Multi-resolution Modeling for Extremely High Resolution 3D Scanned Faces

By: Andrew Thoe Yee Soon, B. Eng.

A thesis submitted to

The Faculty of Graduate Studies and Research

In partial fulfillment of

The requirements for the degree of

Master of Applied Science

Ottawa-Carleton Institute for Electrical and Computer Engineering

Department of Systems and Computer Engineering

Carleton University

Ottawa, Ontario

May 2006

© Copyright 2006, Andrew Thoe Yee Soon

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-16468-6
Our file Notre référence
ISBN: 978-0-494-16468-6

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Abstract

This thesis proposes an approach to model extremely detailed 3D scanned faces in multiple resolutions. The proposed modeling scheme decomposes a scanned face model into several levels of shape and skin detail representations. These representations are structurally consistent, meaning that the structure of the representation at a given level is the same for every scanned face. The shape representations approximate the scanned model to varying degrees, ranging from basic shape capture to the inclusion of fine surface detail. Skin detail representations are derived by calculating the difference between any two levels of shape. The scheme also introduces a "3D skin" representation of the original scanned model, making it possible to reconstruct the original geometry (vertices and their topology) of the scanned model from any shape resolution in the modeling scheme. 3D skin addresses the deficiency of conventionally-used displacement mapping which can only approximate the original geometry.

Three applications of the modeling scheme are presented to demonstrate its benefits. 3D skin transfer, which can be thought of as skin transplantation, shows that the structural consistency of shape representation allows one person's 3D skin to be used with the multi-resolution shape of another person to produce extremely high resolution results. Multi-resolution shape exaggeration and multi-resolution skin detail exaggeration show that shape and skin detail can be manipulated independently from one another and that a higher resolution result can be recomposed by combining shape and skin detail (or 3D skin) components appropriately.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Acknowledgements

I would like to acknowledge the contributions of Dr. Won-Sook Lee to the writing of this thesis in the form of content, encouragement and support. I would also like to thank Won-Sook Lee, Lijia Zhu and Quan Zhou for various contributions to my Master's research.

Numerous other people and bodies deserve recognition for their parts in my Master's studies. I would like to thank Dr. Won-Sook Lee for her supervision, guidance, resources and commitment during my studies. I also express my appreciation to Dr. Samuel A. Ajila for his co-supervision. Carleton University and the University of Ottawa provided me with funding during my graduate studies, for which I am very grateful and appreciative. Special thanks go to my immediate family members for their understanding, support and help. As well, my good friend Abdel-Aziz El-Solh deserves special mention for his optimism, motivation and encouragement during difficult times. Thanks go to Prof. Babak Esfandiari for his role as my program advisor during the first year of my studies. Lastly, but certainly not least, Blazenka Power deserves an honourable mention for her assistance and patience spanning countless written and verbal correspondences.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table of Contents

Abstract iii

Acknowledgements iv

List of Figures ix

List of Tables xiii

Chapter 1 Introduction 1

1.1 A Brief Introduction to 3D Polygonal Surfaces 2

1.2 Motivation 4

1.3 Problem Statement 6

1.4 Proposed Solution 7

1.5 Summary of Empirical Results 8

1.6 Document Organization 9

Chapter 2 State of the Art in 3D Modeling 10

2.1 Face Modeling 10

2.1.1 Sculptures and Plaster Models 11

2.1.2 Digital Sculpting using Arbitrary Photographs as Reference 12

2.1.3 Morphable Model from 3D Face Database 14

2.1.4 Feature Detection with Generic Model Adaptation 16

2.1.4.1 2D Photographs 17

2.1.4.2 3D Range Data 18

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2.1.5 Range Data 18

2.1.5.1 Laser Scanner 18

2.1.5.2 Stripe Generator 19

2.2 Multi-resolution Modeling 20

2.2.1 Displaced Subdivision Surfaces 20

2.2.2 Normal Meshes 21

2.3 Model Parameterization 22

2.3.1 Consistent Mesh Parameterizations 22

2.3.2 High Resolution Model Parameterization 24

2.4 Shape Manipulation 24

2.4.1 Free-Form Deformation 24

2.4.2 Extended Free-Form Deformation 25

2.4.3 Deformation by Radial Basis Functions 26

2.5 Facial Exaggeration 27

2.6 Surface Detail Transfer 28

2.6.1 Normal Mapping 28

2.6.2 Displacement Mapping 29

2.6.3 Image-based Surface Detail Transfer 31

Chapter 3 Multi-resolution Modeling of Scanned Faces32

3.1 Homogeneous Geometric Structure 34

3.2 Feature-based Model 35

3.2.1 Establishment of Feature Point Correspondence 35

3.2.2 Surface Approximation using Radial Basis Function Networks 37

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3.2.2.1 Mathematical Formulation 37

3.2.2.2 Network Training 38

3.2.2.3 Weight Vectors 38

3.2.2.4 Surface Deformation 39

3.3 Representation of Shape and Skin Detail 40

3.3.1 Multi-resolution Shape Construction 43

3.3.2 Shape Refinement 45

3.3.2.1 Vertex Normal Projection 46

3.3.2.2 Cylindrical Projection 47

3.3.3 Multi-resolution Skin Detail Representation 52

3.3.4 Multi-resolution 3D Skin Representation 55

3.4 Database of Scanned Faces 57

Chapter 4 3D Surface Capture and Reconstruction 58

4.1 Geometry Reconstruction using Displacement Mapping 58

4.2 Point-to-Surface Mapping 63

4.3 Model Parameterization 66

4.3.1 Establishment of Mapping Parameters 66

4.3.2 Model Parameterization in Practice 69

4.4 Model Reconstruction 71

4.5 Error Analysis 76

Chapter 5 Applications 81

5.1 3D Skin Transfer 81

5.1.1 Simulation of Aging and Rejuvenation 82

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5.1.2 Augmentation of Mesh Resolution 86

5.2 Multi-resolution Shape Exaggeration 88

5.2.1 Multi-resolution Center Face 89

5.2.2 Exaggeration using Vector-based Scaling 92

5.2.3 Low Resolution Shape Exaggeration 92

5.2.4 Medium and High Resolution Shape Exaggeration 95

5.3 Multi-resolution Detail Exaggeration 96

5.3.1 Multi-resolution Center Detail 96

5.3.2 Exaggeration using Scalar-based Scaling 97

5.3.3 High-Low Skin Detail Exaggeration 99

Chapter 6 Conclusions 101

6.1 Contributions 101

6.2 Discussion 102

6.3 Future Research 105

References 107

Appendix A Multi-resolution Faces 116

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

List of Figures

Figure 2.1: Real-world sculptures used to produce 3D models.. 12

Figure 2.2: Head creation from a template in Sculptor [LeBlanc 91] and TextureFit [Sannier 97]. 14

Figure 2.3: Face morphing [Lee 98; Lee 99] and cloning [Lee 97; Lee W 00]. 15

Figure 2.4: Expressing a face as a linear combination of generic faces [Blanz 99]. 16

Figure 2.5: Example of radial basis function (RBF) deformation. 27

Figure 2.6: Sampling a high resolution surface to create a displacement map. 30

Figure 3.1: Generic model (1,485 triangles). 35

Figure 3.2: Feature point markers placed in the front and side views of a scanned model. 36

Figure 3.3: Construction of a low resolution model from a scanned face. 40

Figure 3.4: Shape and skin detail hierarchy. 42

Figure 3.5: Procedure for creating a shape model in the multi-resolution hierarchy. 43

Figure 3.6: 2D interpretation of shape refinement using point projection. 45

Figure 3.7: 2D interpretation of intersecting rays during vertex normal projection. 46

Figure 3.8: Scanned faces produced during previous work [Soon 06 "Exaggeration"; Soon 06 "shape"] 47

Figure 3.9: Cylindrical projection. 48

Figure 3.10: Shape refinement results. 50

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 3.11: Three feature points (A, B and C) used to calculate the nose projection axis. 50

Figure 3.12: Example of incorrect reconstitution result. 54

Figure 3.13: Database of scanned faces used for experiments. 57

Figure 4.1: Geometry reconstruction result using displacement map sampling. 60

Figure 4.2: Iterative quaternary subdivision. 60

Figure 4.3: Close-up comparison of the left cheek of the original and reconstructed scanned models shown in Figure 4.1. 60

Figure 4.4: Generation and usage of displacement map. 62

Figure 4.5: Example of point-to-surface mapping. 64

Figure 4.6: Point-to-surface mapping using different normals. 65

Figure 4.7: Example of a point for which multiple mappings are possible. 65

Figure 4.8: Bounding box of a 3D triangle ABC and its associated expanded bounding box. 67

Figure 4.9: Reconstruction of a subject model vertex. 71

Figure 4.10: Examples of reconstructing a subject model on the same control model. 73

Figure 4.11: Example of intentionally removing polygons adjacent to an unmapped vertex V. 74

Figure 4.12: Comparison of ways to address unparameterized vertices. 75

Figure 4.13: Sample results from error analysis. 79

Figure 5.1: 3D skin transfer using low resolution shape. 82

Figure 5.2: Example of aging/rejuvenation simulation using 3D skin transfer. 84

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 5.3: 3D skin transfer onto exaggerated low resolution shape..... 85

Figure 5.4: Ten heads obtained from the CAESAR database..... 86

Figure 5.5: Construction of low resolution models from poorly scanned heads. 87

Figure 5.6: Examples of mesh resolution augmentation using 3D skin transfer. 88

Figure 5.7: Formulation of a center face model..... 90

Figure 5.8: Multi-resolution average face models produced from two subject faces..... 90

Figure 5.9: Examples of average face models constructed from varying numbers of faces. 91

Figure 5.10: Example of using low resolution shape exaggeration and 3D skin recovery to produce an exaggerated version of Person B's scanned face. 93

Figure 5.11: Examples of using low resolution shape exaggeration and 3D skin recovery to produce exaggerated versions of Person A's scanned face. 94

Figure 5.12: Example of using medium resolution shape exaggeration and 3D skin recovery to produce an exaggerated version of Person B's scanned face. 95

Figure 5.13: Illustration of skin detail exaggeration..... 98

Figure 5.14: Example of using low resolution shape and exaggerated high-low skin detail to recompose an exaggerated high resolution model of Person A. 100

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 5.15: Example of using low resolution shape and exaggerated high-low skin detail to recompose an exaggerated high resolution model of Person B. 100

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

List of Tables

Table 3.1: Derivation of multi-resolution models..... 44

Table 3.2: Formulations of multi-resolution skin detail. 53

Table 3.3: Formulations of multi-resolution 3D skin. 56

Table 4.1: Statistics for Person B's scanned model..... 78

Table 4.2: Model parameterization settings..... 78

Table 4.3: Error statistics for reconstructing the scanned model of Person B..... 78

Table 4.4: Error statistics for scanned model reconstruction using preliminary medium resolution models. 80

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 1 Introduction

This thesis proposes an approach to model extremely detailed 3D scanned faces in multiple resolutions. The basic shape of the face, which is extracted from the scanned data, is the foundation for the representation. Increasingly more accurate approximations of the scanned face's shape are derived from the basic shape. Several levels of skin detail are established between the various levels of shape. This scheme also provides the necessary elements to reconstruct the original geometry (i.e., the scanned face's vertices and topology) from any shape resolution in the representation. Furthermore, the state (i.e., appearance) of the multi-resolution shape dictates the appearance of the reconstructed model. The reconstruction technique used is able to exactly reproduce the geometry unlike the commonly-used displacement mapping technique which is only capable of reconstructing geometrical approximations.

A multi-resolution modeling scheme for extremely detail faces offers several uses and benefits. In 3D modeling, an object is typically first modeled at a high resolution with all the desired surface detail included. This model is then used to derive lower resolution models which are less detailed. The lower resolution models are necessary for uses such as model manipulation where computational efficiency and visualization of results take precedence over model quality. This is especially true in the case of extremely detailed models, which are unwieldy because of high polygon counts. The proposed representation offers users the flexibility to work with the low resolution models yet still be able to recover the original level of detail. The creation of multiple levels of shape and skin detail representations allows the user to select the level at which

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

changes should be made, i.e., at a low level where only coarse changes can be made or at a high level where fine manipulation is possible. The separation of shape and skin detail also permits a user to perform skin detail operations without affecting shape attributes and vice-versa.

1.1 A Brief Introduction to 3D Polygonal Surfaces

A common way to represent surfaces in 3D space is by using polygons. A polygon is made up of three or more interconnected points in space where each point is considered a vertex of the polygon. The interconnection of vertices forms edges which, in turn, yield polygons. The polygonal representation of a surface can be defined using an indexed vertex list and an unordered/ordered list of polygons, where each polygon is defined by an ordered set of offsets into the indexed list. The most commonly used polygons are triangles and quadrilaterals because they are computationally easier to process than higher order polygons. The *resolution* of a polygonal model increases as the number of polygons increases because a finer level of detail can be modeled with more polygons. The distribution, organization and interconnectivity of the vertices play a large role in defining the *structure* of a model. Two models are said to have the same structure if they have the same number of vertices, the same vertex topology (i.e., interconnectivity) and the same meanings for vertices (i.e., correspondence).

Traditionally speaking, polygons have two faces. However, when used to construct a surface, only one face is visible to indicate that the other face does not contribute to the surface. For example, consider a pyramid constructed from four triangular sidings and a square base. Only the exterior faces of the pyramid contribute to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the representation. The side of a polygon which is visible is referred to as the *frontface* while the other is called the *backface*.

Surface normals are often calculated for lighting (shading) and other operations. The surface normal of a triangle is the 3D vector that is perpendicular to the plane in which the triangle lies. [Note that a quad does not necessarily lie in a plane since its four vertices are not necessarily coplanar.] The surface normal of interest is in the direction of the frontface. The *vertex normal*, which is intended to more accurately represent the normal at a vertex, is derived from surface normals. It is calculated by taking a weighted average of the surface normals of the polygons to which said vertex is adjacent. If each surface normal is given equal weight, the surface is shaded smoothly, i.e., no visible hard edges.

Texture mapping is a popular technique used in computer graphics to add colour and detail to surfaces. Texture mapping is the process of applying an image (*texture map*) onto a surface when it is rendered. It is essentially a 2D parameterization of a 3D mesh that allows colour information stored in a 2D image to be transferred onto the mesh. For instance, a rectangular block can be rendered as a brick by generating an appropriate texture map. However, before texture maps can be prepared, the surface must first be assigned a *UV map*. A UV map describes how a 2D image is wrapped onto a 3D surface. The name "UV" stems from the labels of the axes in the 2D *texture coordinate* system, which is used to specify the location of points on the UV map. In this system, the u axis is horizontal and the v axis is vertical. The values of u and v are restricted to lie in the range [0, 1] with the point (0, 0) located in the bottom left corner of the UV map and the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

point (1, 1) residing in the top right corner. Each vertex of the surface is assigned one or more sets of (u, v) coordinates. The UV map acts as a template for creating texture maps.

1.2 Motivation

Over the years, scanning technology has advanced to the point where it is now possible to acquire extremely high resolution digital 3D models of real-world objects. Such models capture very fine surface detail and are suitable for archival. Unfortunately, the simplicity of using polygons for surface representation is also a detriment in itself. The ability to create accurate representations is directly related to the number of polygons used. In order to accurately model a surface as complex as human skin, a substantially large number of polygons is required. Models with high polygon counts are difficult, if not impossible, to work with directly and are accompanied by significant costs in terms of time and computational resource requirements. A common industry practice is to instead work with a low polygon approximation of the original model to alleviate the computational load. Artists and designers can perform the desired changes on this model and then use texture-based techniques like normal mapping [Blinn 78; Peercy 97] and displacement mapping [Cook 84] to render the altered low resolution surface with the same level of detail as the original model. Films like "The Matrix Reloaded" [Borshukov 03] and video games such as "Doom 3" and "Far Cry" are testaments to the suitability of this practice in delivering the on-screen results needed.

There are, however, some significant drawbacks to this practice.

- Correspondence is generally not achieved between all the low polygon models created. Correspondence means that for a given point (e.g., the tip of the nose of a human face) in one model, the location of the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

corresponding point in any other low polygon model is known. The task of constructing simpler (i.e., lower polygon) models is often accomplished by collapsing edges, minimizing an energy function or removing vertices. These algorithms produce irregular models (i.e., irregular vertex distribution and connectivity) in an indeterministic manner. As a result, there is no relationship between different faces of the same resolution and this outcome precludes any form of automatic geometry processing (e.g., applying the same change to every available face model automatically).

- There is not a great deal of control available to the person making the changes. Typically, only a single low polygon model is produced and its resolution largely dictates the resolution of the changes that propagate to the rendered result. For example, if the low polygon model only captures the very basic shape of the subject's face, a designer or artist cannot alter higher-level detail in the original face (e.g., small creases) because this information is not encoded in the geometry that is altered. The artist would need to construct and alter a higher-resolution model to do that, but it is not a trivial task to amalgamate the changes made to both models to produce a single result and the task becomes increasingly complex as more intermediate models are created for greater control.
- The original scanned model is not physically altered. Many will argue that the only purpose of the scanned model is to geometrically represent a surface at a very high resolution. Once the high resolution detail is extracted from the scanned model, the efficient rendering techniques

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

mentioned previously can reproduce the detail (either in an image or on a visual display) using a compact low resolution model. Hence, the original geometry itself is inconsequential. While this argument is certainly valid for applications such as films or video games in which the rendered result is the sole objective, this is not always the case. For instance, 3D printers generally require geometrical information to be supplied as the source data.

Together, the second and third drawbacks mentioned above present a great challenge. It may be possible to deform the scanned surface at a global level to alter the geometry in a somewhat efficient manner. However, deformations would be very coarse unless a large number of control points are used to manipulate the surface. Increasing the number of control points can greatly increase the complexity of the task and the desired level of control may not even be achievable. Operating on high resolution detail in the geometry would entail working directly with the scanned model, but this is neither practical nor efficient.

1.3 Problem Statement

This thesis addresses the problem of developing a sophisticated multi-resolution modeling scheme for extremely high resolution 3D scanned faces. The requirements for such a scheme are:

- The range of resolutions should be broad enough to capture only low-level shape information at the lowest resolution and a high level of skin detail at the highest. Appropriate levels of shape and detail should be represented in each resolution. Only one face model per shape resolution is required.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- Every face model produced at a particular (shape) resolution is in full correspondence with every other face model of the same resolution.
- Shape manipulation can be performed separately from skin detail manipulation and vice-versa.
- It should be possible to faithfully reconstruct a face at the original scanned resolution using any resolution in the scheme. Any alterations made to the model used to reconstruct the scanned face should be reflected in the reconstruction. The reconstructed face must also resemble real human skin. Localized errors (i.e., surface aberrations) in the reconstruction are tolerable but must be kept to a minimum.
- A user can efficiently manipulate an extremely detailed face by first performing the changes to lower resolution models (derived using the scheme) that are more suitably sized. The altered models are then used to reconstruct the scanned face so that the modifications are propagated to the reconstruction.
- Transitions from one resolution to a higher one are achievable by adding more shape and/or detail information to the former resolution. In other words, one can combine shape and detail elements to form a higher resolution shape. Any changes made to the intermediate elements should be appropriately propagated to the recomposition.

1.4 Proposed Solution

The proposed solution to the problem stated in Section 1.3 is a three-resolution, homogeneously structured representation of shape and skin detail. The representation

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

consists of a low, a medium and a high resolution. The triangle count of the model in each resolution is approximately 3,500, 55,000 and 884,000, respectively. A coarse approximation of the scanned face is derived using shape matching techniques to fit a generic model to the scanned data. A triangular subdivision scheme is iteratively applied to the coarse approximation to obtain preliminary versions of the low, medium and high resolution models. A surface refinement operation is then performed using the preliminary models and the scanned face to arrive at the final model for each resolution. Surface refinement improves the level of approximation so that the refined result more closely resembles the original scanned model. Three levels of skin detail are established: "medium-low" resolution, "high-medium" resolution and "high-low" resolution. The - notation symbolizes a subtraction operation between the two specified resolution models. The concept of "3D skin" is developed to permit faithful reconstruction of the scanned model from any resolution model in the representation. The reconstructed model encompasses the original scanned model's vertex distribution and adjacency information.

1.5 Summary of Empirical Results

The proposed multi-resolution modeling scheme is capable of representing scanned 3D faces with resolutions as high as several million triangles. This representation is used in three different experiments: 3D skin transfer, shape exaggeration and skin detail exaggeration. The following is a summary of the results obtained from the experiments:

- 3D skin representations were computed for four different scanned faces comprising up to three million triangles. Several instances of 3D skin transfer were enacted, which transferred the scanned resolution skin from

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

one person to another. 3D skin transfer can be regarded as a form of skin transplantation where the skin of a donor is applied in a natural way to the shape of the recipient's face. The original level of detail found in the donor's skin was preserved after the transfer was accomplished and only a small number of localized errors appear. 3D skin transfer was used to demonstrate aging/rejuvenation simulation and mesh resolution augmentation (increasing the resolution of low quality scanned faces).

