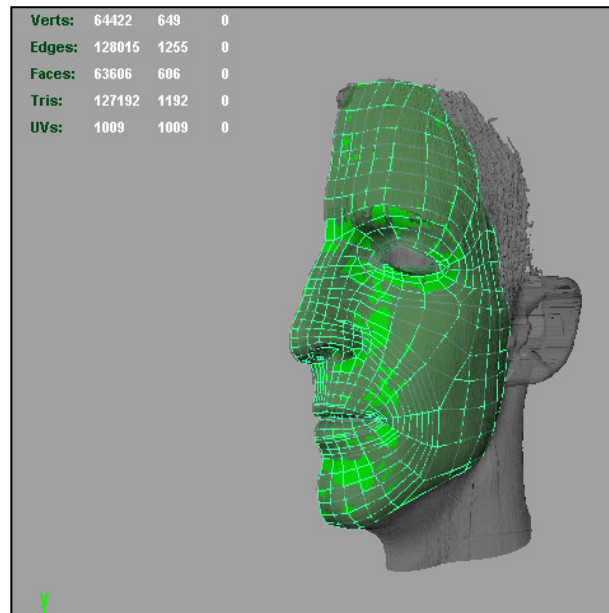


Figure 8. First the left half of the face was modeled, using the neutral scan as a 'template.'

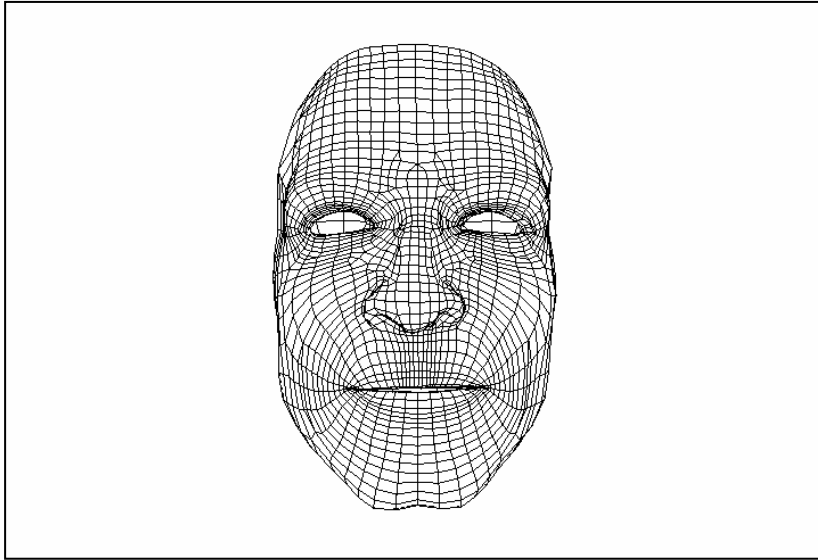
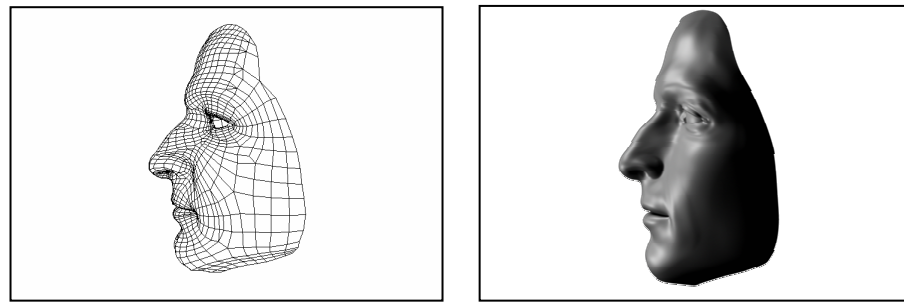


Figure 9. Frontal view of the animatable face topology.



Figure 10. Frontal view of the facial topology rendered as subdivision surfaces with a Blinn shader applied.





(a) Facial topology.

(b) Shaded view.

Figure 11. Side view of face.

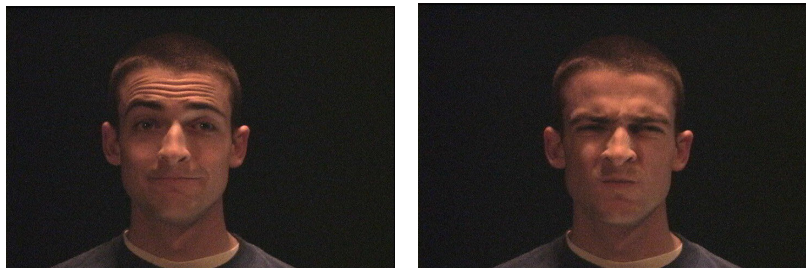


Figure 12. Video reference data of our human reference subject, David Morris.

comparison with generated animated clips. The footage served as a guide for determining which muscle action combinations in superposition we would support. Superposition is used with combinations which include every muscle action in our set.

### 5.5 Muscle Action Blendshapes

After the facial polygon topology was finalized, we modeled individual muscle action blendshapes (see Fig. 13 and Fig. 14). The blendshapes were based on both the 3D scan data and the video reference data.

Individual muscle actions also slightly influence the surrounding surface of the face. We attempted to capture even subtle variations in the surface shape to make muscle

blendshapes appear as realistic as possible. Developing each blendshape involved using a combination of reference from anatomical data, artistic reference, 3D scan data, and video reference data.

First, four brow raiser muscle action blendshapes were modeled individually. Then we began exploring the application of superposition to allow combinations of these muscle actions. Corrective blendshapes were developed for the inner and outer brow raiser muscle actions.

### **5.6 Corrective Blendshapes**

We followed the methodology, previously described, of modeling corrective blendshapes for regions of superposition. Regions of superposition result when muscle blendshapes such as the Left Inner Brow Raiser and the Right Inner Brow Raiser overlap. This occurs when both blendshapes are activated simultaneously.

First we determined which combinations of muscle actions would require corrective blendshapes. It was not necessary to model corrective blendshapes for every muscle action combination. We choose a set of muscle action combinations with superposition interference that needed correction. For the frontalis region, we selected the following combinations.

- Left Inner Brow Raiser, Right Inner Brow Raiser
- Left Inner Brow Raiser, Left Outer Brow Raiser
- Left Inner Brow Raiser, Right Inner Brow Raiser, Right Outer Brow Raiser
- Right Inner Brow Raiser, Left Inner Brow Raiser, Left Outer Brow Raiser
- Left Inner Brow Raiser, Left Outer Brow Raiser, Right Inner Brow Raiser, Right Outer Brow Raiser

In our model, we used a single corrective blendshape for every pair of superpositioned muscle action blendshapes, including the brow raisers. This technique