- Shape exaggeration was performed on several faces at two of the three resolutions using exaggeration factors up to 120%. These exaggerations were used to produce exaggerated versions of the original scanned models.
- Skin detail was exaggerated at one resolution independently of the face shape using exaggeration factors up to 120%. This was performed for two individuals' faces.

1.6 Document Organization

This document consists of six sections. Chapter 2 presents a state of the art in 3D modeling and other topics related to the research. The multi-resolution modeling scheme is discussed in depth in Chapter 3. A technique for capturing and reconstructing 3D surfaces is outlined in Chapter 4. Chapter 5 outlines experiments carried out to demonstrate uses for the proposed multi-resolution representation scheme. Empirical results are also presented in this chapter. The thesis concludes with Chapter 6 which summarizes the key inferences made from the research, provides some general commentary about the methods used, discusses the limitations of the proposed scheme and suggests future areas of research.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 2 State of the Art in 3D Modeling

This chapter presents a review of the state of the art in topics which are closely related to the research presented in this thesis. Section 2.1 outlines numerous ways to obtain 3D face models. Section 2.2 discusses a few techniques to model objects in multiple resolutions and is followed by descriptions of model parameterization techniques in Section 2.3. Section 2.4 covers methods for manipulating the shape of 3D models while Section 2.5 addresses the topic of facial exaggeration. Methodologies for transferring surface detail are presented in Section 2.6.

2.1 Face Modeling

Human face modeling is one of the most challenging topics in computer animation. The challenge in producing convincing human faces lies in the ability to meet an audience's expectations of how real faces should look and behave. Therefore, a modeler's task is to construct visually-realistic faces which are capable of delivering realistic facial animations. It is easy to begin to understand why this can be difficult when one simply considers the large variations in facial structure and skin across gender, age and ethnicity. Although the ultimate goal of modeling is to create the most realistic representation possible, factors such as time, cost and suitability often dictate that a certain degree of realism must be sacrificed. However, producing good results is still non-trivial and face modeling (and animation) continues to be a heavily researched topic.

The subsections that follow outline some techniques to construct face models. To date, there are no fully automatic facial modeling methods. Some human interaction is always necessary, whether it be equipment setup, parameter tweaking or data processing.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The user is required to take a trial-and-error approach in some instances such as for setting equipment parameters or for placing markers in interactive programs.

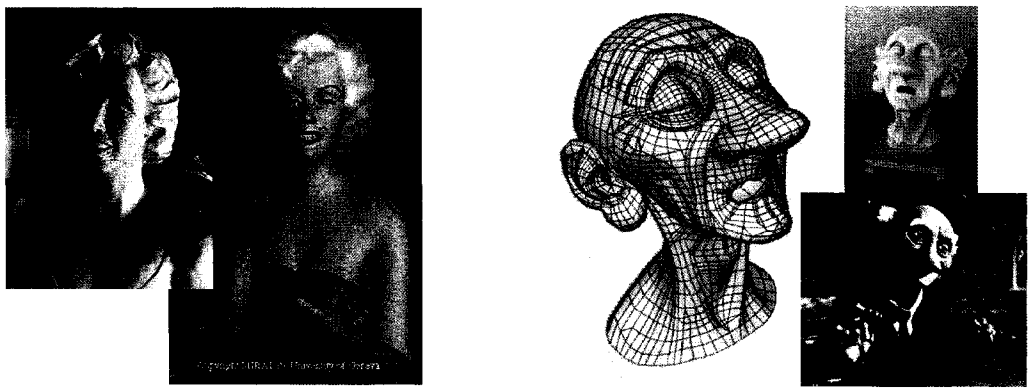
2.1.1 Sculptures and Plaster Models

Even with professional 3D computer modeling tools readily available, many still believe the modeling process begins in the real world with handmade sculptures. This practice demonstrates that many people are still more comfortable producing physical objects with their hands than producing digital models with computer tools. However, one drawback of this approach is that producing good sculptures requires skill. When left to the untrained, results can be unsatisfactory or even mediocre.

Magnenat-Thalmann et al. [Magnenat-Thalmann 87] created face models of Humphrey Bogart and Marilyn Monroe (Figure 2.1(a)) for the 1987 short film "Rendez-vous à Montréal" using digitized photographs of selected facet and vertex markings on plaster models. The 2D coordinate information extracted from the photographs is then used to generate 3D coordinates for the models. The time-consuming task of drawing a mesh on each face must also be performed for this methodology.

Sculptures can also be used to bring fictional characters to life as well. The character Geri in Pixar's animated short "Geri's Game" [DeRose 98] was partially realized as a sculpture before it became a digital model (Figure 2.1(b)). His head and hands were first sculpted and then digitized into coarse polygonal models. The Gollum character in the "Lord of the Rings" movies was also first brought to life in the form of a life-sized plaster model that was then scanned into a computer for digital work [Fordham 03].

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



(a) (b)

Figure 2.1: Real-world sculptures used to produce 3D models. (a) Marilyn Monroe [Magnenat-Thalmann 87]. (b) Geri [DeRose 98].

2.1.2 Digital Sculpting using Arbitrary Photographs as Reference

It is desirable in some instances to create models of people who are not accessible for photographing, filming, etc. For example, Magnenat-Thalmann et al [Magnenat-Thalmann 87] wanted to produce a model of the late actress Marilyn Monroe. The approach of choice in these instances is to work from photographs of the desired subject since photographs are usually available. It is then necessary to calculate 3D information from 2D photographs. The task becomes challenging when there is only one photograph available or several photographs taken at different times and in different environments are used. In the latter case, two options are possible. The first requires laborious human interaction to achieve matching between the many photographs of the subject. The second option is to use pixel comparison calculations that require photographing environment parameters to be manually set. A large database of generic heads is also necessary in the second option.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Just as with hand sculpting, the photograph(s) can also be used as reference for digital sculpting. A software program called Sculptor [LeBlanc 91] uses deformation at both local and global levels to construct 3D objects. The basic functions of Sculptor permit the user to add, remove, modify and assemble triangle meshes. The software is operated through a user interface consisting of a conventional mouse and a specially designed input device called the Spaceball. The Spaceball, which has six degrees of freedom, is used by one hand to handle the sculpted object so that it can be viewed from any angle. The mouse is operated by the other hand to deform and manipulate the sculpture. This configuration closely emulates the experience of traditional sculpting.

Models are texture mapped to improve their appearance at render time. TextureFit [Sannier 97] is a program which lets the user interactively fit a texture map (image) onto a 3D object using a user-friendly interface. The user interactively selects a few features on the 3D object to guide the fitting process and the software determines the closest vertex to each feature. The vertices are projected onto the 2D image to generate texture coordinates for said vertices. The user is able to modify these texture coordinates on the image so that fitting is done correctly. The other vertices in the object are automatically mapped to the image based on the locations of the features on the image.

Figure 2.2 shows a head that was constructed by the MIRALab research group at the University of Geneva using Sculptor and TextureFit. A photograph of a soldier from the Terracotta Army (upper left image in Figure 2.2) served as the reference for the work. The photograph was used to derive a texture map (upper middle image) in Photoshop. The other images in Figure 2.2 depict the 3D head model of the Terracotta figure.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

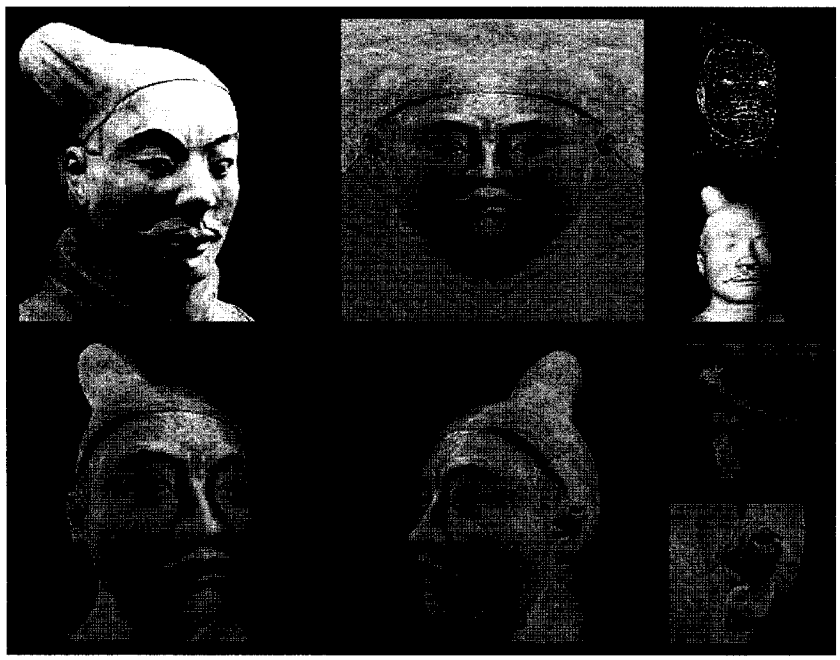


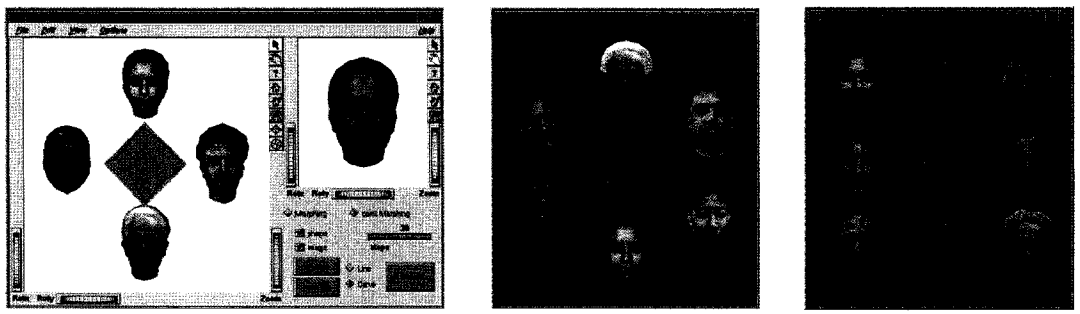
Figure 2.2: Head creation from a template in Sculptor [LeBlanc 91] and TextureFit [Sannier 97]. The upper left image is a photo of a Terracotta Army soldier from which the texture map in the upper middle image was derived. The remaining images show the constructed 3D head model.

2.1.3 Morphable Model from 3D Face Database

Lee and Magnenat-Thalmann [Lee 98; Lee 99] discuss the creation of a population based on morphing techniques. The base models used are created using photo-cloning techniques by Lee and Magnenat-Thalmann [Lee 97; Lee W 00]. Shape and texture parameters are then independently treated to obtain intuitive face modeling using a friendly user-interface. The main methods used are linear-interpolation among the given face data for the shape and image metamorphosis based on triangulation for the texture. The triangulation for the texture is borrowed from the 3D face mesh information. Figure 2.3 illustrates the creation process. The image in (a) shows the user interface for face

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

morphing. Images (b) and (c) show results obtained using cloning and morphing techniques, respectively.



(a) (b) (c)

Figure 2.3: Face morphing [Lee 98; Lee 99] and cloning [Lee 97; Lee W 00]. (a) User interface for face morphing. (b) Face database obtained using photo-cloning techniques. (c) Faces created using face morphing techniques.

Blanz and Vetter [Blanz 99] proposed a method to synthesize natural-looking 3D faces using a morphable model. They compiled a 3D face database consisting of 200 laser scanned young adults with each gender receiving equal representation. The faces all shared a common structure consisting of about 70,000 vertices. Each face is represented by a shape vector $S_i(x, y, z)$ and a texture vector $T_i(r,g,b)$. After establishing full point-to-point correspondence between the faces in the database, novel faces can be expressed as a linear combination of a set of faces in the database as shown in Figure 2.4. To do this, coefficient vectors $a = \{a_1, a_2, \ldots, a_m\}$ and $b = \{b_1, b_2, \ldots, b_m\}$ must be determined to produce the shape and texture vectors $S_{new} = \sum_{i=1}^{m} a_i S_i$ and $T_{new} = \sum_{i=1}^{m} b_i T_i$ for the new face. A new face can be analyzed by supplying either a single photograph or scanned data.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

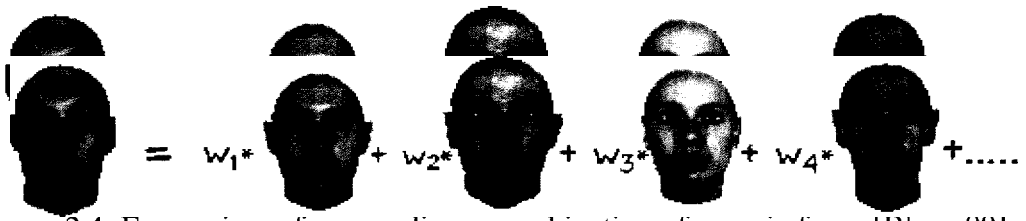


Figure 2.4: Expressing a face as a linear combination of generic faces [Blanz 99].

Rendering parameters such as camera position, object scale, image plane rotation and translation, intensity of ambient light and intensity of directed light must also be determined for each new face to reconstruct an image closely resembling the original. These parameters must be estimated by the user and hence becomes a trial-and-error exercise.

The task of modeling a new face becomes a problem of minimizing an error function E dependent on noise, position, texture and rendering parameters. The optimization of this error metric is iterative and can be time consuming. The authors reported that 10^5 iterations required 50 minutes of computation time on a SGI R10000 processor. A great deal of time is also needed to compile the database of faces (i.e., perform laser scanning), to bring the faces into full correspondence and to estimate parameters.

2.1.4 Feature Detection with Generic Model Adaptation

The premise of feature detection is to detect the location of facial features (eye corners, lip corners, etc.) in input data. These feature points are then used to guide the adaptation of a generic model to approximate the input face. Input data can either be 2D photographs or 3D range data.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2.1.4.1 2D Photographs

A significant drawback of the approach presented in Section 2.1.3 is the time required. It is possible to use more time-saving methods to produce face models from photographs. The photos must, however, be taken or selected carefully so that they meet specific requirements. If an appropriate set of (typically orthogonal) photos for a subject is available, feature detection can be done on them. The 2D facial feature information extracted from the photos is used to calculate 3D feature information to fit a generic model to the face in the given images.

There exist many feature detection algorithms which vary in the level of user interaction required. The method outlined in [Kurihara 91] requires the user to interactively define a small number of control points on input photographs, which are then used to fit a generic model to the input face. The positions of the non-control points are adjusted to fit the face using interpolation. Due to the use of interpolation and a small number of feature points, this method produces models which do not fit the input face very well, which in turn leads to problems during texture fitting. Ip and Yin [Ip 96] proposed a method to automatically extract facial features from photographs using a local maximum curvature tracking algorithm. The two front and side view input photographs are blended using linear interpolation to produce a texture map for the 3D face. Lee and Magnenat-Thalmann [Lee W 00] also used front and side view photographs to detect feature points semi-automatically and the texture map is created using feature matching-based multi-resolution image mosaic. A generic model containing animation structure is then adapted to approximate the input face.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2.1.4.2 3D Range Data

Several works have also demonstrated the principle of adapting a generic model to input data in the 3D domain. Range data obtained using scanning technology such as the Cyberware[1] scanner can produce good quality models of real-world objects relatively quickly. Lee et al. [Lee 93; Lee 95] adapted a generic model to range and reflectance data captured by a scanner. The model is subsequently equipped with tissue, muscle and bone structure to form a physics-based model capable of displaying realistic facial expressions. Zhang et al. [Zhang Y 04] use a combination of global and local adaptation in conjunction with a small set of feature points to match a generic model to input scan data.

2.1.5 Range Data

Range data acquisition is used extensively to digitize static objects for movies, video games and scientific research. Specialized hardware is often required and their cost ranges depending on the technology and level of accuracy needed. Some technologies are more popular than others because they can produce very accurate scans. The proceeding subsections present several technologies that are used to collect range data.

2.1.5.1 Laser Scanner

A laser scanner is an example of a non-contact active 3D scanner. Active scanners emit radiation to digitize an object. In this case, light from a laser emitter is directed at

[1] http://www.cyberware.com

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the subject and the reflected light is used to determine the point of contact. Time-of-flight range scanners determine the point of impact between the laser beam and the object by measuring the time elapsed before the light reflected off the object is detected. Triangulation scanners use knowledge of the distance between the laser emitter and a camera (used to detect the laser dot on the subject), the orientation of the emitter and the orientation of the camera (determined by analyzing the location of the laser dot in the camera's field of view) to fully describe a triangle which, in turn, identifies the location of the laser beam's point of impact. Laser scanners produce point clouds which can be used to create polygonal models.

Cyberware's laser scanning products are widely used to collect range data. Some of their 3D colour digitizers can acquire scans very quickly but at the cost of resolution. Examples of work that has used Cyberware scanning technology are [Lee 93; Lee 95; Guenter 98; Blanz 99]. A more elaborate and sophisticated laser scanning setup is described in [Taylor 02]. This scanning technology captures surface detail at lateral and depth resolutions of 50 and 10 microns, respectively. It has been used to digitize patches on several of Michelangelo's sculptures [Levoy 00; Godin 01] and to produce realistic models of actors' faces for the movie "The Matrix Reloaded" [Borshukov 03]. Scanning time is quite long, but the results are unparalleled.

2.1.5.2 Stripe Generator

Stripe (structured light) generators are also non-contact active scanners that project light onto their subjects. A projector or laser projects an organized pattern (line or grid) of light and a camera captures the changes in the pattern. The images taken by the camera are then processed to determine the shape of the subject(s). If a single stripe is

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

projected, the triangulation principle used for laser scanning can also be applied to recognize the subject. If a grid pattern is used, more complex algorithms such as the one presented in [Zhang 02] are needed to correctly detect the subject.

Proesmans et al. [Proesmans 97] constructed face models by projecting square grid patterns onto a moving face and analyzing the captured video. Texture information is also captured at the same time as the geometry. Zhang et al. [Zhang L 04] constructed a system consisting of four monochrome video cameras, two colour video cameras and two stripe pattern generators to record video sequences of moving faces. A spacetime stereo algorithm computes depth maps from the video sequences. The depth maps are combined with optical flow to produce models suitable for animation.

The strengths of structured light capture are its relatively low cost and its speed. It is affordable because general purpose off-the-shelf equipment can be integrated into capture systems. Speed gains are achieved over other methods like laser scanning when 2D grid patterns are used to capture the subject.

2.2 Multi-resolution Modeling

Multi-resolution modeling is a useful technique in 3D modeling that produces representations of an object in multiple resolutions. Multi-resolution models are suitable for a wide range of applications from animation to efficient real-time rendering.

2.2.1 Displaced Subdivision Surfaces

Lee et al. [Lee A 00] proposed an efficient representation of high resolution geometry called displaced subdivision surfaces (DSS) which essentially represents the high resolution detail as a scalar displacement field over a smooth domain surface. The

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

DSS representation is constructed in three steps: building a control model, optimizing the control model and generating the displacement map. [Displacement mapping is described in Section 2.6.2.] First, a low polygon mesh is produced from the dense geometry. This mesh, which serves as the control model for DSS, is subdivided to obtain a smooth domain surface. The control model is optimized so that the resulting domain surface more closely approximates the original dense model. The displacement map is generated by computing the offsets from the smooth domain surface.

DSS is suitable for mesh editing, animation and other applications which can benefit from using an efficient representation. It is efficient because surface detail is stored as scalar values instead of vector-valued 3D points in a dense mesh. It is also unnecessary to store any dense triangle structure information as it is derived from the control mesh using subdivision. These benefits let a user efficiently work with the control mesh and then recover the surface detail later by re-sampling the displacement map. However, like standalone displacement mapping, the cost of rendering the control model is high compared to other techniques (e.g., normal mapping) because subdivision can increase the number of triangles by very large factors. A limitation of DSS is that it cannot be used to reconstruct the topology of the original high polygon model. DSS is only intended to approximate the original surface.

2.2.2 Normal Meshes

The normal mesh [Guskov 00] of a surface is a multi-resolution representation in which the mesh for each level is defined as a normal offset of a lower level (coarser) mesh. The procedure applies mesh simplification on the surface to build a base domain whose vertices contribute to a patch network. Each patch is parameterized to a triangle

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and this initial parameterization gives way to the normal mesh representation. Each level of the normal mesh is constructed by subdividing the previous level to form new vertices (and, consequently, finer patches) and offsetting the new vertices in the normal direction.