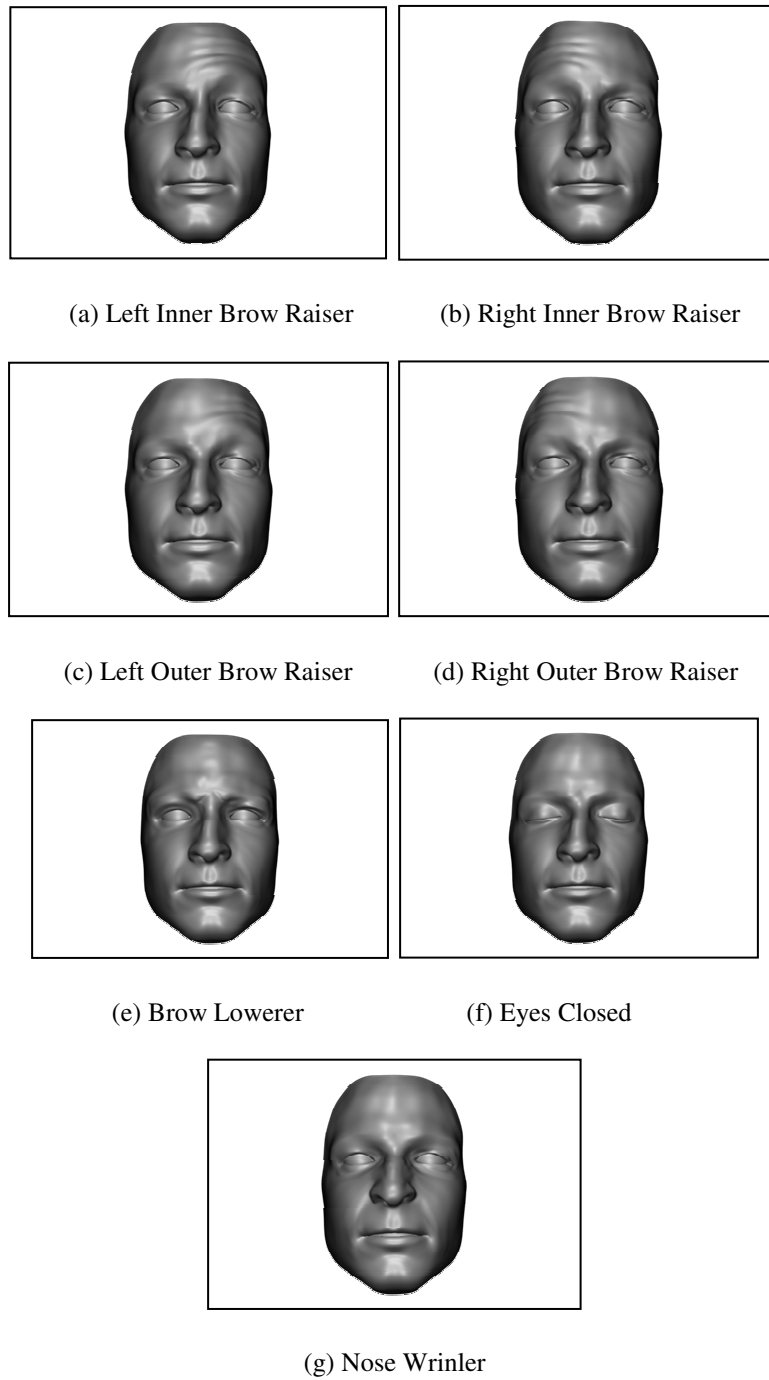
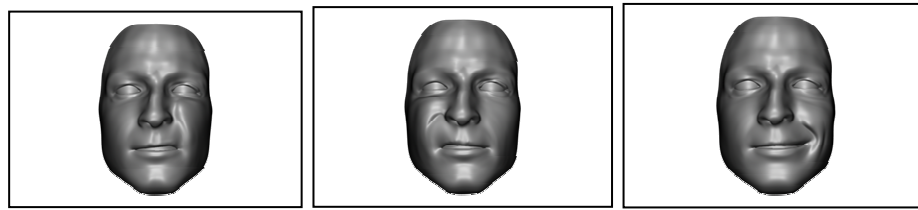
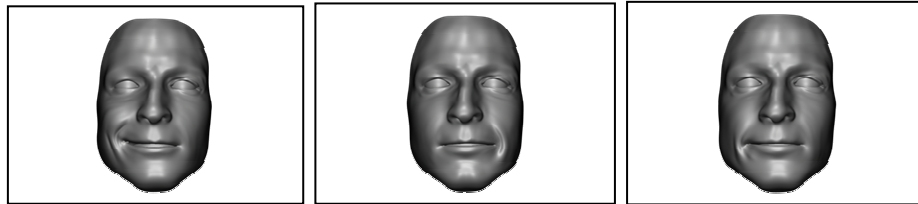


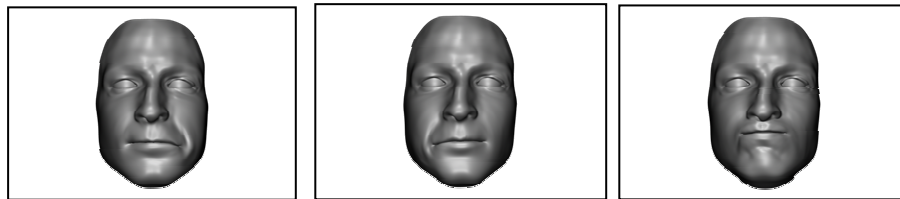
Figure 13. Upper facial muscle actions.



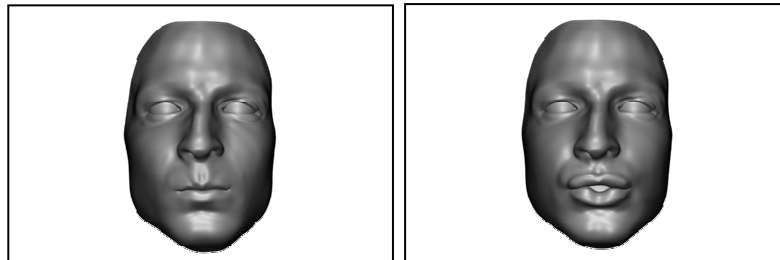
(a) Left Upper Lip Raiser (b) Right Upper Lip Raiser (c) Left Lip Corner Puller



(d) Right Lip Corner Puller (e) Left Lip Stretcher (f) Right Lip Stretcher



(g) Left Lip Corner Depressor (h) Right Lip Corner Depressor (i) Chin Raiser



(j) Lip Tightener

(k) Lip Puckerer

Figure 14. Lower facial muscle actions.

will be explained first. We then explain how combinations using more than two muscle action blendshapes are handled.