The primary contribution of normal meshes is a way to represent surfaces using only a single value per vertex as opposed to the traditional three per vertex (i.e., a floating point value for each coordinate of 3-space). The multi-resolution nature of normal meshes also makes it suitable for a wide range of applications. However, a normal mesh only approximates the irregularly-connected original geometry with a semi-regular mesh and thus is not suitable for topology recovery. Additionally, because it is approximative, the representation may not be able to represent very finely detailed surfaces (e.g., human skin) even with a very large number of levels.

2.3 Model Parameterization

Model parameterization refers to the act of expressing one surface with respect to another. Two types of parameterization are described in the proceeding subsections: consistent mesh parameterizations and high resolution model parameterization.

2.3.1 Consistent Mesh Parameterizations

Lee and Magnenat-Thalmann [Lee W 00] used mesh adaptation to automatically transfer animation structure onto range data. A unified method is applied on input photographs as well as range data input. Examples of other works that employed mesh adaptation are [Lee 93; Lee 95; Zhang Y 04], which are outlined in Section 2.1.4.2.

Praun et al. [Praun 01] proposed an algorithm to generate parameterizations for a set of meshes. The parameterizations are said to be consistent because they share a

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

common base domain (point connectivity) and respect the features of the domain (e.g., eyes of a human head). Features are defined on each mesh in the set using either feature detection schemes described in the literature or human interaction. The feature vertices that result are used to define patches on the mesh's surface. A network of patch boundaries referred to as a "net" is then established such that the net is topologically equivalent to the base domain's connectivity and that the resulting patches are fair (well distributed, smooth boundaries, absent of artifacts, etc.).The result of performing the algorithm on all the models is a topologically consistent set of nets.

The strength of consistent mesh parameterization is in its ability to establish a common base domain between very different objects and in the use of its results in applications. The authors demonstrated that the same base domain can be used for both a human and a horse. Although this particular scenario has limited use in practice, one might see a more practical use in establishing the same layout for a human body and a fictional creature endowed with both human and animal physical attributes. A byproduct of establishing consistent mesh parameterizations is exact one-to-one point correspondence across an entire set of models. This trait can then be exploited in applications such as animation, texture transfer and principal component analysis.

Lee [Lee 02] focused on human body consistent parameterization with various types of human body data such as photographs, laser scanned meshes, and designer-created meshes. Two steps of mesh adaptation are used to ensure the robust parameterization. The first step uses body feature points and the second uses silhouette information.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2.3.2 High Resolution Model Parameterization

The parameterization of a high resolution model with respect to a low resolution model is mentioned in [Hilton 02]. Parameterization amounts to using point-to-surface mapping [Sun 01; Zhang Y 04] to map the vertices in the high resolution model onto the low resolution surface. In both [Sun 01] and [Zhang Y 04], point-to-surface mapping is used to map the high resolution vertices to the low resolution surface so that a displacement map can be generated. The following is stated in [Hilton 02] with regard to high resolution model parameterization: "Having computed this mapping the model can be exactly reproduced and animated from the low-resolution model, M^L, together with the high-resolution vertex parameters and mesh topology." However, this idea was not tested for that publication. The applicability of this idea is addressed in Chapter 4 alongside a technical discussion of the idea in practice.

2.4 Shape Manipulation

It is often desirable to change the shape of an existing model for animation, simulation and other applications. There are many techniques described in the literature to accomplish this. This section presents a few of these techniques.

2.4.1 Free-Form Deformation

Free-form deformation [Sederberg 86], or FFD, is a technique for deforming objects by manipulating the space in which they are embedded. A FFD lattice structure serves as a space enclosure and is formed by tricubic Bezier hyperpatches. The analogy drawn by the authors is to imagine the lattice structure as a plastic parallelepiped whose

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

shape can be controlled by a set of control points. Deforming the structure (by moving the control points) would then cause the enclosed object to be deformed as well.

One of the strengths of FFD is its applicability for both global and local deformation. Global deformation is accomplished by enclosing the entire object within a single FFD structure while localized deformation requires the use of several interconnected structures to enclose the object. Continuity across lattice boundaries (i.e., adjacent lattices) can also be preserved. It is possible to control or preserve the volume of the object when performing FFD. The limitations and disadvantages of FFD include the inability to control the shape of the boundary between a deformed region and a non-deformed region, the inability to produce certain types of surface deformations (e.g., small bumps) and high computational cost which results from calculating new positions of points.

2.4.2 Extended Free-Form Deformation

Extended free-form deformation (EFFD) was proposed by Coquillart [Coquillart 90] as an extension to FFD. One of the biggest differences between the two is the use of non-parallelepiped lattices in EFFD. Parallelepiped lattices restrict the type of deformations that can be achieved because of its shape. An EFFD lattice can be prismatic or non-prismatic (e.g., spherical) and prismatic lattices can further be classified as either elementary or composite (several elementary prismatic lattices). Contrary to FFD where the object is deformed by manipulating lattice control points, changes are made to the surface by directly reshaping an EFFD lattice. The generalized lattice empowers a modeler with the flexibility to make arbitrary deformations to a surface. Additionally, the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

shape of the deformed region is not constrained by the position of neighboring control points as in FFD.

The use of differently shaped lattices can lead to problems in continuity across lattices. The work states, however, that surface continuity is often preserved even though lattice continuity may not. The results of EFFD are not as natural as those obtained using physical simulation, but it is much more computationally efficient to use EFFD.

2.4.3 Deformation by Radial Basis Functions

Radial basis functions (RBF) [Powell 87] networks are a powerful and efficient mathematical tool for surface approximation. RBF depend on a special set of points which lies on the surface known as centers. A reasonable approximation of a surface can be achieved by carefully selecting a good suitable set of points to represent it.

RBF may also be employed for deformation purposes if the locations of centers are determinable at all times. Many works [Fidaleo 00; Noh 00; Noh 01; Bui 03] have shown its suitability in algorithms for facial animation applications. After surface approximation is established, the internal parameters of the RBF can be updated to reflect any changes to the centers. The RBF can then approximate the effect of the changes to the rest of the surface. This technique allows the surface to be manipulated using only a (typically) small set of points. This concept is exemplified in Figure 2.5. The flat grid in the left image consists of 81 points and the chosen centers are each marked with a red square. The deformed grid on the right demonstrates the effect of translating the center in the middle of the flat grid upward while anchoring the other centers. The new positions of the non-centers produce an overall smooth curved surface.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(a) (b)

Figure 2.5: Example of radial basis function (RBF) deformation. Centers are indicated by red squares. (a) Before. (b) After.

RBF is used to produce models in the multi-resolution modeling scheme and the technical details are described in Section 3.2.2.

2.5 Facial Exaggeration

Brennan laid the groundwork for creating digital caricatures [Brennan 85] by manually defining polyline drawings from face images and then increasing each face drawing's deviation from a calculated mean (average) face. Many 2D (image) and 3D (model) exaggeration methodologies were developed thereafter.

Liang et al. [Liang 02] proposed an example-based approach to generate caricatures of face images. Sample drawings supplied by a caricaturist are needed so that their system can be trained to acquire an understanding of distinctive feature selection as well as the style in which exaggeration is done. After the shape of an input face has been exaggerated by their system, texture style transfer is performed to produce the final caricature. Chiang et al. [Chiang 05] also developed a system to produce caricatures which follow the style of a caricaturist's hand-drawn work. The system extracts features from an image of the subject, identifies the unique features and selects a suitable exaggeration factor. A sample of a caricaturist's work is then morphed to produce the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

stylized individualized caricature. Later work by Mo et al. [Mo 03] improved upon Brennan's original idea by also taking into consideration not only feature difference (from the mean) but also feature variance.

A standard caricaturing algorithm is typically employed in the case of 3D exaggeration. Face data is restructured to conform to a homogeneous (morphable) structure. Point correspondence is established across all faces and an average face can be easily calculated. Differences between each input face and the average face are computed and exaggerated to produce the caricatured face. This approach has been adopted by many researchers [O'Toole 97; Blanz 99; Fujiwara 01]. A technical description of the standard caricaturing algorithm is presented in Section 5.2.2.

2.6 Surface Detail Transfer

There are many examples of interesting and beneficial uses of surface detail transfer. Aging or rejuvenation can be simulated by transferring wrinkled or youthful skin onto different subjects. This concept is used by companies and agencies involved in locating missing persons. Transferring surface detail can also be an efficient way of rendering the same geometry in different ways. For example, a rectangular block could be rendered as a brick, a bar of gold or block of wood simply by transferring the appropriate surface detail onto it during rendering. This reduces the number of models that need to be produced.

2.6.1 Normal Mapping

Normal mapping [Blinn 78; Peercy 97] is a render-time technique used to create the illusion of detail on a surface. The surface's normals are replaced completely by those

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

stored in a normal map prior to performing shading calculations and, by doing so, adds bumps and other surface detail that is absent in the geometrical representation. The normal vectors in the normal map are typically taken from a higher-detail version of the surface to be rendered. The x, y and z components of these vectors are stored in the RGB fields of an image (red for x, green for y, blue for z). The UV mapping assigned to the surface determines which value is retrieved from the image for a particular point on the surface.

Normal mapping has grown in popularity for real-time applications such as video games because it is a relatively inexpensive way of adding a high level of detail to 3D objects. As its popularity grew, manufactures of 3D graphics hardware began implementing normal mapping support into their products. The major disadvantage of normal mapping is that there are no physical changes to the geometry of the rendered surface. As a result, bumps that are added by normal mapping do not cast shadows correctly and can ruin the illusion. Similarly, the silhouette of the 3D object can also reveal the use of normal mapping (e.g., a "bumpy" surface has a smooth silhouette).

2.6.2 Displacement Mapping

Displacement mapping [Cook 84] is a technique used in computer graphics to augment the appearance of a surface. Displacement mapping yields a more detailed surface by perturbing a base surface along its normals. The perturbation values are retrieved from a displacement map that contains either scalar or vector values. A displacement map is essentially a texture map with the offsets stored in the colour channels of the image's pixels. In the case of vector displacements, each component of the vector is assigned to its own channel (e.g., R for the x component, G for the y

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

component and B for the z component). The values are retrieved from the displacement map according to the base surface's UV map. The same principle is used to create the displacement map, i.e., the offset stored in a pixel is obtained by mapping said pixel onto the base surface and calculating the distance between the base and the high resolution surface along the surface normal at the mapped point . A simple example of displacement map computation is shown in Figure 2.6.

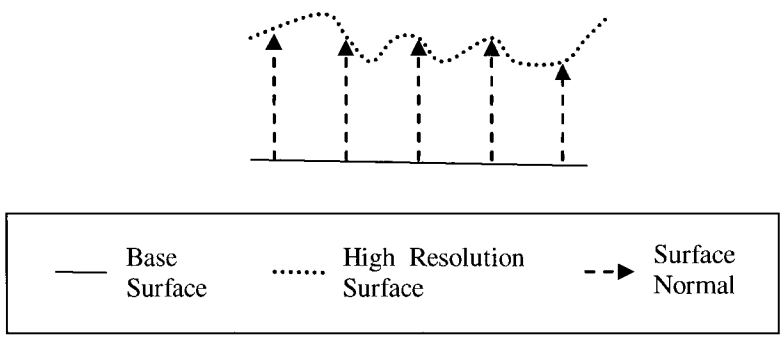


Figure 2.6: Sampling a high resolution surface to create a displacement map.

The benefits of producing actual deformed geometry (e.g., proper shadow casting and occlusion) are offset by the cost of the technique. The low polygon base surface must be subdivided to produce a large number of smaller polygons whose vertices can then be displaced to achieve the desired effect. This process increases the cost of rendering the object significantly and is not a practical choice for real-time rendering. Another disadvantage of displacement mapping is that it is not suitable for reconstructing extremely high resolution geometry. Displacement mapping is approximative by definition and this limitation is magnified when dealing with very high resolution surfaces. An example of this deficiency is presented in Section 4.1.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2.6.3 Image-based Surface Detail Transfer

The technique of image-based surface detail transfer (IBSDT) was proposed by Shan et al. [Shan 04] as a way of transferring geometric detail from one image to another without any knowledge of the first image's 3D geometric information. The fundamental principle upon which this approach operates is that image smoothness corresponds to geometric smoothness when surface reflectance is smooth. Geometric detail is expressed independently from surface reflectance; thus, the recipient of the transferred detail retains its own reflectivity attributes. Given two source images – one acting as the detail donor and the other as the detail recipient – the surfaces depicted in the images are brought into alignment by performing transformations (scaling, translation and rotation) or by performing image warping. The images are then smoothed using a Gaussian filter whose σ value controls the degree of smoothness achieved, which in turn affects the scale of detail that can be transferred. A small value of σ must be used to transfer fine surface detail whereas only coarse detail can be transferred with a large value of σ.

The primary strengths of IBSDT are its simplicity and its requirement of only one image for each participating object. The researchers also demonstrated human skin detail transfer to simulate aging. The limitations of IBSDT are the need for similar lighting conditions in the two input images and the need for smooth reflectance in the objects.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 3 Multi-resolution Modeling of Scanned Faces

Today, there are many companies that offer shape acquisition services commercially. The models which arise from these services are usually structured in an efficient manner with no relationship between structures of different models. With no guarantee that the data yielded by the scanning service is structured consistently, performing operations on the models can be difficult. For example, calculating the average of a set of faces requires correspondence between the elements of the set, i.e., for each face, the vertex corresponding to a particular location (e.g., nose tip) is known. For this reason, it is desirable to include a restructuring step during data preparation to ensure that all scanned faces share the same structure.

This chapter presents a multi-resolution modeling scheme suitable for extremely high resolution scanned faces. The scheme proposed is a three-level hierarchy (i.e., three resolutions) although the choice of the number of resolutions was made arbitrarily and the building block of the hierarchy allows for any number of resolutions. The scheme consists of four main components.

- *Feature-based Model*: The foundation of the hierarchy is a feature-based model which captures the coarse shape of the scanned face.
- *Shape Models*: Three higher order shape models – one for each resolution – are derived from the feature based model. The resolution, accuracy and level of detail of the shape models increase with the order. Each resolution is characterized by its own geometric structure that is derived from a *generic model* and every face model of a particular resolution inherits the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

same structure. This ensures structural homogeneity for each resolution and, in turn, full point correspondence between all faces of a particular resolution.

- *Skin Details*: Skin detail is loosely modeled as the "difference" between the models of two different resolutions; hence, there are three levels of skin detail in a three-resolution scheme. Another way to conceptualize skin detail is that it is "added" to a shape model of resolution X to obtain the shape model (of the same person) of resolution Y, where X is lower than Y. Thus, adding skin detail to a shape model increases its resolution. Structural consistency is also an integral part of skin detail.
- *3D Skins*: 3D skin is a representation of an extremely high resolution scanned face that is also loosely modeled as the "difference" between the scanned face and a shape model. The principle behind it is similar to that behind skin detail. The result of "adding" or "applying" 3D skin onto a shape model is an extremely detail face model which bears the shape of the shape model and the surface detail of the scanned face from which the 3D skin representation is derived. In this regard, 3D skin is similar to skin detail, but there are some prominent differences that distinguish one from the other. These differences are also discussed later in this chapter.

Section 3.1 describes the concept of the homogeneous structure and the generic model from which it is derived. The procedure carried out to construct the feature-based model is outlined in Section 3.2 proceeded by a discussion of the other components of the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

multi-resolution hierarchy (shape, skin detail and 3D skin) in Section 3.3. In Section 3.4, a small database of extremely high resolution 3D faces is showcased.

3.1 Homogeneous Geometric Structure

The goal of homogeneous representation is to use a common geometric structure to represent models of the same resolution. The geometric structure of a model encompasses vertex distribution and topology, i.e., vertex adjacency. The models of the multi-resolution hierarchy are derived from a single *generic model* in a deterministic way. By doing so, structural homogeneity (consistency) is achieved among the models.

The generic model, comprising 1,485 triangles, is a derivative of the one described in [Lee W 00]. The eyes and mouth of the generic model have been closed to match the states of the scanned faces (refer to Section 3.4). The upper entities (eyelid and lip) come into contact with their lower counterparts at piecewise linear 3D lines so that there is no overlap and/or self-penetration. Additionally, the animation structure in the original generic model has been removed, though it can be re-included fairly trivially.

The generic model has efficient triangulation with finer triangles over the highly curved and highly articulated regions of the face and larger triangles elsewhere as shown in Figure 3.1. A subset of the generic model's vertices is classified as feature points which represent the most characteristic points used for human recognition. There are 172 such vertices and each is further categorized as either major or minor. Twenty-eight feature points are major and the remaining 144 are minor. These points are shown in Figure 3.2. Feature points come into play during feature-based model construction. Feature points and feature-based model construction are described in the section that follows.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

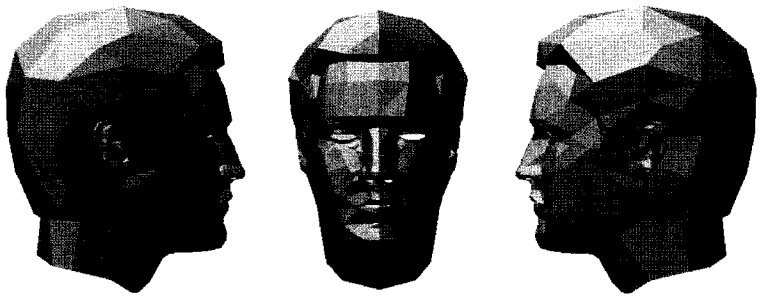


Figure 3.1: Generic model (1,485 triangles).

3.2 Feature-based Model

The first step in establishing the hierarchy is to achieve a coarse approximation of the scanned data. The generic model's shape is manipulated to roughly match that of the scanned face. Higher-order shape representations can then be constructed from this coarse approximation. Feature points guide a deformation algorithm that accomplishes the required rough matching. For this reason, the resulting deformed generic model is referred to as the *feature-based model.*

3.2.1 Establishment of Feature Point Correspondence

Feature point correspondence must be established between the generic model and the scanned data in order to perform deformation. This means that the same set of feature points defined for the generic model must also be defined on the scanned surface as well. A feature point detection procedure [Lee W 00] is performed to establish these points. However, unlike the generic model, scanned model feature points need not necessarily be vertices– it is sufficient for them to merely lie on the scanned surface. Detection requires a user to interactively place feature point markers on the front and side views of the scanned data visualization as illustrated in Figure 3.2. Of the 172 feature points defined

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

for the generic model, only 135 are visible in the front view and 23 of these are major points. Marker placement is considered a semi-automatic procedure because the positions of minor feature points can be automatically calculated (using affine/similarity transformations) based on the manual placement of major feature point markers. The positions of minor feature point markers in the side views are also automatically calculated based on the depth of the scanned model. Feature point marker placement on the front view can be performed automatically if texture information is available [Pötzsch 96; Goto 02]. Once established, the 2D marker positions are used to calculate 3D feature points which lie on the surface of the scanned model.

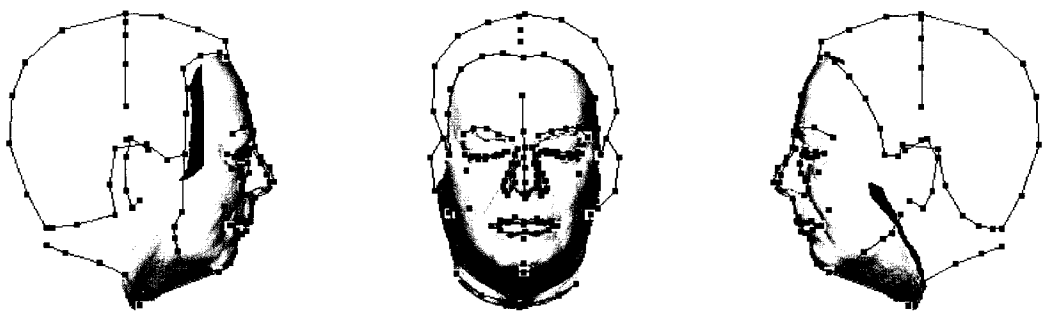


Figure 3.2: Feature point markers placed in the front and side views of a scanned model. Major feature points are marked in both yellow and red while minor points are only marked in red.

Note that since the scanned model only captures the face of the subject, rough estimations have to be made for the remainder of the head during feature detection. As a result, these areas are very coarsely approximated in the feature-based model and they cannot be improved in the shape refinement step (discussed in Section 3.3.2).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3.2.2 Surface Approximation using Radial Basis Function Networks

Section 2.4.3 gives an overview of the concept of deforming surfaces using radial basis functions (RBF) [Powell 87]. Works such as [Fidaleo 00; Noh 00; Noh 01; Bui 03] have reported the success of using RBF networks for face manipulation. The text that follows describes the manner in which these networks are prepared and used to deform the generic model so that the deformed result roughly resembles the scanned face.

3.2.2.1 Mathematical Formulation

For the purpose of deforming the generic model, one RBF network is created for each coordinate of 3-space. The general expression for each RBF network $\mathrm{R}_k(\bar{x})$ is

$$\mathrm{R}_k(\bar{x}) = \sum_{i=1}^{N} \mathbf{W}_{k,i}\phi_i(\bar{x}), \quad k = 1,2,3. \tag{3.1}$$

N represents the number of feature points in the generic model (i.e., $N = 172$) and the set of generic model feature points $\overline{\mathbf{X}} = [\bar{x}_1 \ \cdots \ \bar{x}_N]$ is used as centers by all three networks to compute their respective weight vectors $\mathbf{W}_k$. [The term "center" is defined in Section 2.4.3. Further discussion about the vectors follows the coverage of network training.] The notation $\bar{x}$ denotes a point in 3-space, i.e., the point has x, y and z coordinates. The Hardy multiquadric basis function

$$\phi_i(\bar{x}) = \sqrt{\|\bar{x} - \bar{x}_i\|^2 + s_i^{\,2}} \tag{3.2}$$

is used, where $\|\bar{x} - \bar{x}_i\|$ denotes the Euclidean distance separating the point $\bar{x}$ and the ith center $\bar{x}_i$. The s_i term is a stiffness constant whose value is chosen during the training stage.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3.2.2.2 Network Training

The training stage initializes the stiffness constants to reflect the separation between the generic model's feature points. The networks are trained using the feature point set $\overline{\mathbf{X}}$. The values of the stiffness constants are chosen to be

$$s_i = \min_{j \neq i} \left\| \overline{\mathbf{X}}_j - \overline{x}_i \right\| \tag{3.3}$$

as suggested in [Eck 91], where $\overline{\mathbf{X}}_j$ denotes the jth entry in $\overline{\mathbf{X}}$. These values for the stiffness constants yield more subtle deformations for feature points with greater separation and more pronounced deformations for feature points located more closely to one another.

An N x N matrix $\boldsymbol{\varphi}$ containing the evaluations of the basis functions may be built to aid computation of the weight vectors $\mathbf{W}_k$ and is given by

$$\boldsymbol{\varphi} = \begin{bmatrix} \phi_1(\overline{\mathbf{X}}_1) & \cdots & \phi_N(\overline{\mathbf{X}}_1) \\ \vdots & & \vdots \\ \phi_1(\overline{\mathbf{X}}_N) & \cdots & \phi_N(\overline{\mathbf{X}}_N) \end{bmatrix}. \tag{3.4}$$

3.2.2.3 Weight Vectors

The values in the weight vectors $\mathbf{W}_k$ reflect the current state of the centers. When deforming a surface, the new locations of the centers (i.e., the feature points defined on the scanned face) are used to set the values in the vectors. If the scanned face's feature points are given by $\overline{\mathbf{X}}' = [\overline{x}_1' \quad \cdots \quad \overline{x}_N']$, $\mathbf{W}_k$ is calculated as

$$\mathbf{W}_k = \boldsymbol{\varphi}^{-1} \overline{\mathbf{X}}_k', \tag{3.5}$$

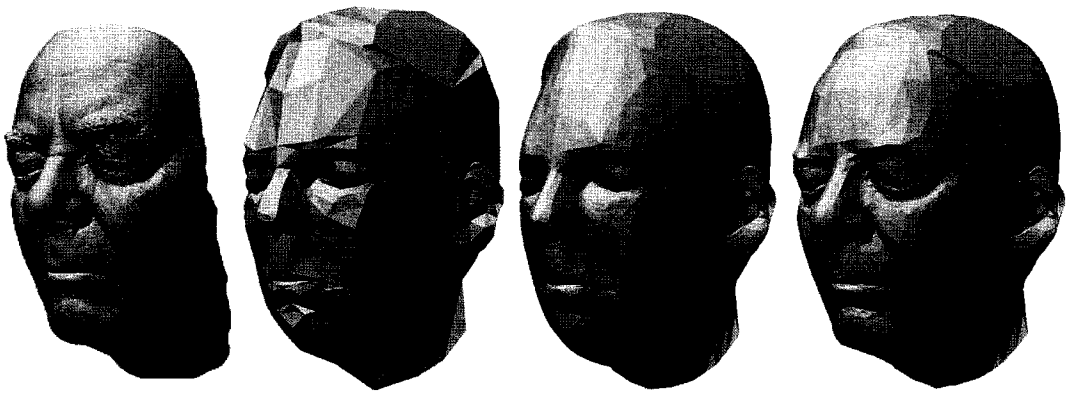
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

where $\boldsymbol{\varphi}^{-1}$ denotes the inverse of $\boldsymbol{\varphi}$ and $\overline{\mathbf{X}}'_k$ is a column vector containing only the kth coordinate of every feature point. Accordingly, $\mathbf{W}_k$ is a column vector of N values. Note that $\boldsymbol{\varphi}$ remains constant; thus, $\boldsymbol{\varphi}^{-1}$ is static as well. The matrix $\boldsymbol{\varphi}^{-1}$ can be decomposed using LU decomposition to make computation more efficient if the weight vectors need to be recomputed multiple times.

3.2.2.4 Surface Deformation

The generic model is deformed to approximate the scanned surface after network training and weight vector initialization have been performed. The RBF networks $\mathrm{R}_k(\overline{x})$ are evaluated using each vertex of the generic model as the input point $\overline{x}$. In doing so, the RBF networks interpolate the new positions of the generic model's non-feature points to define the vertices for the feature-based model. Note that $(\mathrm{R}_1(\overline{\mathbf{X}}_j), \mathrm{R}_2(\overline{\mathbf{X}}_j), \mathrm{R}_3(\overline{\mathbf{X}}_j)) = \overline{\mathbf{X}}'_j$ as expected, i.e., the RBF networks evaluate the generic model feature point $\overline{\mathbf{X}}_j$ to the scanned model feature point $\overline{\mathbf{X}}'_j$. Figure 3.3(b) shows the feature-based model which results from deforming the generic model using the feature points detected on the scanned surface shown in Figure 3.3(a). The derivations of the models shown in Figure 3.3(c) and (d) are discussed in a later section.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



(a) (b) (c) (d)

Figure 3.3: Construction of a low resolution model from a scanned face. (a) Scanned face (1,000,000 triangles). (b) Full head feature-based model (1,485 triangles). (c) Full head preliminary low resolution model (5,940 triangles). (d) Full head low resolution model obtained after shape refinement (5,940 triangles).

3.3 Representation of Shape and Skin Detail

The feature-based model described in the previous section is the principal building block of the multi-resolution hierarchy. This modeling scheme represents the scanned model's shape and skin detail information at several resolutions. Figure 3.4 shows the shape and skin detail decomposition hierarchy.

The feature-based model is the first in a series of models which is iteratively increased in resolution to produce the subsequent model in the series. This chain yields three shape models – a *low resolution model*, a *medium resolution model* and a *high resolution model* – with each model capturing shape information for its respective resolution. The transition from one model to the subsequent one is made by applying subdivision and then shape refinement. Skin detail (also referred to as simply "detail") elements are represented with the symbols $D_{H\text{-}L}$, $D_{H\text{-}M}$ and $D_{M\text{-}L}$. They can be loosely

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

defined as the difference between two models in the hierarchy. 3D skin elements, denoted by the symbols $S_{O\text{-}H}$, $S_{O\text{-}M}$ and $S_{O\text{-}L}$, loosely represent the differences between the original scanned model and a model in the hierarchy. At this time, the distinction made between detail and 3D skin can be difficult to comprehend since they are both described as the difference between two models. However, it is a conscious decision to distinguish one from the other because they may be similar in loose description but are vastly different in mathematical definition. This distinction becomes clear in a later section.

Before continuing with the discussion about shape and detail, it should be remarked that none of the models depicted in this thesis are texture-mapped. Traditionally, skin detail is added to 3D faces at render time by applying an appropriate texture map. Only the raw geometry of the polygonal models is visualized in this thesis and all visible skin detail is modeled using polygons. Thus, the remarks about "applying" or "adding" skin detail or 3D skin are purely in the geometrical sense. Neither skin detail nor 3D skin is a texture map.

Note that the entities in Figure 3.4 are enclosed in either a rectangle or an ellipse. The rectangular notation signifies that the entity is structurally consistent across all instances. [Here, an instance refers to the entity that corresponds to or is derived from a particular scanned model.] For skin detail, the meaning of "structural consistency" is somewhat similar to that for the models. Detail is structurally consistent because the quantity and organization of the encapsulated information (i.e., its structure) is identical for all instances of a particular level (e.g., $D_{M\text{-}L}$). The structure of skin detail, by itself, does not reveal the identity of the specific model instances from which it is derived. For geometry, the resolution (low, medium or high) is considered structurally consistent if all

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

instances are topologically equivalent and they contain the same number of vertices and polygons. The elliptical enclosure denotes that the structure of the entity is context dependent, varying from one instance to another. For example, the structure of $S_{O\text{-}L}$ can differ from one face to another and is dependent on the original scanned model being represented. Scanned models are not structurally consistent because they are arbitrarily structured and their resolutions can vary.

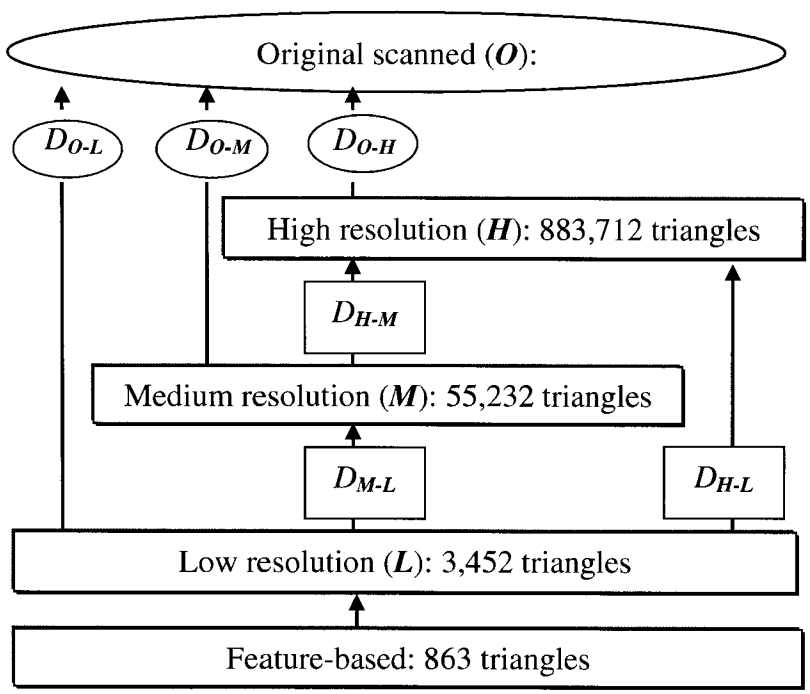


Figure 3.4: Shape and skin detail hierarchy. $D_{A\text{-}B}$ represents skin detail calculated between resolutions A and B. $D_{O\text{-}A}$ is referred to as the 3D skin representation of the original scanned model (O) computed on resolution A. The rectangle denotes structural consistency across all instances and the ellipse indicates structural inconsistency (i.e., different structures) among instances. Triangle counts correspond to the models that retain only the face region.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3.3.1 Multi-resolution Shape Construction

The feature-based model is a coarse approximation of the scanned model and serves only as the starting point for multi-resolution (shape) model construction. The feature-based model and its derivatives must undergo increases in resolution (i.e., polygon count) in order to realize the hierarchy. A simple way to accomplish this is to repeatedly apply a subdivision scheme on the model's polygons to achieve the desired resolutions. Subdivision splits each polygon into four smaller ones and also smoothes the resulting overall surface. However, if subdivision was the only step performed to produce the low, medium and high resolution models, the accuracy of these models would be unsatisfactory since the feature-based model itself is admittedly not very accurate. The accuracy of the subdivided models can be improved quite easily by using point projection to ensure that their vertices lie on the actual surface being represented.

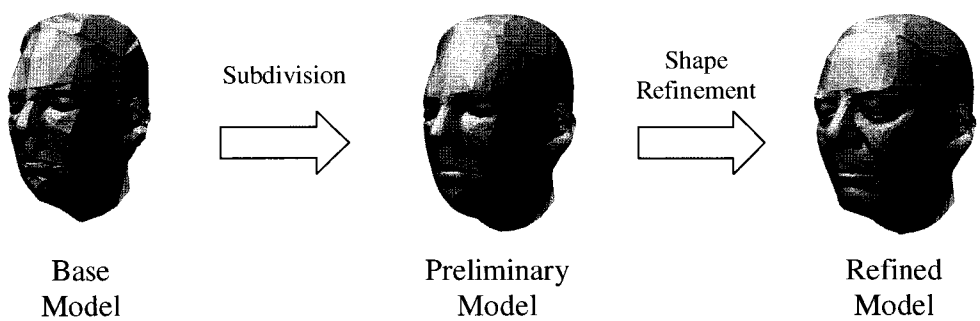


Figure 3.5: Procedure for creating a shape model in the multi-resolution hierarchy.

The sequence of steps taken to produce each of the three multi-resolution models is essentially identical except for variations in the number of subdivisions used. The general procedure is illustrated in Figure 3.5 and Table 3.1 lists the resolution-specific subdivision statistics.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Base Model	Number of Subdivisions Applied to Base Model	Refined Model
Feature-based	1	Low
Low	2	Medium
Low or Medium	4 or 2	High

Table 3.1: Derivation of multi-resolution models.

Consider the construction of the low resolution model. The feature-based model, which serves as the *base* model, is subdivided once using Loop's scheme [Loop 87] to produce the *preliminary* low resolution model (also illustrated in Figure 3.3(c)). The triangle count increases by a factor of four to 5,940. Shape refinement (discussed in Section 3.3.2) is then performed on the preliminary low resolution model to produce the *refined* model, which in this instance is the low resolution model (also depicted in Figure 3.3(d)).

Two points deserve special mention. Firstly, the high resolution model may be created in two different ways: it may be derived from the low resolution model (subdivided four times) or the medium resolution model (subdivided twice). The final resolution is the same in either case, but there is a minor difference in the final surface appearance. Secondly, since the scanned models (refer to Section 3.4) capture only the face region of the subjects, the back of the head and the ears of each feature-based model are actually removed prior to subdivision in experiments. The full head models have been used in illustrations to this point rather than the actual experimental models to demonstrate that full heads could be created when only face data is available. All hierarchy models depicted in later chapters are the true models of the hierarchy, capturing only the face area. Although this choice yields less visually attractive models, removing

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the superfluous portions of the feature-based model leads to a much more compact hierarchy, considering that each step of Loop subdivision increases the triangle count by a factor of four. The triangle counts given in Figure 3.4 are for the true models.

3.3.2 Shape Refinement

Shape refinement is performed on a preliminary model to obtain a closer resemblance to the scanned model. It is accomplished by projecting each vertex of the preliminary model onto the scanned model's surface. The basic premise is to associate a ray with each vertex of the preliminary model. Each vertex is then displaced to the intersection point of its ray and the scanned surface to obtain the vertex for the refined model. The intersection point does not necessarily correspond to a vertex of the scanned surface. The 2D interpretation of point projection is illustrated in Figure 3.6.

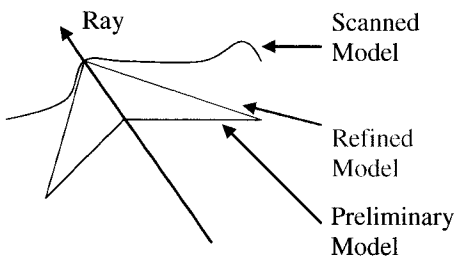


Figure 3.6: 2D interpretation of shape refinement using point projection.

Two commonly used projection schemes are vertex normal projection and cylindrical projection. Cylindrical projection was chosen for multi-resolution modeling because previous work [Soon 06 "Exaggeration"; Soon 06 "Shape"] demonstrated that vertex normal projection led to poorer results. The two subsections that follow describe each scheme in more detail.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3.3.2.1 Vertex Normal Projection

For vertex normal projection, each ray passes through its associated vertex and is in the direction of the vertex normal at said vertex. Using this scheme to refine preliminary low resolution models leads to errors in areas such as the eyes and mouth because those areas are coarsely approximated in the feature-based model. Since the shape is rough, the preliminary model may be quite far from the scanned model, which leads to intersecting rays as shown in Figure 3.7. After shape refinement, the refined model can suffer from self-folding and self-penetration.

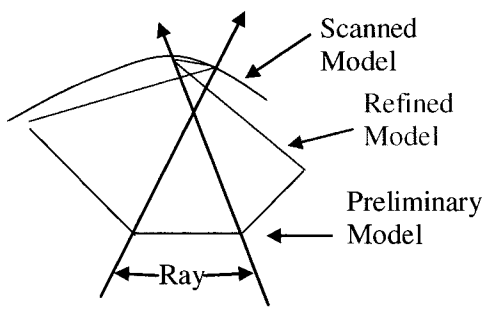


Figure 3.7: 2D interpretation of intersecting rays during vertex normal projection. The resulting refined model suffers from self-folding and self-penetration.

This type of problem was avoided in previous work [Soon 06 "Exaggeration"; Soon 06 "Shape"] by imposing a global threshold on the distance by which a vertex can be displaced as a result of projection. Consequently, areas which are coarsely approximated in the preliminary low resolution model remained so in the refined model. This had an adverse effect on the reconstruction of scanned models because another small displacement threshold (associated with model reconstruction) was also used to ensure that self-folding and self-penetration would not corrupt the reconstruction results. [Model reconstruction is discussed in Chapter 4.] The reconstructed scanned models contained

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

many large holes because of this safeguard. A more liberal displacement threshold led to fewer and smaller holes but at the cost of more self-penetration problems. Figure 3.8(a)-(c) show the original scanned surface, a reconstructed scanned surface with many large holes (but no self-penetration) and a reconstructed surface with many self-penetration problems (but few small holes), respectively, that exemplify the experiences from previous work. Previous work required the use of a simplified model to avoid the problems in Figure 3.8(b) and Figure 3.8(c).

(a) (b) (c)

Figure 3.8: Scanned faces produced during previous work [Soon 06 "Exaggeration"; Soon 06 "shape"]. (a) Original scanned face. (b) Reconstructed scanned face obtained using a small displacement threshold to avoid self-penetration issues. (c) Reconstructed scanned face obtained using a large displacement threshold to avoid large holes.

3.3.2.2 Cylindrical Projection

For cylindrical projection, each ray passes through its associated vertex and is perpendicular to a *center axis*. Figure 3.9 shows how cylindrical projection works for the preliminary low resolution model. The center axis of the cylinder is established by finding the line which satisfies two requirements: 1) it passes through the center of the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

full head feature based model; and 2) it is parallel to the y axis. Each vertex of the preliminary low resolution model is displaced to the intersection of its ray and the scanned surface as long as the displacement is below a user-defined threshold. Figure 3.3(d) and Figure 3.5 show an example of the low resolution model obtained from applying cylindrical projection to a preliminary low resolution model.

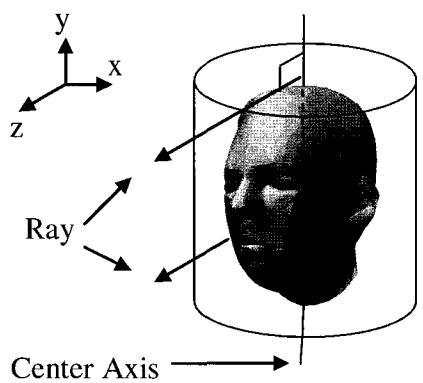


Figure 3.9: Cylindrical projection.