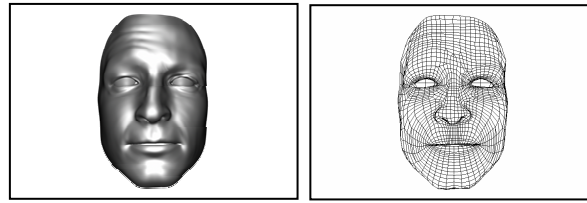
In our explanation of corrective blendshapes for superpositioned pairs, we cite as

an example, the combination: Left Inner Brow Raiser, Right Inner Brow Raiser (see Fig. 15). In this example we refer to the Left Inner Brow Raiser as *IBR\_lft* and the Right Inner Brow Raiser as *IBR\_rt*. When *IBR\_lft* and *IBR\_rt* are applied simultaneously, the principle of superposition states that the resultant shape is the vector sum of the two blendshapes. Because these two blendshapes share vertices in the middle of the forehead, this region exhibits the effects of both blendshapes. We see surface interference in the form of extreme wrinkling and bunching where the muscle actions meet (see Fig. 15(c)).

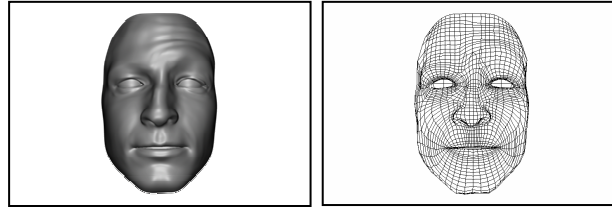
To correct this surface interference, we apply another level of superposition to move the distorted vertices back to their proper positions. We would like to see a shape that looks natural when both muscle blendshapes are applied.

To do this, we made two copies of the geometry with *IBR\_lft* and *IBR\_rt* applied simultaneously (see Fig. 15(c)). With one of the copies, we adjusted the geometry to look realistic based on 3D scan and video reference data (see Fig. 15(d)). This adjusted for the surface interference (in some cases we already had facial geometry modeled for the superposition region, created during our topology flexibility tests.).

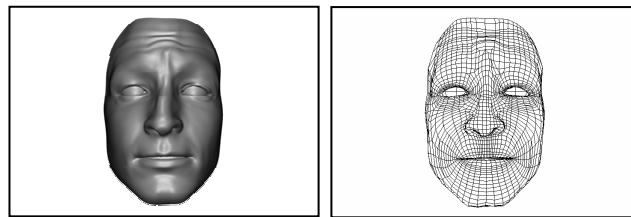
We apply as blendshapes both the uncorrected geometry and the corrected geometry to a copy of the neutral expression (see Fig. 9). We move the coefficient value for the uncorrected blendshape to a value of -1.0, and the coefficient value for the corrected blendshape to value of 1.0. This new combination of blendshapes yields the vector difference between the geometry with surface interference and the corrected geometry, which we see in Fig. 15(e). This forms our corrective blendshape for *IBR\_lft*



(a) Left Inner Brow Raiser.



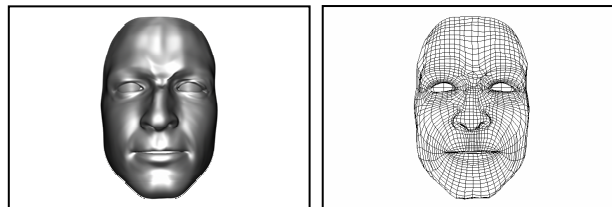
(b) Right Inner Brow Raiser.



(c) Interference when muscle actions (a) and (b) are applied simultaneously.



(d) The inner brow region adjusted to correct surface interference.



(e) Corrective blendshape containing the vector difference between (c) and (d).

Figure 15. Applying a corrective blendshape for a pair of muscle action blendshapes: left inner brow raiser, right inner brow raiser.

and *IBR\_rt*. This difference information is retained in a copy of the facial geometry to be used as a blendshape.

When we apply the difference blendshape (see Fig. 15.e) to the neutral geometry along with *IBR\_lft* and *IBR\_rt*, the former corrects the surface interference and we see the corrected geometry (Fig. 15.d). This is our desired target shape.