Unlike vertex normal projection, self-penetration and self-folding issues are not encountered when refining the preliminary low resolution model because vertices are well distributed and the rays are defined relative to the same center axis. The latter property means that rays do not intersect one another unless they are parallel. Thus, cylindrical projection would not experience the same shortfall shown in Figure 3.7 despite the large separation between the scanned model and the preliminary model. This also means that a larger displacement threshold can be used for cylindrical projection than for vertex normal projection. Cylindrical projection produces better refined models than vertex normal projection, which in turn produce better reconstructed scanned models. Model reconstruction is addressed in Chapter 4 and some reconstructed scanned models are shown in Section 4.5.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

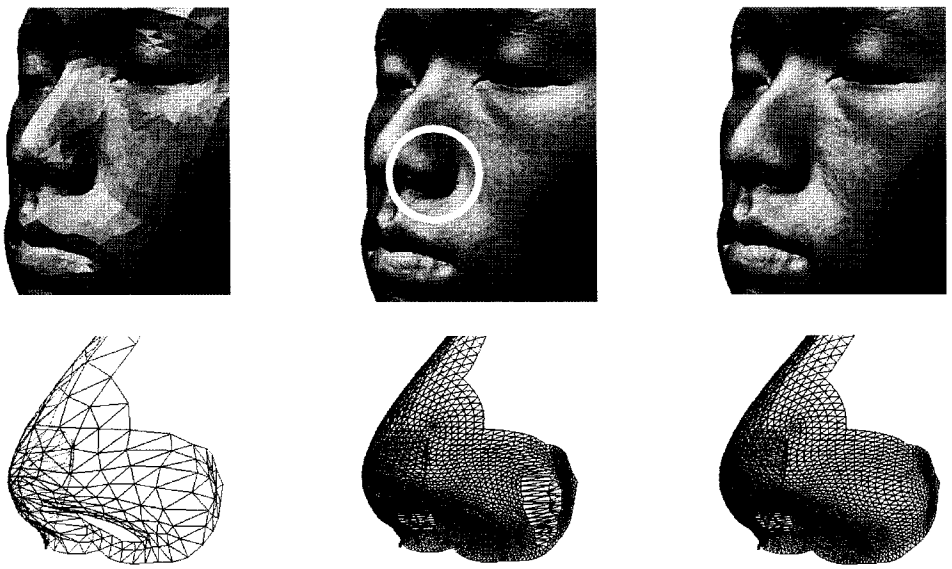
Note that the full head is not necessary at the time when refinement is actually performed. After creating a full head feature-based model, the center axis can be calculated and then the feature-based model can be cropped to retain only the face. Since neither subdivision nor cylindrical projection requires the cropped feature-based model to be relocated, the same center axis is valid for refining all preliminary models. It is also possible to ensure that every feature-based model created is centered about the same center axis, thus removing the need to calculate an axis for every head.

This projection scheme works well with preliminary low resolution models. The refined models bear closer resemblances to the scanned models which they are intended to represent. However, empirical results demonstrate that errors can arise when the scheme is performed on higher resolution preliminary models. These errors appear on the nose alae near the alar grooves. Figure 3.10(b) shows an example of this type of error in a medium resolution model. Such errors arise because points on the alae can be incorrectly projected onto the cheek or vice-versa. The distribution of vertices in the error area is noticeably different from the rest of the nose. This type of problem is not encountered when producing low resolution models as exemplified in Figure 3.10(a). It is possible to use post-processing techniques to remedy such errors, but experiments demonstrated that it is also possible to circumvent the problem by introducing a second cylinder axis for projection in the nose region. The nose region is refined using this second axis while the remainder of the model uses the center axis. The second axis shall be referred to as the *nose projection axis* to avoid confusion.

Like the center axis of the cylinder, the nose projection axis is parallel to the y axis. However, it must be situated nearer to the nose so that points are not erroneously

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

projected onto the cheeks or vice-versa. A suitable location can be calculated using the 3D feature point information obtained during feature detection (see Section 3.2.1).



(a) (b) (c)

Figure 3.10: Shape refinement results. (a) Low resolution model achieved using only center axis. (b) Medium resolution model achieved using only center axis. (c) Medium resolution model achieved using both center and nose projection axes.

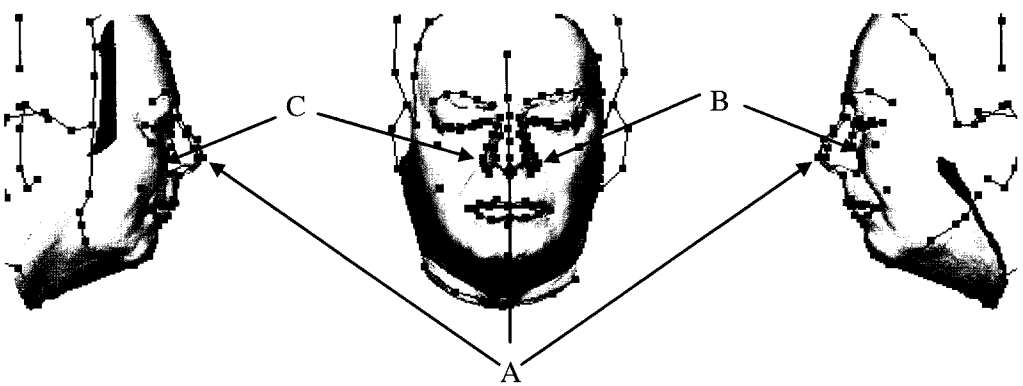


Figure 3.11: Three feature points (A, B and C) used to calculate the nose projection axis.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Three feature points are required for this exercise: the tip of the nose (A) and the leftmost and rightmost points of the nose alae (B and C) as depicted in Figure 3.11. The point in space denoted by

$$(x, y, z) = (A_x, A_y, 0.75 \times \min(B_z, C_z) - 0.25 \times A_z) \tag{3.6}$$

is calculated, where the subscript notation is used to indicate a specific coordinate (i.e., x, y or z) of a point. The nose projection axis is then defined as the line which passes through this point. Note that the derivation of the z value in Equation (3.7) may be rewritten as

$$0.75 \times \min(B_z, C_z) - 0.25 \times A_z = \min(B_z, C_z) - 0.25 \times (A_z - \min(B_z, C_z)), \tag{3.7}$$

to more clearly indicate that the z value is extrapolated from the z values of A, B and C. The resulting point has a smaller z value than any of A, B and C, meaning that the nose projection axis is located between the center axis and the nose. The value of 0.25 used in Equation (3.7) was chosen based on trial-and-error experimentation.

Figure 3.10(c) shows the result of shape refinement obtained when both the center and nose projection axes are used. The two-axis scheme does not suffer from the same problem as the one-axis scheme because the nose projection axis is closer to the nose region, meaning that points are not erroneously projected.

Shape refinement can be expedited by using a lower resolution version of the scanned model. The execution time of the refinement procedure should be faster since a lower polygon count translates into fewer triangles to check for ray intersection. The compromise in resolution is worthwhile for refining the preliminary low resolution model since a high level of detail is not required. Depending on the choice of the lower resolution used, the same can be said for refining the preliminary medium resolution

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

model as well. Models containing about 100,000 triangles were used in experiments to refine the preliminary low and medium models.

3.3.3 Multi-resolution Skin Detail Representation

As stated earlier, the skin detail elements $D_{M\text{-}L}$, $D_{H\text{-}M}$ and $D_{H\text{-}L}$ loosely represent the difference between two models of differing resolutions. This "difference" is defined more rigorously here. The information in Table 3.1 indicates that it is possible to evolve a base model into a refined model using a combination of subdivision and shape refinement. This may be written loosely as

$$\textit{refined model} = \textit{base model} +_{pseudo} \textit{subdivision} +_{pseudo} \textit{shape refinement} \tag{3.8}$$

to facilitate discussion, where the $+_{\text{pseudo}}$ symbol denotes pseudo-addition, i.e., subdivision is "added to" (performed on) the base model and then shape refinement is "added to" (performed on) the resulting subdivided model. Since the preliminary model is simply the result of applying subdivision to the base model, it follows that

$$\textit{refined model} = \textit{preliminary model} +_{pseudo} \textit{shape refinement}, \tag{3.9}$$

Since shape refinement does not alter the preliminary model's structure, one can see that there is an exact one-to-one point correspondence between the preliminary and refined models. More precisely, the models share the same geometric structure. Equation (3.9) may be rewritten to derive the "detail" term as

$$\textit{detail} = \textit{shape refinement} = \textit{refined model} -_{pseudo} \textit{preliminary model}. \tag{3.10}$$

Because the $+_{pseudo}$ symbol used in Equations (3.8) and (3.9) denotes pseudo-addition, the $-_{\text{pseudo}}$ operator in Equation (3.10) accordingly denotes pseudo-subtraction, i.e., the result of "subtracting" the preliminary model from the refined model is detail. However,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

because of the point correspondence between the two models, the $-_{\text{pseudo}}$ operator in Equation (3.10) can have a mathematical meaning as well: it can be interpreted as vertex subtraction which calculates vector offsets. Model subtraction in the present context may then be more formally written as

$$\textit{refined model} -_{pseudo} \textit{preliminary model} = \{v_{i,r} - v_{i,p},\ i = 1, \ldots, N\} \tag{3.11}$$

where $v_{i,r}$ is the ith vertex in the refined model, $v_{i,p}$ is the ith vertex in the preliminary model and N is the total number of vertices in the refined/preliminary model. The result of the $-_{\text{pseudo}}$ operation is always a set of N offset vectors, regardless of the face which is being represented by the models in question, provided that the resolutions of the preliminary and refined models remain constant. For this reason, this form of detail is considered structurally consistent.

Table 3.2 lists the formulations for the three skin detail levels. Recall that the high resolution model may be created either by subdividing the medium resolution model twice or by subdividing the low resolution model four times. The implied method is the one which leads to one-to-one point correspondence between the models in the operation.

Skin Detail Level	**Formulation**
$D_{M\text{-}L}$	Medium resolution $-_{\text{pseudo}}$ Preliminary medium resolution
$D_{H\text{-}M}$	High resolution $-_{\text{pseudo}}$ Medium resolution subdivided twice
$D_{H\text{-}L}$	High resolution $-_{\text{pseudo}}$ Low resolution subdivided four times

Table 3.2: Formulations of multi-resolution skin detail.

Note that

$$\textit{refined model} = \textit{preliminary model} +_{pseudo} \textit{detail}, \tag{3.12}$$

which can be interpreted as a reconstitution of shape (subdivided base model) and skin detail to form a higher resolution shape (i.e., refined model). Here, the $+_{\text{pseudo}}$ symbol

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

means that the ith offset vector in the detail representation is added to the ith vertex of the preliminary model, for $i = 1, \ldots, N$. This recomposition holds true for any two shape levels of the hierarchy and their associated skin detail.

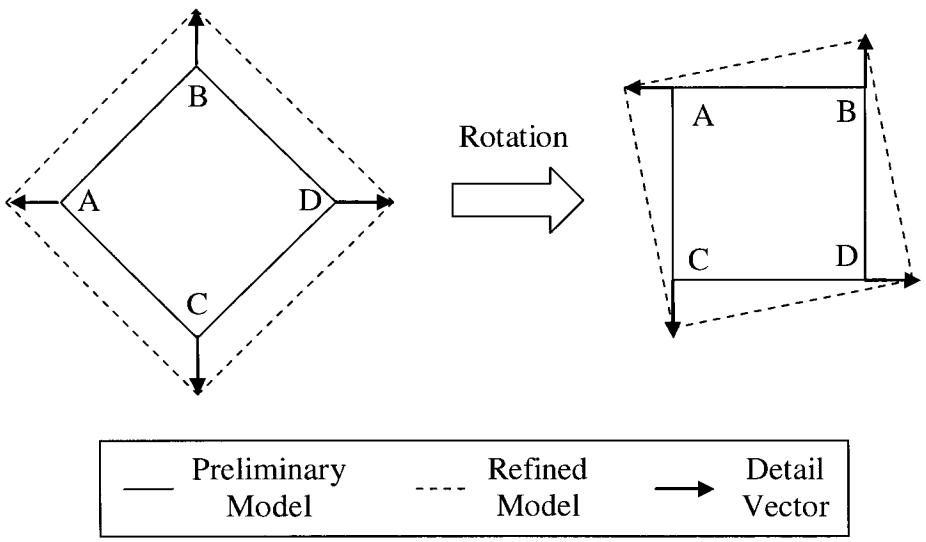


Figure 3.12: Example of incorrect reconstitution result. The preliminary model is altered by rotation, but the reconstituted model is not rotated similarly.

The reader may notice that the direction vectors of the offset vectors are statically defined, i.e., they are calculated based on the vertices of the refined and preliminary models. Therefore, the result of performing recomposition as given in Equation (3.12) may be incorrect if the preliminary model is altered (by way of altering the base model). This is demonstrated with a trivial example in Figure 3.12. The preliminary model in the left image is rotated 45 degrees clockwise prior to applying detail vectors on the four vertices A, B, C and D. The expected result is a refined model that is also rotated 45 degrees clockwise. However, the orientation of the refined model that is actually constructed is different from that of the preliminary model. The detail vectors do not account for any change in the preliminary model. Hence, when adding skin detail for

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

reconstitution, the preliminary model must be in its original form to ensure that valid results are achieved.

3.3.4 Multi-resolution 3D Skin Representation

3D skin is also said to be the difference between the original scanned model and a shape model. As with skin detail, it would be useful to express 3D skin as

$$\textit{3D skin} = \textit{original scanned model} -_{pseudo} \textit{preliminary model} \tag{3.13}$$

in a loose way, where the $-_{pseudo}$ operator represents pseudo-subtraction, i.e., the result of "subtracting" the preliminary model from the original scanned model is 3D skin. Following this statement, pseudo-addition is defined such that 3D skin is "added to" the preliminary model to obtain the original scanned model. However, the original scanned model, unlike the refined model in the skin detail discussion, is of arbitrary resolution and topology. Hence, there is no one-to-one point correspondence between the two models in the 3D skin context as there is in the previous case. The notion of model subtraction stated in Equation (3.10) is inapplicable here.

For 3D skin, pseudo-subtraction has no mathematical interpretation, but it is defined to be a *model parameterization* operation which maps the scanned model's vertices onto the surface of the preliminary model. The parameterization result and the scanned model's original vertex connectivity information constitute a 3D skin representation of the scanned model defined on the preliminary model. Pseudo-addition is then defined to be a *model reconstruction* operation which reconstructs the vertices of the scanned model by applying the parameterization stored in the 3D skin representation back onto the preliminary model. The reconstructed vertices and the vertex connectivity information stored in the 3D skin representation constitute the reconstructed scanned

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

model. Model parameterization and reconstruction are discussed fully in Chapter 4. Table 3.3 lists the three formulations used to produce multi-resolution 3D skin.

3D Skin Level	Formulation
$D_{O\text{-}H}$	Original scanned model $-_{\text{pseudo}}$ High resolution model subdivided twice
$D_{O\text{-}M}$	Original scanned model $-_{\text{pseudo}}$ Preliminary high resolution model
$D_{O\text{-}L}$	Original scanned model $-_{\text{pseudo}}$ Preliminary medium resolution model

Table 3.3: Formulations of multi-resolution 3D skin.

There are three major distinctions between skin detail and 3D skin. Firstly, pseudo-subtraction in the skin detail case has a mathematical interpretation, but not in the 3D skin case. While skin detail is established by exploiting the one-to-one point correspondence to perform vertex subtraction, 3D skin requires a model parameterization operation to be performed. Secondly, the model which results from "adding" skin detail to a preliminary model has the same structure as the preliminary model, but the result for the 3D skin case has the structure of a scanned model. Thirdly, from a visual appearance standpoint, 3D skin is able to show more surface detail than skin detail, although that is not to say that skin detail cannot inject a good level of surface detail into a preliminary model.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3.4 Database of Scanned Faces

Figure 3.13 shows a small database of four scanned faces used for experiments. The faces were scanned and contributed by the Ottawa-based company XYZ RGB[2], which specializes in extremely high resolution 3D scanning services. The faces provided by XYZ RGB range in resolution from about one million to three million triangles. All of the models capture only the face region of the subject. These faces will be referenced in later chapters by their labels (Person A, Person B, etc.). Appendix A illustrates the low, medium and high resolution models for each of the four subjects.

Person A	Person B	Person C	Person D
1.0 million	2.2 million	3.0 million	1.5 million

Figure 3.13: Database of scanned faces used for experiments. The resolution stated under each face is measured in triangles.

[2] http://www.xyzrgb.com

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 4 3D Surface Capture and Reconstruction

It is stated in Section 2.6.2 that texture-based displacement mapping [Cook 84] is inadequate for reconstructing extremely high resolution geometry from low resolution (base) models. This chapter demonstrates an example of its failure and outlines a suitable technique that can successfully accomplish this task in such a way that the base model's shape guides the reconstruction procedure. Although this capture technique is used in the context of the multi-resolution representation to calculate 3D skin, it is considered to be a general purpose procedure because it neither makes assumptions about the resolutions of the two surfaces nor does it assume that the surfaces correspond to human faces.

Section 4.1 outlines the test which led to the realization that displacement mapping cannot faithfully reconstruct very high resolution scanned faces. The fundamental mechanism of the successful surface capture procedure is a point-to-surface mapping scheme that is discussed in Section 4.2. The act of capturing a surface is accomplished by performing model parameterization on that surface using point-to-surface mapping and is covered in Section 4.3. The result of model parameterization is combined with the scanned model's topology (i.e., vertex connectivity) information to establish a 3D skin representation of the scanned model defined on the base model. Surface reconstruction is explained in Section 4.4. An error analysis of model parameterization and reconstruction follows in Section 4.5.

4.1 Geometry Reconstruction using Displacement Mapping

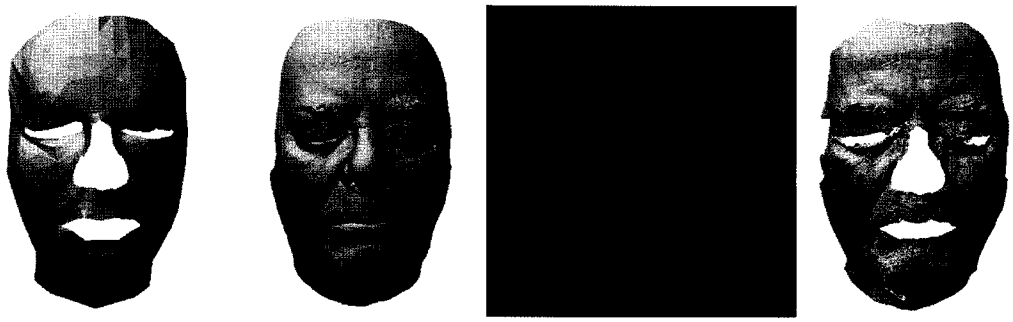
As discussed in Section 2.6.2, displacement mapping is a technique used to capture the geometrical detail of a high resolution surface by expressing it as a field of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

scalar/vector displacements over a low resolution surface. A displacement map is used to alter the geometry of the low resolution surface so that it approximates the high resolution surface. In previous work [Soon 06 "Shape"], the technique described in [Zhang Y 04], which combines hierarchal mesh refinement and displacement map sampling, was put into practice to evaluate displacement mapping's applicability for reconstructing very dense geometry. The surface detail of a 2,000,000 triangle scanned face model was captured in a high resolution displacement map using a low resolution face as the base model (Figure 4.1(a)-(c)). The nose and mouth were removed from the pictured low resolution model so that it could be trivially UV mapped with as few overlapping textures as possible. To reconstruct the scanned face, the low resolution face model was iteratively subdivided five times using the quaternary scheme (Figure 4.2) and every vertex in the final subdivided model was displaced according to the values in the map. The reconstructed model's final resolution was 2,002,944 triangles and it is shown in Figure 4.1(d).

Disregarding the glaring aberrations in the reconstructed model, a close-up comparison (Figure 4.3) with the original scanned model also showed imperfections at a finer level. The triangles of the reconstructed model can be seen much more clearly than those of the scanned model, which made the reconstructed surface look unnatural. Furthermore, this deficiency was evidenced in many regions of the face in which point density was high. Overall, the reconstructed model's appearance was not comparable to the original scanned model.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



(a) (b) (c) (d)

Figure 4.1: Geometry reconstruction result using displacement map sampling. (a) Cropped low resolution face (3,780 triangles). (b) Scanned face (2,000,000 triangles). (c) High resolution RGB displacement map (2048 x 2048 pixel resolution). Positive displacements are encoded in the red channel while negative displacements are stored in the green channel. (d) Reconstructed model (2,002,944 triangles).

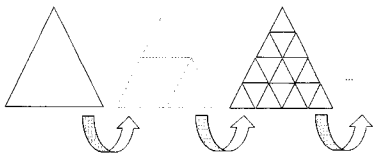
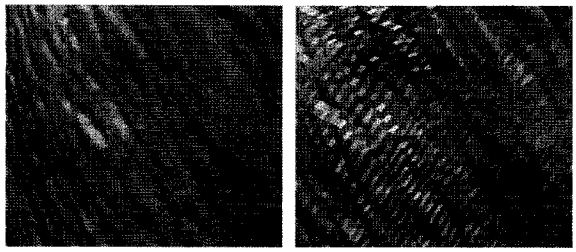


Figure 4.2: Iterative quaternary subdivision. Each existing edge is split at its midpoint. Existing vertices are not displaced.



(a) (b)

Figure 4.3: Close-up comparison of the left cheek of the original and reconstructed scanned models shown in Figure 4.1. (a) Original model. (b) Reconstructed model.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The chief reason for displacement mapping's failure to satisfactorily reconstruct the scanned model geometry is its inability to properly recreate the vertices of the model, which ultimately determines the appearance of the reconstructed surface. This inability can be attributed to two factors. The first is the way in which a displacement map is generated. A displacement value is calculated for each pixel of the (cropped) low resolution model's UV map. Each pixel corresponds to a point on the surface of the low resolution model and the displacement is measured between the low resolution model and the scanned model in the direction of the interpolated normal at said point. Since it is impossible to record a displacement for every point on the scanned model surface, some scanned model vertices are not accounted for in the displacement map (Figure 4.4(a)). As a result, reconstructing the model using the displacement map can produce a substandard result (Figure 4.4(b)). The second factor attributed to the downfall is the way in which the reconstructed model is achieved. As mentioned previously, quaternary subdivision is used to increase the resolution of the low resolution model whose surface is then perturbed by sampling the displacement map at each vertex. Even if the location of every vertex in the scanned model was successfully encoded in the displacement map, there is no guarantee that the scanned model vertices were captured using the low resolution model vertices (Figure 4.4(c)). As a result, the reconstructed model will differ from the original model.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

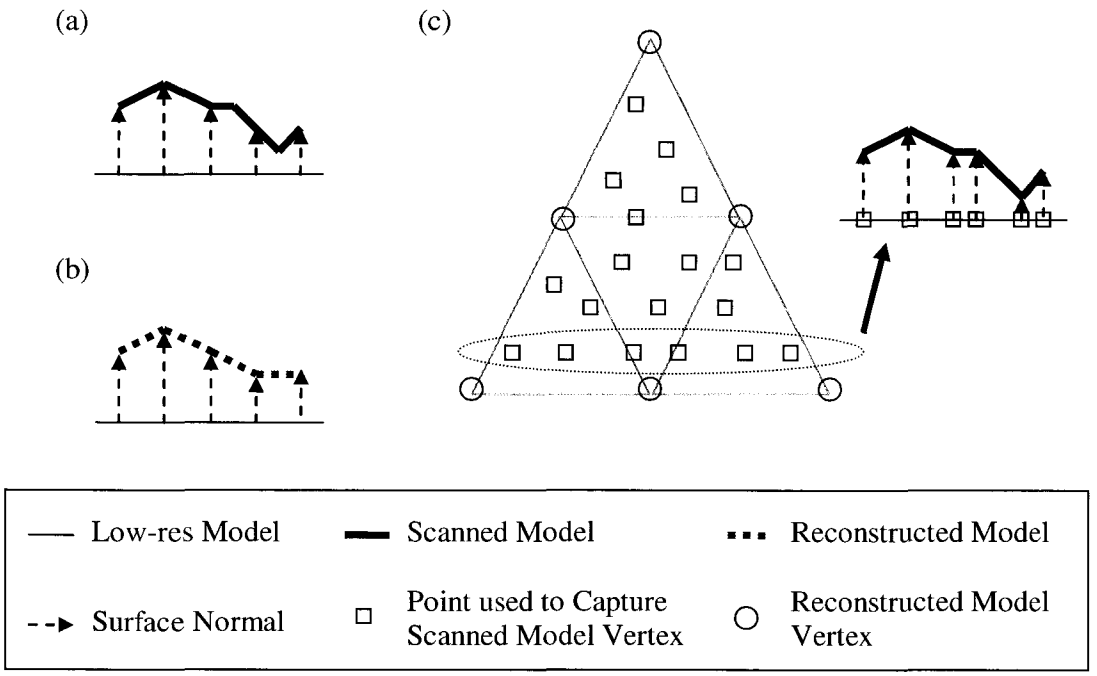


Figure 4.4: Generation and usage of displacement map. (a) Displacement map generation (sampling at regular intervals). (b) Scanned model reconstruction using displacement map from (a). (c) Illustration of inability to reconstruct original scanned model vertices despite having the necessary values in the displacement map.

Essentially, a different technique which can capture (and then consequently reproduce) the original vertices of the scanned model is needed. One might argue that the limitation shown in Figure 4.4(c) may be overcome by recording the locations of the points used to capture the scanned model vertices. However, one must keep in mind that in such a case, there is a great deal of computation time needed to produce a displacement map and much of this work is essentially wasted since only a relatively small set of displacement values would actually be used during model reconstruction.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4.2 Point-to-Surface Mapping

The purpose of point-to-surface mapping [Sun 01; Zhang Y 04] is to map a point to a polygonal surface (both of which reside in 3D space) using an interpolated vertex normal of the surface. This idea is based on simplification envelopes first proposed in [Cohen 96]. Consider the case where a point V is mapped to a triangle ABC belonging to a surface (Figure 4.5). The ray origin P and the interpolated vertex normal $\overline{\mathbf{N}}_{\mathbf{P}}$ which result from this mapping are given by

$$\begin{aligned} \mathrm{P} &= (1-u-v)\mathrm{A} + u\mathrm{B} + v\mathrm{C} \\ \overline{\mathbf{N}}_{\mathbf{P}} &= (1-u-v)\overline{\mathbf{N}}_{\mathbf{A}} + u\overline{\mathbf{N}}_{\mathbf{B}} + v\overline{\mathbf{N}}_{\mathbf{C}} \end{aligned}, \tag{4.1}$$

where u and v are the 2D barycentric coordinates of P with respect to $\Delta\mathrm{ABC}$ and $\overline{\mathbf{N}}_{\mathbf{A}}$, $\overline{\mathbf{N}}_{\mathbf{B}}$ and $\overline{\mathbf{N}}_{\mathbf{C}}$ are the vertex normals at A, B and C respectively. The barycentric coordinates express the position of P relative to the vertices of $\Delta\mathrm{ABC}$. For all points lying within or along the edges of $\Delta\mathrm{ABC}$, the constraints $u, v \in [0,1]$ and $u + v \leq 1$ are satisfied. For all other points, either or both of these constraints are violated. The position of V can then be expressed as

$$\mathrm{V} = \mathrm{P} + d\frac{\overline{\mathbf{N}}_{\mathbf{P}}}{\left\|\overline{\mathbf{N}}_{\mathbf{P}}\right\|}, \tag{4.2}$$

where d is the signed distance from P to V in the direction of $\overline{\mathbf{N}}_{\mathbf{P}}$ and $\| \ \|$ denotes vector magnitude.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

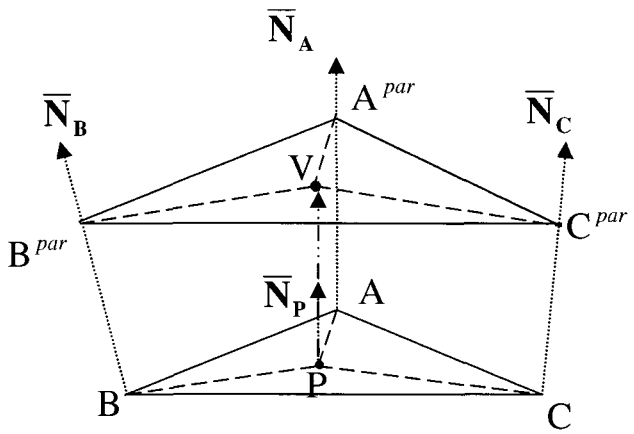


Figure 4.5: Example of point-to-surface mapping.

To calculate the values of u and v in Equation (4.1), a new triangle $A^{par}B^{par}C^{par}$ that is parallel to ΔABC and whose vertices are coplanar with V is defined (Figure 4.5). The vertices of this new triangle are obtained by finding the intersection of $\overline{\mathbf{N}}_A$, $\overline{\mathbf{N}}_B$ and $\overline{\mathbf{N}}_C$ with the parallel plane. Computing the barycentric coordinates of V with respect to $\Delta A^{par}B^{par}C^{par}$ yields the values of u and v. In other words,

$$V = (1-u-v)A^{par} + uB^{par} + vC^{par}. \tag{4.3}$$

Note that vertex normal interpolation ensures that the normals transition smoothly and continuously between adjacent triangles (Figure 4.6a). On the other hand, if the same surface normal is used at all three triangle vertices during mapping, the transition between adjacent triangles can be discontinuous (Figure 4.6b). This, in turn, means that points situated in these voids cannot be mapped onto the surface.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

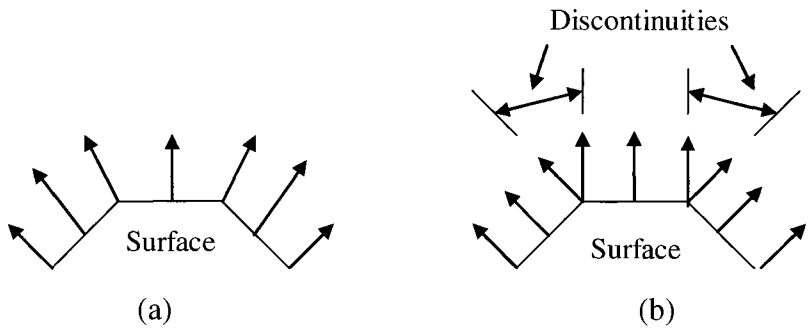


Figure 4.6: Point-to-surface mapping using different normals. (a) Vertex normals (no discontinuities). (b) Surface normals (discontinuities can occur).

In some instances, it is possible to obtain multiple mappings using different triangles where $u, v \in [0,1]$ and $u + v \leq 1$ are satisfied. An example of one such instance is shown in Figure 4.7 where the point V can be mapped to three different locations on the surface. If only one mapping is desired, heuristics can be used to guide the selection process. The procedure used to select the mapping for a particular vertex is covered in the next section.

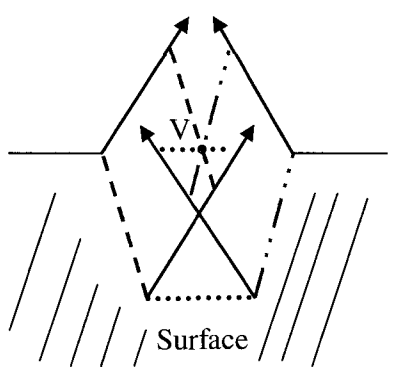


Figure 4.7: Example of a point for which multiple mappings are possible. Each pair of parallel planes is marked with a different line style.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4.3 Model Parameterization

As stated in Section 2.3.2, high resolution model parameterization is mentioned in [Hilton 02]. This process can be generalized to apply to two models of arbitrary resolution. This generalization can simply be referred to as model parameterization where a *subject model* is parameterized with respect to a *control model.* Parameterization is accomplished by performing point-to-surface mapping to obtain a set of mapping parameters (I,u,v,d) for each vertex in the subject model, where I is an identifier for the control model triangle onto which the vertex is mapped (e.g. index into an ordered triangle list), u and v are the barycentric coordinates of the mapped surface point relative to triangle I and d is the signed displacement along the interpolated vertex normal. Additionally, the vertex connectivity of the subject model must be stored separately to recover the topology later. Together, the sets of mapping parameters and the vertex connectivity information establish a 3D skin representation for the subject model defined on the control model.

4.3.1 Establishment of Mapping Parameters

A two-step process is required to find triangle I for each scanned model vertex V. The first step consists of accumulating a virtual "bin" of subject model vertices for each control model triangle. Only subject model vertices in the vicinity of a triangle are added to its bin. This judgment is made by checking whether the vertices are situated within the boundaries of an expanded triangle bounding box (Figure 4.8) with the size of the expansion user-definable. Bins are not mutually exclusive so that multiple triangles are checked for each vertex.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

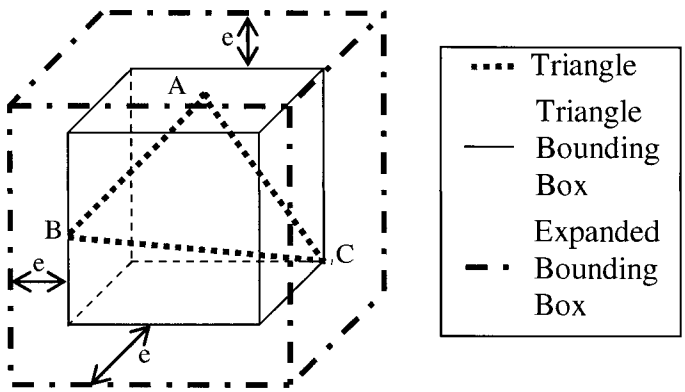


Figure 4.8: Bounding box of a 3D triangle ABC and its associated expanded bounding box. Each side of the triangle's bounding box is expanded outward by a value of e, where e is user-definable.

Once the bins have been populated, each control model triangle is paired with each of the subject model vertices in its bin. For instance, a bin containing ten vertices yields ten vertex-triangle pairs. The point-to-surface mapping scheme is then performed iteratively for all vertex-triangle pairs to calculate a set of mapping parameters (I,u,v,d) for each pair. As the iterative process progresses, each newly calculated parameter set for a particular subject model vertex is compared to the incumbent set for the same vertex. The incumbent set is the best set of mapping parameters encountered at any given point in the iterative process and there is at most one incumbent per vertex. A sequential set of heuristics is used to decide whether the newly calculated set should become the new incumbent. If there is no incumbent for a particular subject model vertex when the heuristics are used, the heuristics which compare the new set to the incumbent set are skipped.

- If the incumbent set's barycentric coordinates parameters satisfy the constraints $u,v\varepsilon[0,1]$ and $u+v\leq 1$, the new set is removed from

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

consideration if its parameters do not satisfy either constraint. This heuristic ensures that a set of mapping parameters which maps the vertex to a point lying outside of a triangle does not replace the incumbent set which maps the same vertex to a point within a triangle. If the vertex is mapped to a point lying outside of the triangle, the vertex may not be reconstructed properly should the state of the triangle change.

- The new set is removed from consideration if its barycentric coordinates parameters do not satisfy either $u, v \varepsilon [\alpha, \beta]$ or $u + v \varepsilon [\alpha, \beta]$, where and are user-definable. This heuristic allows the subject model vertex to be mapped to a point outside of a triangle when $\alpha < 0$ and/or $\beta > 1$. The vertex can still be reconstructed fairly well as long as α and β do not deviate too far from 0 and 1, respectively. As with the previous heuristic, the new set should not replace the incumbent if it maps the vertex outside of the accepted range.
- The new set is removed from consideration if its displacement parameter d exceeds a user-definable threshold. This heuristic favours vertex-triangle pairs in which the two elements are close to one another. These pairs typically lead to better reconstructions.
- The reconstruction error for each parameter set is the Euclidean distance between the reconstructed vertex and the original scanned model vertex. [Vertex reconstruction is discussed in Section 4.4.] The new set is removed from consideration if the reconstruction error exceeds a user-definable threshold.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

- The new set is removed from consideration if the absolute value of its displacement parameter d is larger than that of the incumbent. Again, this favours a triangle that is closer to the vertex.
- The new set is removed from consideration if its associated reconstruction error is larger than that of the incumbent.
- The new set becomes the incumbent if the preceding heuristics do not remove the new set from consideration.

After all vertex-triangle pairs are checked, the incumbent set for each vertex is used in the subject model's parameterization. If there is no incumbent for a vertex, that vertex has no mapping parameters in the model parameterization and, therefore, cannot be reconstructed. These cases require special attention and are discussed in Section 4.4.

4.3.2 Model Parameterization in Practice

Full model parameterization is generally not achievable (i.e. it is not always possible to obtain a set of mapping parameters for every vertex in the subject model) when the control model has open edges (e.g. those around the border of a scanned face). Typically, some of the subject model vertices near the open edges are missed because the control model's vertex normals along the open edges are not pointing in the optimal direction. This is not of great concern since vertices along open edges are not particularly significant, but the control model can be constructed such that its borders extend beyond those of the subject model to ensure proper detection.

The ability to capture vertices in the subject model greatly depends on the shape of the control model and, in particular, its approximation of the subject. The subject model is poorly parameterized if the control model is not a satisfactory approximation.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Previous work [Soon 06 "Exaggeration"] demonstrated the use of an intermediate simplified model between a low resolution model and a high resolution model to improve results. The premise is that the simplified model is better able to capture the high resolution surface than the low resolution surface. The simplified surface is parameterized with respect to the low resolution surface and the high resolution surface is parameterized with respect to the low resolution surface. With this configuration, changes can be propagated from the low resolution model to the simplified model and, in turn, to the high resolution reconstruction.

With regard to constructing 3D skin in the multi-resolution hierarchy, the face region is well captured, but the eyes and mouth areas are known to be problematic. Some vertices in these areas are missed during model parameterization. Section 4.4 discusses two ways in which this problem can be handled.

In theory, the constraints $u, v \in [0,1]$ and $u + v \leq 1$ should be enforced so that vertices are mapped to points lying within the boundaries of triangles. However, these constraints are overly stringent in most cases, causing vertices to not be mapped at all. In practice, the values of α and β (mentioned in Section 4.3.1) are not chosen to be 0 and 1, respectively. The values of α used to produce empirical results ranged from -0.1 to -10^{-7} while the values of β range from $(1 + 10^{-7})$ to 1.1.

The control model should be a smooth surface as suggested in [Lee A 00] so that the reconstructed subject model does not appear faceted if the control model is changed drastically. A smooth control model can be obtained by subdividing the control model before model parameterization. For example, if one wishes to use the low resolution model (of the multi-resolution modeling scheme) to capture the scanned surface, the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

preliminary medium resolution model may be used for model parameterization since it is a subdivided version of the low resolution model. An illustration of the impact of a smooth control surface is shown in Figure 4.10.

4.4 Model Reconstruction

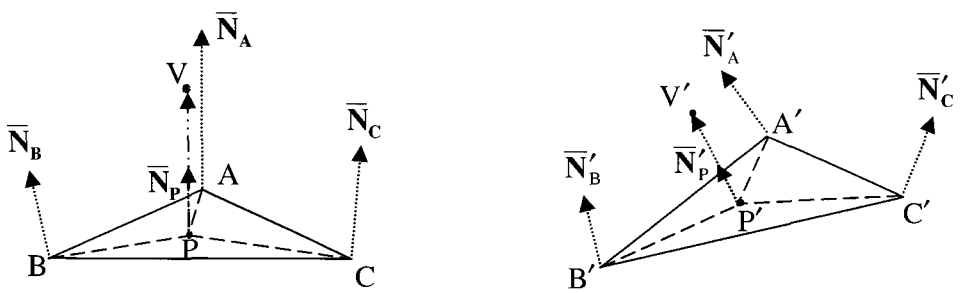


Figure 4.9: Reconstruction of a subject model vertex.

The premise of the model reconstruction algorithm is to use the subject model's parameterization (i.e. sets of mapping parameters) to reproduce the vertices of the subject model. Each set of mapping parameters (I, u, v, d), where I is the control model triangle identifier, u and v are the barycentric coordinates of the ray origin and d is the displacement along the interpolated vertex normal, is applied to the control model to reconstruct each subject model vertex (Figure 4.9). Equations (4.1) and (4.2) can then be restated as

$$\begin{aligned} \mathrm{P}' &= (1-u-v)\mathrm{A}' + u\mathrm{B}' + v\mathrm{C}' \\ \overline{\mathbf{N}}'_{\mathrm{P}'} &= (1-u-v)\overline{\mathbf{N}}'_{\mathrm{A}} + u\overline{\mathbf{N}}'_{\mathrm{B}} + v\overline{\mathbf{N}}'_{\mathrm{C}} \end{aligned} \tag{4.4}$$

and

$$\mathrm{V}' = \mathrm{P}' + d\frac{\overline{\mathbf{N}}'_{\mathrm{P}'}}{\left\|\overline{\mathbf{N}}'_{\mathrm{P}'}\right\|}, \tag{4.5}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

respectively. The set of all reconstructed vertices along with the original vertex adjacency information constitute the reconstructed subject model.

Note that the model reconstruction algorithm is only concerned with the present state of the control model and not with its state when the parameterization was obtained. The new labels assigned to the entities in Figure 4.9 are meant to reflect this fact. The use of vertex normals ensures that the shape of the control model influences the reconstruction of the subject model. An example of this principle is demonstrated in Section 5.2 where shape exaggeration is discussed.