The next step allows for automatic animation of the corrective blendshape when *IBR\_lft* and *IBR\_rt* are applied together. Using Maya's expression editor, we apply the following expression to the corrective blendshape.

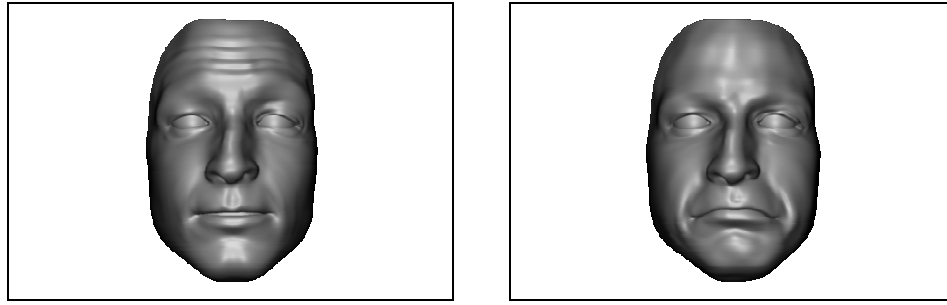
$$\{\text{corrective blendshape coefficient}\} = \\ \{\text{IBR\_lft coefficient}\} * \{\text{IBR\_rt coefficient}\};$$

Applying this expression to the corrective blendshape is one of the most important steps because the correction is now controlled automatically. This blendshape is now always actuated in proportion to the level of superposition between *IBR\_lft* and *IBR\_rt*. Superposition only occurs when both muscle actions are applied. The automatic expression ensures a smooth and seamless correction of surface interference so that we always see a realistic superposition of muscle actions.

### **5.7 Three or More Muscles in Superposition**

If three or more muscle actions are in superposition, we first apply appropriate corrections for each pair of blendshapes in the combination (see Fig. 16 and Fig. 17).

Once pair-wise corrections are in place, we can activate three or more overlapping



(a) Outer Brow Raiser Left, Inner Brow Raiser Left, and Inner Brow Raiser Right.

(b) Chin Raiser, Lip Corner Depressor Right, and Lip Corner Depressor Left.

Figure 16. Three muscle action blendshapes in superposition.



Figure 17. Four muscle actions in superposition: outer brow raiser left, inner brow raiser left, inner brow raiser right, and outer brow raiser right.

blendshapes together. This will yield new surface interference even with the pair-wise corrections activated. Therefore, we must model another corrective blendshape using the technique described above:

Make two copies of the geometry with surface interference, adjust one copy to look realistic, apply the corrected geometry and the geometry with surface interference as



blendshapes to a neutral mesh to yield the vector difference between the two. Then apply the difference geometry as a corrective blendshape.

The expression used for the three-way combination correction will be slightly different than before. This time we activate the three-way correction in proportion to the coefficients of the two-way corrective blendshapes.

$$\{\text{three-way correction coefficient}\} = \\ \{\text{two-way correction A coef.}\} * \{\text{two-way correction B coef.}\};$$

This expression follows the same format as a two-blendshape correction expression, except that by using the correction coefficients instead of individual muscle action coefficients, it is activated only when more than two muscle actions are applied. *{two-way correction A coef.}* handles two blendshapes and *{two-way correction B coef.}* handles a different pair-wise combination. But, since we are dealing with three superpositioned blendshapes in this example, we know that one of the muscles actions must be in both two-way expressions. This is most likely the muscle action located between the other two. Since both two-way coefficients must have non-zero values to activate the three-way expression, we know that the three-way correction is applied only when all three muscle actions are applied in some combination.

### **5.8 Hierarchy of Corrective Blendshapes**

This technique of a hierarchy of corrective blendshapes can be extended to include combinations of more blendshapes. Figure 17 shows four muscle actions in

superposition. Since the complexity of this method can grow quickly with the number of muscle action combinations, it is more practical to choose a specific set of muscle action combinations to support.

### **5.9 Corrective Blendshapes for the Other Facial Muscles**

Once the technique of corrective blendshapes was fully developed for the brow raiser muscle actions, we modeled blendshapes for the rest of our muscle action set for the face.

We applied the corrective blendshape technique described above to other regions of superposition. We chose specific combinations of muscle actions from our muscle action set listed earlier. These regions included the nose, cheeks, mouth, and chin.

In cases where muscle actions were used for which we could not easily refer to 3D scan data, we used video reference data instead. The video reference held additional muscle actions and combinations which normally arise only during active facial movement.

In Maya, we applied selected video frames as texture maps to a planar set of polygons. In a frontal orthographic view, this texture plane was placed in front of the face, scaled to match the size of the face mesh, and aligned with the face model. This approach allowed careful adjustment of the x and y vertex positions. Artistic judgment was used to adjust the z depth value.

In Maya, we are able to adjust the opacity of the texture. When the opacity is lowered to approximately 50 %, the face model could be seen through the video frame. Inconsistencies between the geometry for a particular muscle action and a reference

frame now could be easily noted. Staying in the frontal view, we could manipulate the facial geometry to closely match the video frame. Wrinkles were realigned, corners of the mouth were moved, and various adjustments of a similar nature were made. These video textures served as guides when modeling muscle actions and muscle action combinations.

### **5.10 Supported Muscle Action Combinations**

The following list includes all supported muscle action.

- Left Inner Brow Raiser, Right Inner Brow Raiser
- Left Inner Brow Raiser, Left Outer Brow Raiser
- Left Inner Brow Raiser, Right Inner Brow Raiser, Right Outer Brow Raiser
- Right Inner Brow Raiser, Left Inner Brow Raiser, Left Outer Brow Raiser
- Left Inner Brow Raiser, Left Outer Brow Raiser, Right Inner Brow Raiser, Right Outer Brow Raiser
- Brow Lower Left, Brow Lowerer Right
- Brow Lower Left, Brow Lowerer Right, Nose Wrinkler
- Left Upper Lip Raiser, Right Upper Lip Raiser, Nose Wrinkler
- Left Lip Corner Puller, Right Lip Corner Puller
- Left Lip Corner Puller, Right Lip Corner Puller, Left Upper Lip Raiser, Right Upper Lip Raiser, Nose Wrinkler
- Left Lip Stretcher, Right Lip Stretcher, Left Lip Corner Puller, Right Lip Corner Puller
- Left Lip Corner Depressor, Right Lip Corner Depressor
- Left Lip Corner Depressor, Right Lip Corner Depressor, Chin Raiser
- Left Lip Stretcher, Right Lip Stretcher, Left Lip Corner Depressor, Right Lip Corner Depressor, Chin Raiser
- Chin Raiser, Lip Tightener
- Chin Raiser, Lip Tightener, Left Upper Lip Raiser, Right Upper Lip Raiser, Nose Wrinkler

### **5.11 Test Animation**

Since there are infinitely many levels of activation of various muscle actions in a face, it becomes necessary to pick a specific set of expression sequences to test our

model. We selected video reference footage to mimic as a means of testing the model (see Fig. 18). Our goal was to imitate selected clips of video reference data as closely as possible using the set of muscle actions in our model. The selected test sequences show most of the supported muscle action combinations.

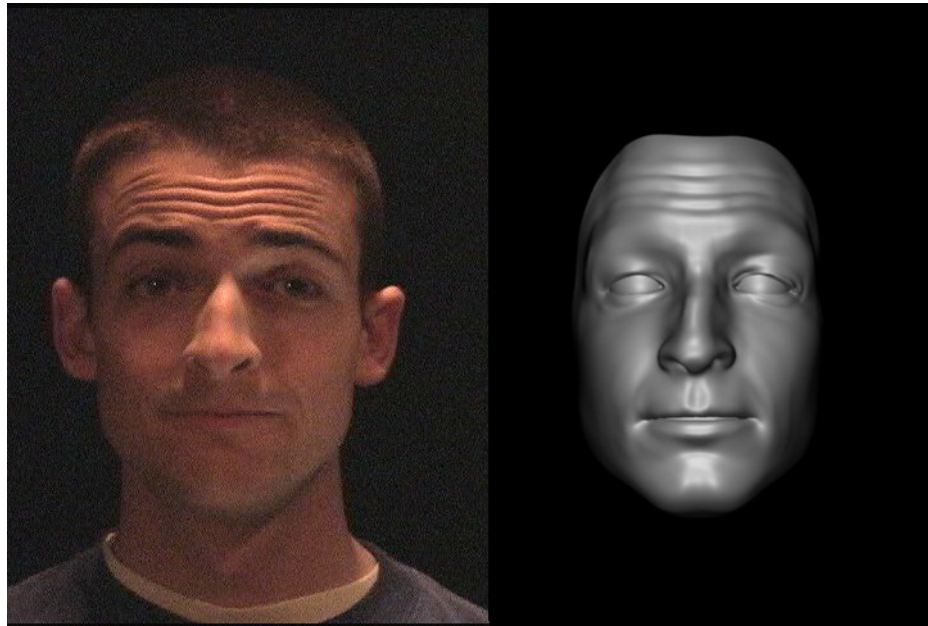


Figure 18. Animation was patterned after video reference data.