As mentioned in the previous section, it is important that the control model's surface be smooth when reconstructing the subject model. Thus, the control model must be subdivided the same number of times prior to model reconstruction as it was before model parameterization. This ensures that the same structure is used to capture and reconstruct the subject. The images in Figure 4.10 illustrate the effect of subdivision on the control model. The control model shown on the left in image (a) is highly deformed from its original state (a natural-looking face) to emphasize the effect. Without subdivision, the edges of the control model appear very clearly in the reconstruction shown to its right. The model on the left in image (b) is produced by subdividing the control model twice. The reconstruction depicted on the right shows that the polygon edges are less visible because the control surface is much smoother than in (a).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(a)

(b)

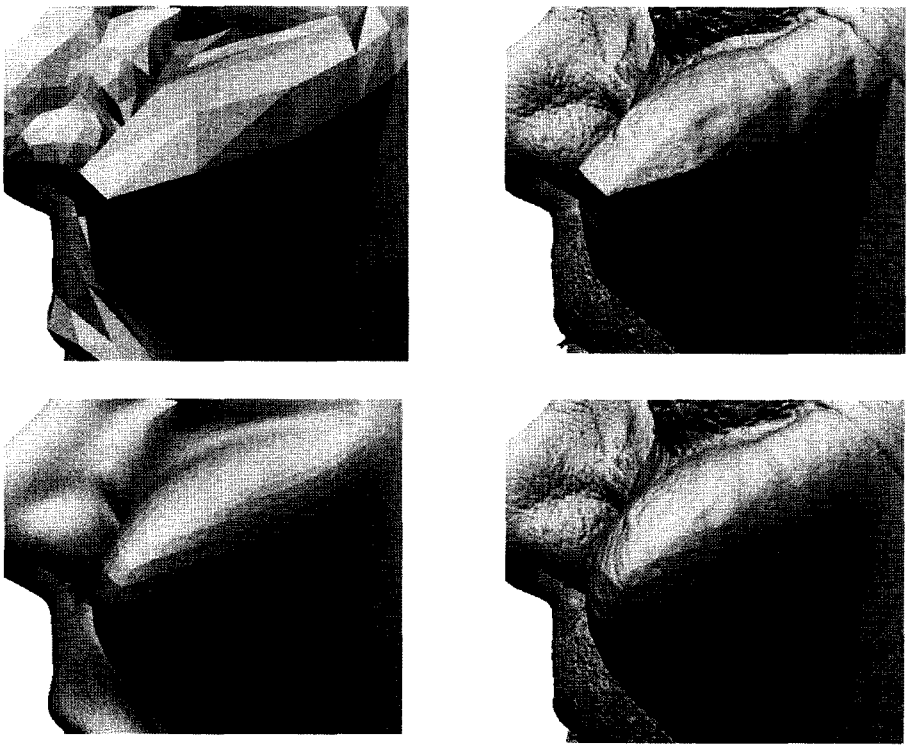


Figure 4.10: Examples of reconstructing a subject model on the same control model. The control model was greatly altered from its original state. Left to right: control model, reconstructed model. (a) Non-subdivided control model. (b) Subdivided the control model twice.

Unparameterized vertices (refer to Section 4.3.1 for the process used to establish parameters) pose a problem for reconstruction because without a proper mapping, a vertex cannot be reproduced properly and neither can the immediate area surrounding the vertex. One of two actions can be taken to handle unparameterized vertices. They and their adjacent polygons can be intentionally removed altogether from the reconstruction or some special technique can be employed to attempt to remedy the issue.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

First, consider the case in which the vertices and their adjacent polygons are intentionally removed, thereby producing holes. The reasoning for this decision is that it is better to produce a hole than it is to reproduce the surface incorrectly. The presence of holes in the reconstructed model is undesirable, but a designer may feel more comfortable using commercial applications like Maya and 3DS Max to patch the holes. An example of the routine used to remove polygons is shown in Figure 4.11. The image in (a) shows the original subject model and one of its vertices V. Suppose that V could not be parameterized and its adjacent polygons (marked with dashed edges) are to be removed. Figure 4.11(b) shows the reconstructed subject model with point V and the identified polygons removed.

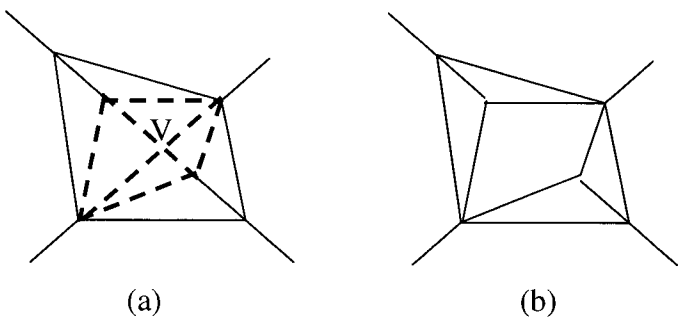


Figure 4.11: Example of intentionally removing polygons adjacent to an unmapped vertex V. (a) Original subject model. Adjacent polygons are marked with dashed edges. (b) Reconstructed subject model. V and adjacent polygons are omitted.

Now consider the case where a technique is employed to resolve the issue. One possible way to solve the problem is to use RBF networks to interpolate new positions of vertices that could not be parameterized. During model parameterization, these vertices can be marked and the following steps can be taken for each marked vertex during reconstruction.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1. Find all neighbouring vertices in the original subject model that are both: a) within a certain (perhaps user-definable) radius from the marked vertex, and b) not marked.
2. Prepare RBF networks (as described in Section 3.2.2) using the neighbours as centers. The initial positions of these centers are given by their positions in the original subject model. The result is that the networks approximate the original surface surrounding the marked vertex.
3. Set the RBF networks' weight vectors using the new positions of the centers obtained by performing model reconstruction. The networks now approximate the reconstructed surface surrounding the marked vertex.
4. Evaluate the RBF networks using the marked vertex as the input point.

(a) (b)

Figure 4.12: Comparison of ways to address unparameterized vertices. (a) Reconstruction with unparameterized vertices and adjacent polygons intentionally removed. (b) Reconstruction achieved with RBF interpolation.

Figure 4.12 contrasts the results obtained using each of the two courses of action described in the preceding paragraphs. The reconstructions shown in this thesis used the first course of action to demonstrate that holes are typically quite small, they occur

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

infrequently and they are almost always located in specific areas of the face. The error analysis section that follows gives some statistics to support these claims.

4.5 Error Analysis

The accuracy of the model parameterization scheme is evaluated by analyzing the error resulting from model reconstruction. Because 3D models can be visualized in a meaningful and easily understandable way, reconstruction error can be stated by judging the quality of the reconstructed surface (e.g., its resemblance to the original and the presence of holes). Reconstruction error may also be quantified by calculating the number of vertices and triangles that are missing in the reconstructed model.

The goal of the error analysis described here is to apply the aforementioned error measurement methods to reconstructions of an original (i.e., unmodified) subject model. Person B's scanned face was chosen from the database of scanned faces (Section 3.4) as the subject model to be reconstructed and some of its relevant statistics are given in Table 4.1. Two test scenarios were considered. In the first scenario, Person B's scanned face was reconstructed using two control models (of differing resolutions) created by Vizup[3]. In the second scenario, Person B's preliminary medium resolution and preliminary high resolution models served the roles of control models for reconstruction. No changes were made to the control models in either scenario, but the control models in the first scenario

[3] Vizup (http://www.vizup.com) is a shareware program that uses model simplification to construct a low resolution model from a high resolution model.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(also referred to as simplified models) were subdivided twice prior to parameterization and reconstruction to simulate actual usage in applications. Note that the preliminary models in the second case were already subdivided when they were created. The same set of settings (i.e., thresholds and ranges described in Section 4.3.1) was used to parameterize the scanned model in the four tests. These values are given in Table 4.2.

Table 4.3summarizes the control models' statistics and the test results that were obtained on an AMD Opteron 252 processor. Generally, the original scanned model was well reconstructed in each of the four cases. The second simplified model produced the lowest percentage of missing vertices and triangles (about 0.06%) while the preliminary medium resolution model produced the highest (about 2.2%). Of the two test scenarios, the first produced the better reconstructed models in terms of the percentage of missing vertices and triangles. In each test scenario, the higher resolution control model captured more vertices than its lower resolution counterpart, thus producing a reconstructed model with a higher number of triangles. This was due to the higher resolution model's closer approximation of the scanned face. Execution times for model parameterization ranged from 11 s for the smallest control model to 277 s for the largest. The time required to reconstruct the scanned model was fairly constant at about 2 s for all cases.

The images shown in Figure 4.13 are for (a) the simplified model with 32,016 triangles and (b) the preliminary medium resolution model test cases. Each row of images shows the control model, the reconstructed model and the reconstructed models placed on top of the original scanned face, respectively. In the last image of each row, the original scanned surface is shown in red and the reconstructed surface is shown in green. The reconstructed model in each case resembles the original as expected. There are no

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

missing triangles in the face area of the reconstructions obtained using the simplified models, but the reconstructions produced by the preliminary models have holes in the eyes and mouth areas. These holes are very small and thus are not clearly visible in the images in (b).

Vertices	1,081,676
Triangles	2,160,470
Bounding Box	X: [-0.4106, 0.4030] Y: [-0.7676, 0.5944] Z: [0.2215, 0.8207]
Bounding Box Diagonal	1.6957

Table 4.1: Statistics for Person B's scanned model.

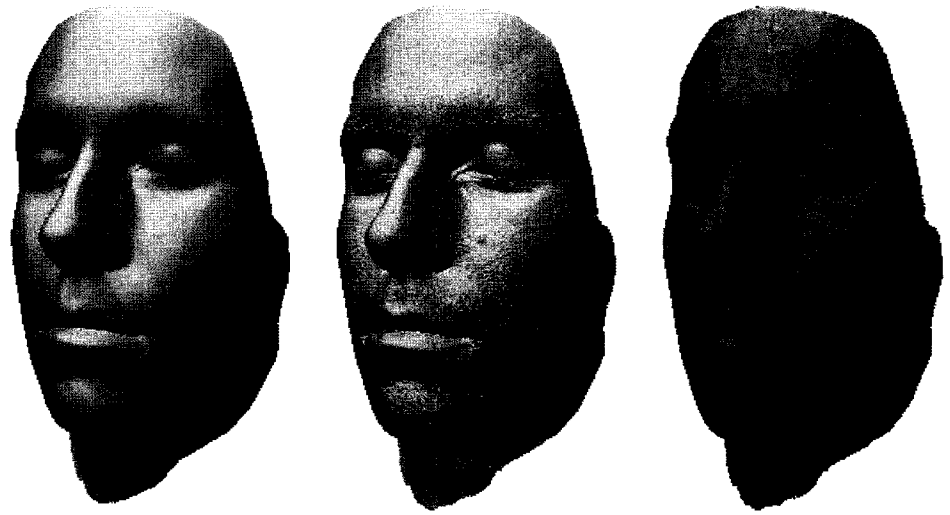
Bounding Box Expansion Parameters	0.015
Acceptable Barycentric Coordinates	[-0.1 1.1]
Displacement Threshold	-
Error Threshold	0.01

Table 4.2: Model parameterization settings.

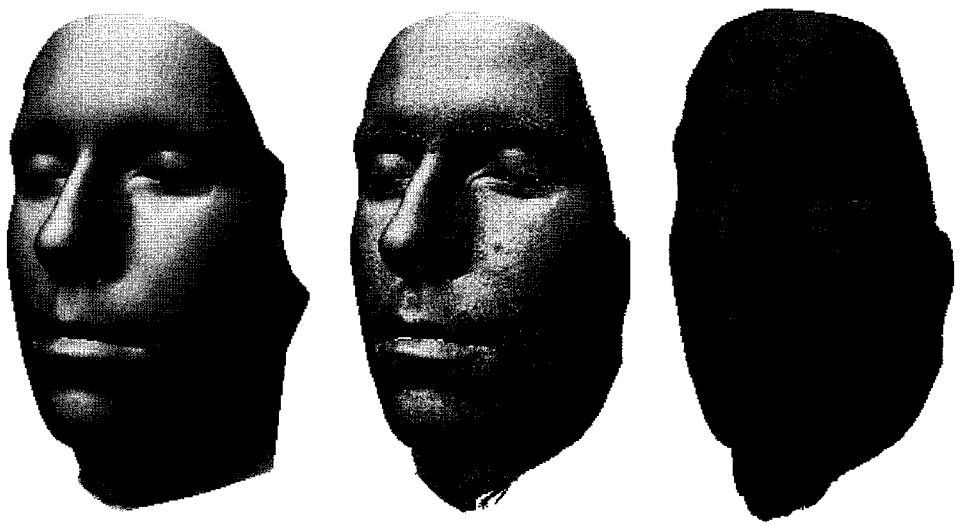
Control Model	Subdivided Simplified #1	Subdivided Simplified #2	Preliminary Medium	Preliminary High
Resolution (triangles)	32,016	80,048	55,232	883,712
Reconstructed Subject Model (Vertices/Triangles)	1,079,779 / 2,156,443	1,081,094 / 2,159,119	1,057,955 / 2,111,525	1,063,032 / 2,122,076
Missing Vertices (%)	0.175376	0.053805	2.192986	1.723621
Missing Triangles (%)	0.186395	0.062532	2.265479	1.777113
Parameterization Execution Time (s)	11	23	20	277
Reconstruction Execution Time	2	2	2	2

Table 4.3: Error statistics for reconstructing the scanned model of Person B.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.



(a)



(b)

Figure 4.13: Sample results from error analysis. From left to right for each row of images: control model, reconstructed model, reconstructed model (green) placed on top of original model (red). (a) Simplified model case (32,016 triangles). (b) Preliminary medium resolution model case (55,232 triangles).

On a related note, Table 4.4 lists the error statistics collected for each of the four scanned faces in the database. Each scanned face was reconstructed using its

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

corresponding preliminary medium resolution model consisting of 55,232 triangles. Person C's reconstruction achieved the smallest error (about 1.7%) while Person A accounted for the largest (about 6.0%). As with the reconstructed scanned models shown above, these reconstructed models also had only a few small holes in the face area. In the case of Person A, there were also some holes in the neck area which contributed to the error.

	Person A	**Person B**	**Person C**	**Person D**
Reconstructed Subject Model (Vertices/Triangles)	47,1463 / 937,761	1,057,955 / 2,111,525	1,490,096 / 2,962,853	741,929 / 1,479,937
Missing Vertices (%)	5.965243	2.192986	1.604369	1.975864
Missing Triangles (%)	5.944084	2.265479	1.654530	2.088644

Table 4.4: Error statistics for scanned model reconstruction using preliminary medium resolution models.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 5 Applications

This chapter presents three applications for the multi-resolution modeling scheme presented in Chapter 3 and Chapter 4. These applications demonstrate the primary advantages of the scheme, namely the ability to alter shape and skin detail separately from one another, the ability to combine unaltered shape and detail elements to form higher resolution models and the ability to reconstruct high resolution geometry using 3D skin. Section 5.1 presents the concept of 3D skin transfer in which one person's high resolution 3D skin is applied onto another person's face. The second application, discussed in Section 5.2, is multi-resolution shape exaggeration, i.e., exaggeration of the models in the multi-resolution hierarchy. Multi-resolution skin detail exaggeration is also possible and the concept is outlined in Section 5.3.

5.1 3D Skin Transfer

Although 3D skin is not structurally consistent due to the arbitrary nature of scanned model resolution and topology, it is defined on the models of the hierarchy which are homogeneously structured. Since model reconstruction using 3D skin relies on the topology of the control model and not the identity of the person represented by the control model, the homogeneous structure can be exploited to perform 3D skin transfer (or simply skin transfer) whereby 3D skin is applied on any face.

An interesting use for skin transfer is to simulate aging and rejuvenation where a face model receives skin that makes it appear older or younger. This has applications in the field of medicine where doctors or surgeons can show the results of surgical skin procedures such as transplants to patients prior to performing them.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Another possible use for skin transfer is to increase the resolution of models. When acquiring highly accurate 3D models of faces, the capture process can be very time-consuming and expensive (e.g., extremely accurate laser scanning). However, "skin recycling" allows a larger number of highly detailed 3D faces to be constructed by transferring existing skin onto new face shapes. Basic shape capture techniques that are faster and cheaper can be used to construct low resolution models in the multi-resolution scheme, which can then be subjected to skin transfer to obtain new high resolution models. Although the skin does not belong to the faces being modeled, this may be adequate for some applications and it is a cheaper way to obtain face models with a high level of skin detail.

5.1.1 Simulation of Aging and Rejuvenation

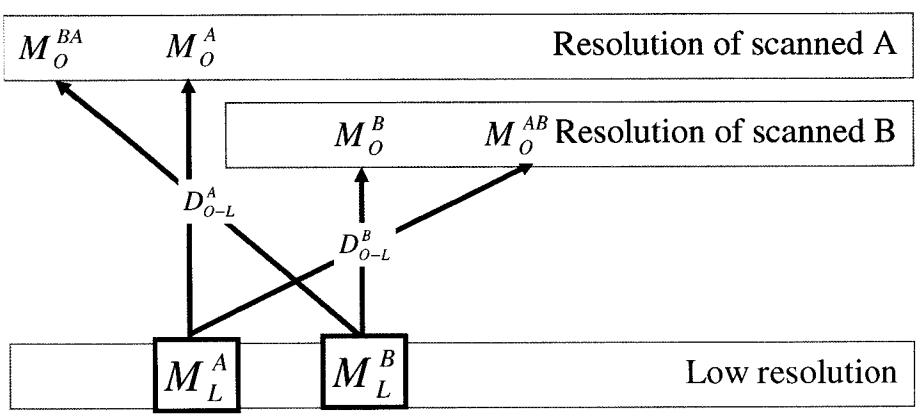


Figure 5.1: 3D skin transfer using low resolution shape. 3D skin is swapped between Persons A and B to create two new models.

Consider the examples of applying Person A's 3D skin onto Person B's low resolution model and vice-versa. Following the discussion in Section 3.3.4, a non-rigorous expression for Person A's scanned face is

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$M_O^A = M_L^A +_{pseudo} D_{O-L}^A, \tag{5.1}$$

noting that M_L^A should be subdivided appropriately. The expression for transferring Person B's 3D skin onto Person A's low resolution model might then be written as

$$M_O^{AB} = M_L^A +_{pseudo} D_{O-L}^B. \tag{5.2}$$

This particular example of 3D skin transfer is illustrated in Figure 5.1, which also indicates the final resolution of the models post skin transfer. The number of triangles in M_O^{AB} depends on the number of mapping parameters held by D_{O-L}^B, which is at most equal to the number of vertices in M_O^B.

The models M_O^{AB} and M_O^{BA} are shown in the top-right and bottom-right images, respectively, in Figure 5.2 and compared to the original scanned models of each person. As expected, Person A (a senior Caucasian male) looks younger with the new skin while Person B (a young Caucasian male) appears to have aged because of the transfer. This result mirrors what one might expect if this type of skin transplantation occurred in real life. The skin retains its original properties, but the skin would conform to the shape of the transplant recipient's face.