We adjusted muscle blendshape coefficients over time to imitate the muscle actions observed in the video reference data. Individual muscle action coefficients were key-framed to match the video reference. Blendshape coefficient interpolation curves were adjusted using Maya's graph editor to capture subtle muscle transitions as closely as possible. All corrective blendshapes were activated automatically using expressions. The appendix consists of a Quicktime file with the resulting test animation.

## **CHAPTER VI**

### **EVALUATION AND SUMMARY**

We have succeeded in developing a model for creating fairly lifelike human facial animation. The model is implemented in Maya as we intended. We were successful in using a muscle action taxonomy similar to the Facial Action Coding System to control our blendshapes.

The main achievement of this project lies in the successful use of muscle action blendshapes in superposition. In regions of superposition, we found a way to extract the vector difference between a surface with interference and our intended target shape. By connecting a blendshape holding this difference information to the muscle blendshapes in superposition using an expression in Maya, we are able to automatically remove surface interference. This technique makes it possible to create more realistic facial animation than with conventional use of blendshapes.

Since the focus of the model is on the shape of the face, we applied a basic surface shader with simple lighting to clearly show the contours of the face. This has worked well for presenting the facial expressions.

Some areas which need improvement are addressed in the Future Work chapter. A principal bottleneck in our process is the time required to model muscle blendshapes by hand. While we benefit from an artist's touch, our principal focus is to match 3D scan and video reference data objectively. It is a somewhat mechanical process to click and drag vertices on the facial model to positions near the surface of a 3D scan. In the case of

2D reference, it is also somewhat mechanical to adjust facial geometry to match a video texture which is aligned with the facial model. We were able to use Maya's polygon sculpting tool to speed up this process. We have come quite close to the likeness of our reference subject, however there are limits on the level of realism we achieved because of various factors: time constraints, artist skill level, imperfect reference data, etc. Some wrinkles did not go deep enough or there were not enough of them. Some blendshapes did not move as naturally as others. These are examples of some of the subtle but noticeable shortcomings observed in our technique. This is not to say that our model falls short of achieving its goal. There is simply room for improvement. It would be beneficial to not need the service of a trained modeler and animator in this process.

The animations we created are a fairly good match to the clips of video reference we attempted to mimic. There were minor differences in the performance due to the fact that our muscle blendshapes were not modeled directly from the video. These differences do not, however, distract from the effective demonstration of muscle blendshape superposition. The animation looks quite realistic, but could still be improved. We would like to capture even more subtle muscle actions specific to each performance.

## **CHAPTER VII**

### **FUTURE WORK**

It would be beneficial to extend the method to address some of the issues mentioned in the Evaluation chapter. Improvements to the facial model could allow yet more realistic facial animation, faster generation of muscle blendshapes, more intuitive control for animators, and applicability to animation of other parts of the body.

The following is a list of possible future directions.

1. Develop an algorithm to automatically find a best fit for superpositioned blendshapes to closely match 3D scanned expressions. A learning algorithm might iteratively readjust blendshapes to closely match 3D scans based on the differences between current blendshape targets and scanned expressions. Coupling the learning algorithm with computer vision might allow blendshape adjustments to be automatically made based on video reference data.
2. Make the facial model more realistic by:
  - Including jaw rotation
  - Adding eye tracking
  - Supporting speech postures
  - Using more realistic lighting and shading
  - Including displacement maps for wrinkles
  - Adding hair

3. Develop a set of graphical control handles for the face which use set-driven keys to drive muscle blendshape coefficient values. This kind of interface might be more intuitive for animators.
4. Extend the subset of FACS muscle actions included in the facial model.
5. Develop an approach for automatically transferring a muscle blendshape set from one facial model to another. If we begin by remodeling an animatable face which has had a full set of superpositioned blendshapes applied to it to the shape of another actor's face, it may be possible to scale each blendshape in the process to roughly match the new face. This might speed up the process of modeling muscle blendshapes for multiple actors. It would be necessary to use a facial topology capable of capturing all of the target expressions of multiple actors.
6. Incorporate dynamic effects into the facial model. It would be useful to be able to layer various facial animation techniques on a single geometry. For example, we could make use of a volume-preserving muscle rig, a muscle rig which deals with skin shifting, and a model with springy mesh properties, in addition to our muscle blendshape model. We could then develop an approach that allows us to layer the effects of each technique onto a single copy of the facial geometry with varying intensities, analogous to the way Adobe Photoshop handles 2D layers. One feature which would allow high-fidelity adjustments of the geometry would be 'per-vertex' controls of the various layers and their interaction.
7. Apply this approach to a full-body animation by using muscle action superposition in other regions of the body.



## REFERENCES

- DUCHENNE, G. B. 1862. *The Mechanism of Human Facial Expression*. Jules Renard.
- EKMAN, P. 1973. *Darwin and Facial Expressions*, Academic Press.
- EKMAN, P., AND FRIESEN, W.V. 1978. *Manual for the Facial Action Coding System*. Consulting Psychologists Press, Inc.
- FAIGIN, G. 1990. *The Artist's Complete Guide to Facial Expressions*. Watson-Guptill.
- FRIED, L.A. 1976. *Anatomy of the Head, Neck, Face, and Jaws*. Lea and Febinger.
- HABER, J., TERZOPOULOS, D. editors. 2004. Course Notes 8: Facial Modeling and Animation. from *ACM SIGGRAPH Conference Presentations DVD-ROM Set*.
- HJORTSJO, C. H. 1969. *Man's Face and the Mimic Language*, Studentlitteratur, Lund, Sweden.
- HYNEMAN, W., ITOKAZU, H., WILLIAMS, L., ZHAO, X. 2005. Human Face Project. From Course 9: Digital Face Cloning. from *ACM SIGGRAPH Conference Presentations DVD-ROM Set*.
- ITOKAZU, H. 2005. Private communication.
- JOSHI, P., TIEN, W. C., DESBURN, M., AND PIGHIN, F. 2003. Learning Controls for Blend Shape Based Realistic Facial Animation. In D. Breen and M. Lin, editors, *ACM Transactions on Graphics*, 37, 187-192.
- MARSCHNER, S. R., GUENTER, B., AND RAGHUPATHY, S. 1999. Modeling and Rendering for Realistic Facial Animation. In Bernard Peroche and Holly Rushmeier, editors, *Rendering Techniques 2000 (Proceedings 11<sup>th</sup> Eurographics Workshop on Rendering)*, 19, 231-242.
- NOH, J., AND NEUMANN, U. 1999. *A Survey of Facial Modeling and Animation Techniques*. University of Southern California, Los Angeles, CA, USC Technical Report, 99-705.
- OSIPA, J. 2003. *Stop Staring: Facial Modeling and Animation Done Right*. Sybex Inc.
- PARKE, F. I. 1972. Computer generated animation of faces. *Proceedings of ACM Nat'l Conference*, 1:451-457.

- PARKE, F. I. 1982. Parameterized Models for Facial Animation. *IEEE Computer Graphics and Applications*, 2(9):61-68
- PARKE, F. I., AND WATERS, K. 1996. *Computer Facial Animation*. A K Peters.
- PLATT, S. M., AND BADLER, N. I. 1981. Animating Facial Expressions. In *Computer Graphics (SIGGRAPH '81 Conference Proceedings)*, 15, ACM, 245 – 252.
- WATERS, K. A. 1987. Muscle Model for Animating Three-Dimensional Facial Expression. In *Computer Graphics (SIGGRAPH '87 Conference Proceedings)*, 21, ACM, 17 – 24.
- WILLIAMS, L. 1990. Performance-Driven Facial Animation. In Forest Baskett, editor, *Computer Graphics (SIGGRAPH '90 Conference Proceedings)*, 24, ACM, 235-242.
- WILLIAMS, L. 2005. Case Study: Gemini Man. From Course 9: Digital Face Cloning, from *ACM SIGGRAPH Conference Presentations DVD-ROM Set*.

## **APPENDIX**

A Quicktime video file containing test animation should be downloaded with this document. This animation compares our facial model using blendshapes in superposition with video reference footage. There is also animation of the model from a side view.

**VITA**

Name Andrew Patrick Smith

Education Texas A&M University  
B.S. in Computer Science  
May 2003

Texas A&M University  
M.S. in Visualization Sciences  
December 2006

Address 928 N. Commerce St.  
Gainesville, TX 76240