Figure 5.3 shows another example of 3D skin transfer in which the recipient's face is exaggerated prior to the transfer. The original scanned model of Person C (young Asian male) is shown in image (a). Person C's low resolution model is exaggerated by a factor of 60% and his own 3D skin is used to construct the model shown in image (b). Person A's 3D skin is also applied onto the exaggerated low resolution model to produce the model M_O^{CA} shown in images (c) and (d).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

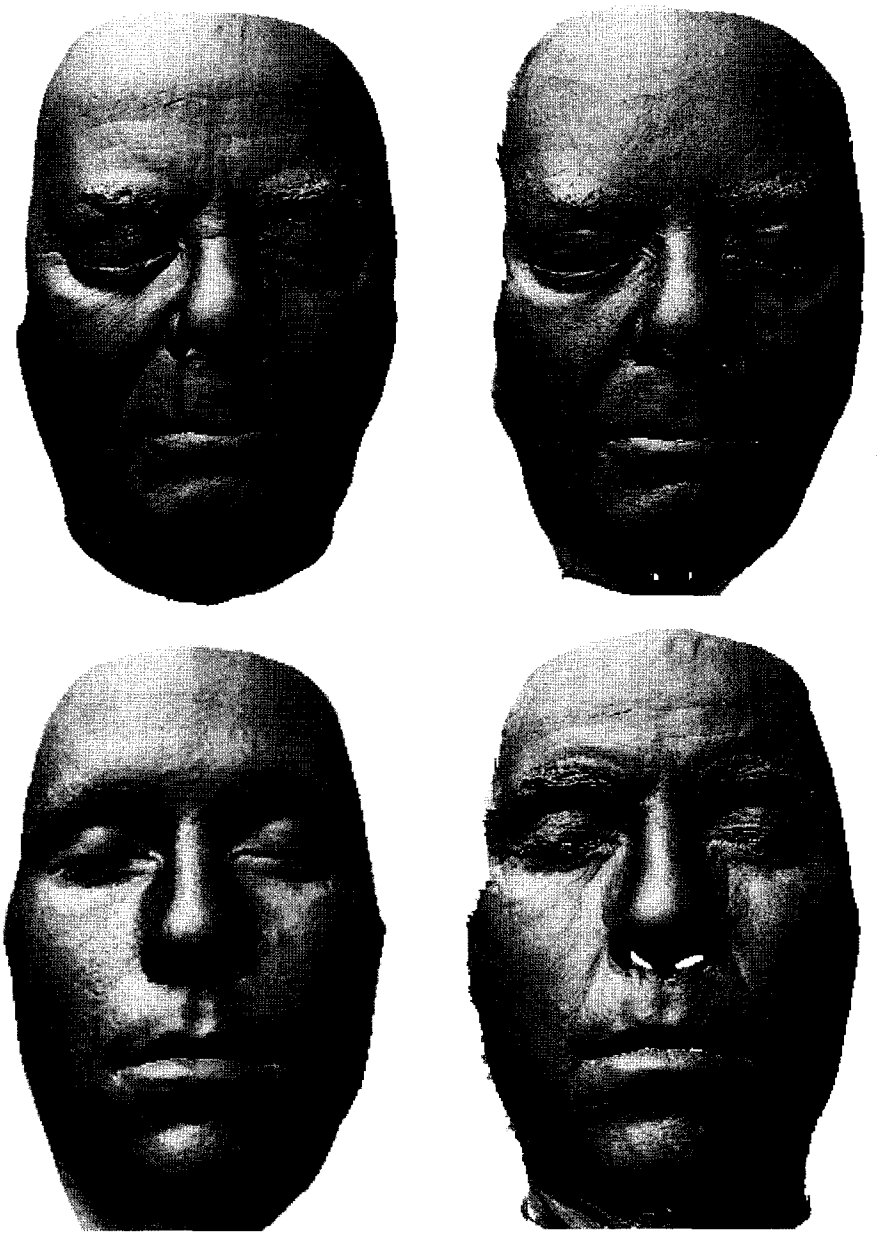


Figure 5.2: Example of aging/rejuvenation simulation using 3D skin transfer. 3D skin is transferred between Persons A and B. Low resolution shape is used as the base models for the transfer. Top-left: Original Person A. Top-right: Person B's 3D skin applied to Person A. Bottom-left: Original Person B. Bottom-right: Person A's 3D skin applied to Person B.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

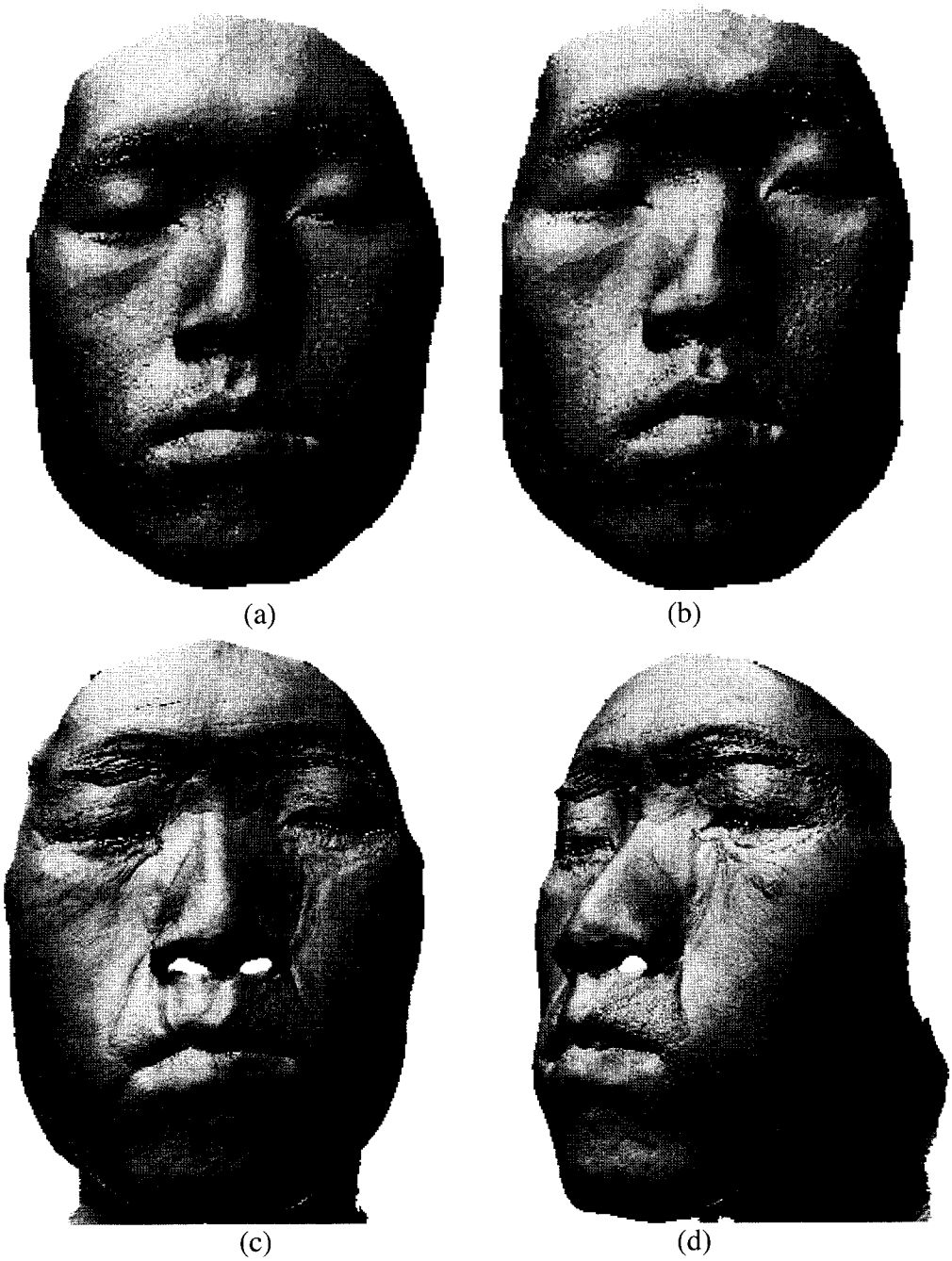


Figure 5.3: 3D skin transfer onto exaggerated low resolution shape. (a) Original scanned face of Person C. (b) 60% exaggerated reconstruction using 3D skin of Person C (young Asian male). (c),(d) 60% exaggerated reconstruction using 3D skin of Person A (senior Caucasian male).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5.1.2 Augmentation of Mesh Resolution

There is a large database of 3D human bodies for the Civilian American and European Surface Anthropometry Resource Project (CAESAR). The human (male and female) subjects being represented by the models are between the ages of 18 and 65. These subjects were scanned very quickly, but the capture process still produced models which capture the global shapes of the subjects. Although these models provide valuable information such as measurements for various body parts, an up-close inspection of their faces reveals that they are not suitable for visualization, computer animation or any other applications in which visual appearance is important.

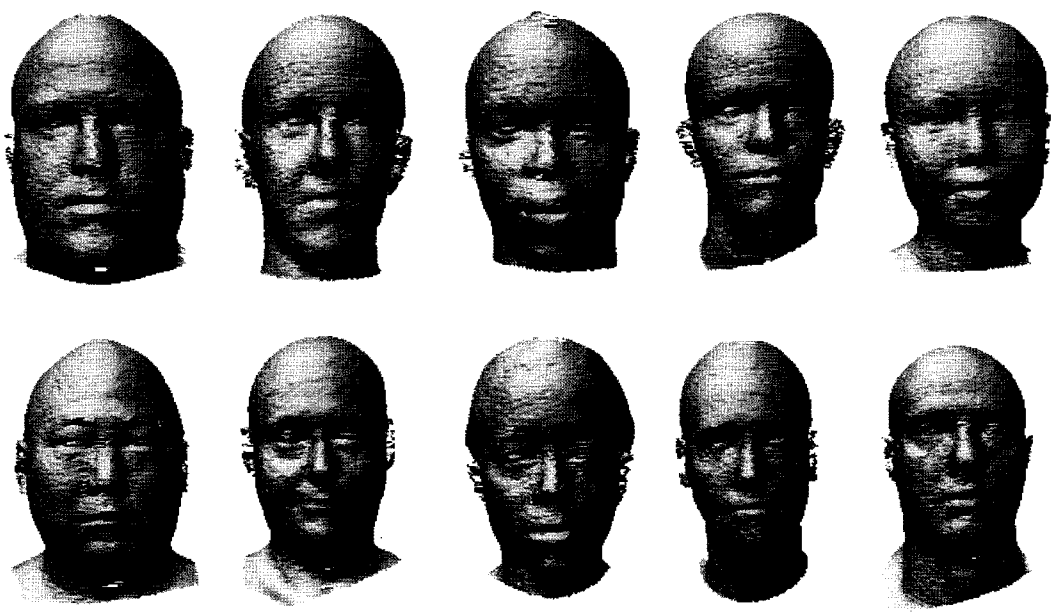


Figure 5.4: Ten heads obtained from the CAESAR database.

Figure 5.4 shows ten heads obtained from the CAESAR dataset. These heads were compiled by removing the body portions of the ten available models. Each head consists of about 20,000 triangles. The general shape of the faces is good, but horizontal sliding lines are quite visible and there are many missing points in the face region for

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

some faces. Additionally, the ears and the back of the head are incomplete to some degree for every person.

(a)

(b)

Figure 5.5: Construction of low resolution models from poorly scanned heads. (a) Three CAESAR head models. (b) Corresponding low resolution models.

These models were used as the scanned surfaces to produce a low resolution model for each of the subjects. Figure 5.5(a) shows three different CAESAR head models and Figure 5.5(b) shows their corresponding low resolution models. Note that the low resolution models still only capture the face area despite the availability of head data.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Because the existing database of high resolution models still only contains face models, no real purpose is served by making the change to full head low resolution models. Figure 5.6(a) shows the original scanned model of the 3D skin transfer recipient while Figure 5.6(b) and (c) show the results of applying the 3D skin of Person A and Person C onto the low resolution model of the recipient.

(a) (b) (c)

Figure 5.6: Examples of mesh resolution augmentation using 3D skin transfer. (a) Original scanned model. (b) Reconstruction using Person A's 3D skin. (c) Reconstruction using Person C's 3D skin.

5.2 Multi-resolution Shape Exaggeration

Exaggeration is one of the key principles of animation [Lasseter 87] and is commonly used in art and entertainment. A character is made more attractive and distinctive by exaggerating its shape, colour, emotion, or actions. The multi-resolution representation of scanned data discussed in Chapter 3 makes it possible to perform shape exaggeration at any resolution (from low to high) independently from skin detail. The

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

exaggeration algorithm, which was originally proposed by Brennan [Brennan 85], automatically identifies distinctive features based on a comparison with a center (i.e., average) face. The computed difference between the subject face and the center face are scaled in a vector-based manner and then used to derive the exaggerated model at the desired resolution. 3D skin can be applied to obtain an exaggerated version of the original scanned face.

5.2.1 Multi-resolution Center Face

The purpose of calculating a center (average) face model is to obtain some measure of an average human face which is also void of individual characteristics. Individual faces are exaggerated by making comparisons with the center face to determine which features are distinctive (i.e., those that deviate from the center). The true average human face can only be calculated by considering the world's population in its entirety. This is clearly not feasible. Hence, a limited number of faces is used to obtain the average.

A center face M^{center} is defined as the average of a set of F subject faces $\{M^1, M^2, \ldots, M^F\}$ of a common resolution. The center face is endowed with the average of the facial features found in the set. The averaging operation occurs at the vertex level, exploiting the homogeneous structure found in every model. The average of the models is formulated as

$$M^{center} = \frac{1}{F}\sum_{i=1}^{F} M^i = \left\{ \frac{1}{F}\sum_{i=1}^{F} v_j^i, j = 1,\ldots,N \right\}, \tag{5.3}$$

where v_j^i is the jth vertex in M^i and N is the number of vertices in each model. The center face concept is illustrated in Figure 5.7.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

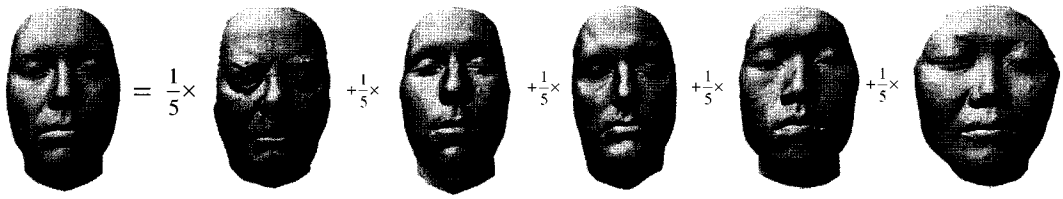
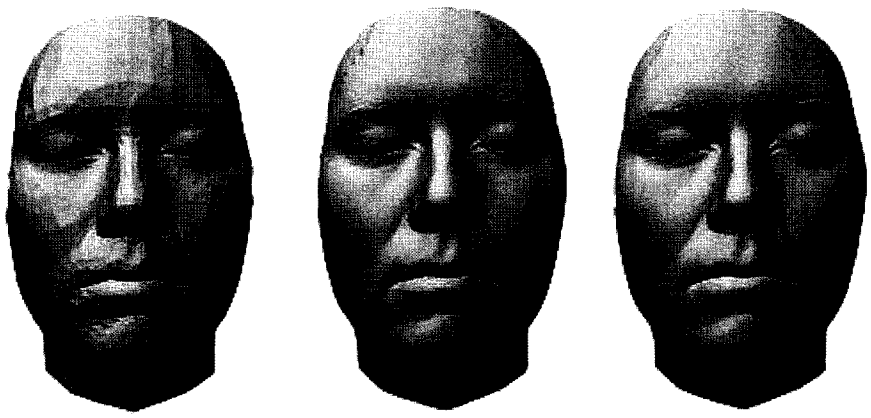


Figure 5.7: Formulation of a center face model.

Multi-resolution center faces are needed before multi-resolution shape exaggeration can be performed. A center face must be derived for each of the three resolutions from a group of subject models. An example of center faces produced from the multi-resolution representations of two subjects' scanned faces is shown in Figure 5.8.



(a) (b) (c)

Figure 5.8: Multi-resolution average face models produced from two subject faces.

(a) Low. (b) Medium. (c) High.

It may be desirable to construct multiple center faces at the same resolution for specific demographics, e.g., males, females, Caucasians, etc. This strategy can prevent undesired features from contributing to the center faces.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Some may consider that using a small set of faces to construct an average face is inaccurate and thus produces poor exaggerations. Although not optimal, a small number of faces may still accomplish the desired goal of removing individual characteristics. Figure 5.9 shows examples of average faces produced using two, three, four and five faces from previous work [Soon 06 "Shape"]. The two-face and three-face averages contain some features inherited from their respective sets. The four-face and five-face averages are quite similar and do not bear significant resemblance to any of their respective constituents.

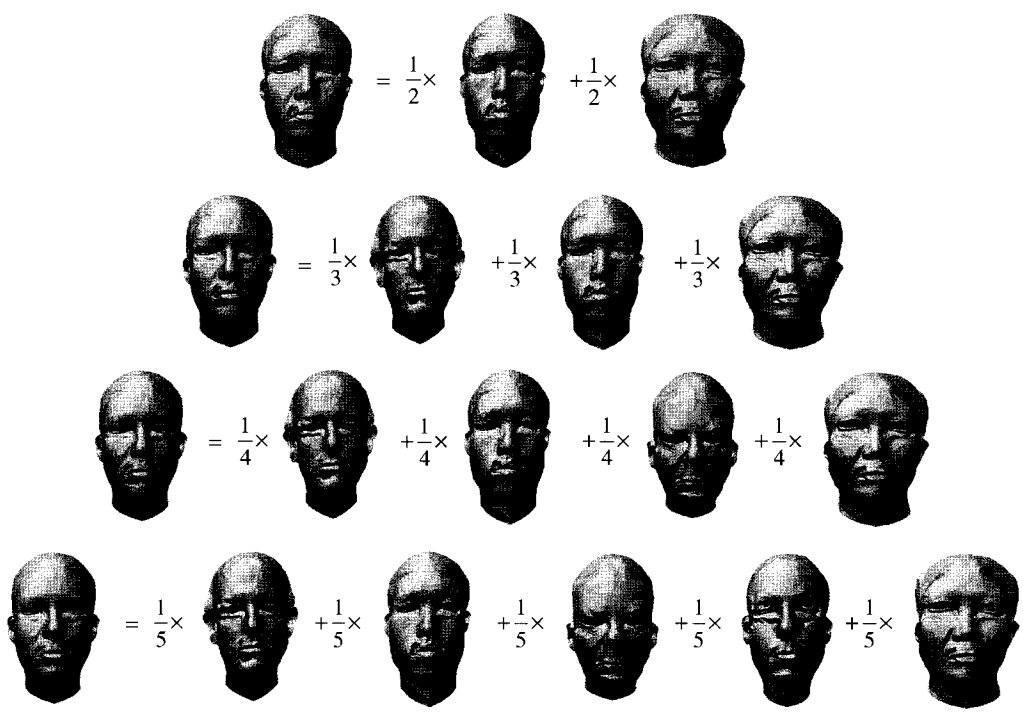


Figure 5.9: Examples of average face models constructed from varying numbers of faces.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5.2.2 Exaggeration using Vector-based Scaling

The same exaggeration algorithm is used regardless of the center face model's resolution. The algorithm is vector-based and begins with a calculation of the difference between a subject face and the center face of the same resolution. This difference calculation yields a set of feature vectors $\{\vec{v}_1,...,\vec{v}_N\}$ where each feature vector $\vec{v}_i$ is defined as

$$\vec{v}_i = v_i^{person} - v_i^{center}, \tag{5.4}$$

where v_i^{person} is point i in the subject model and v_i^{center} is point i in the center face. Each feature vector is scaled by a constant exaggeration factor ef to derive the point $v_i^{exaggeratedPerson}$ in the exaggerated model, namely

$$v_i^{exaggeratedPerson} = v_i^{center} + (1+ef)\vec{v}_i. \tag{5.5}$$

It then follows that

$$M^{exaggeratedPerson} = \{v_i^{exaggeratedPerson}, i = 1,...,N\}. \tag{5.6}$$

Using the $-_{pseudo}$ symbol from Equation (3.11) and a similarly defined $+_{pseudo}$,

$$M^{exaggeratedPerson} = M^{person} +_{pseudo} ef(M^{person} -_{pseudo} M^{center}). \tag{5.7}$$

5.2.3 Low Resolution Shape Exaggeration

Low resolution shape exaggeration refers to the process of exaggerating a low resolution model which captures only the low level shape of a face. In many cases, the person performing the exaggeration does not wish to exaggerate skin detail, so using higher resolution models for the work is wasteful. It may also be necessary to visualize the results of the exaggeration in a timely manner, making it even more important to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

perform as few calculations as possible so as to expedite the process. With a triangle count of only roughly 3,500, the low resolution model is a suitable choice.

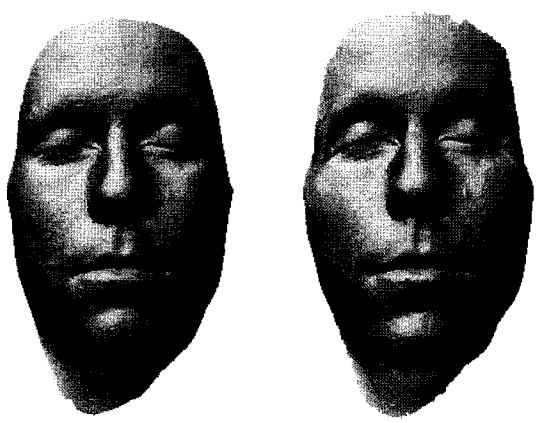
The discussion from Chapter 5.2.2 can be contextualized such that Equation (5.7) becomes

$$M_L^{exaggeratedPerson} = M_L^{person} +_{pseudo} ef(M_L^{person} -_{pseudo} M_L^{center}). \tag{5.8}$$

The exaggerated low resolution model $M_L^{exaggeratedPerson}$ can then be used to produce an exaggerated version of the scanned model $M_{LO}^{exaggeratedPerson}$ by using 3D skin and model reconstruction (refer to Section 4.4). The procedure can be written as

$$M_{LO}^{exaggeratedPerson} = M_L^{exaggeratedPerson} +_{pseudo} D_{O-L}, \tag{5.9}$$

where the notation $+_{pseudo}$ is loosely used to represent model reconstruction. [Recall from the discussion in Chapter 4 that the model used for parameterization and reconstruction must be a smooth surface to ensure good results are achieved.]



(a) (b)

Figure 5.10: Example of using low resolution shape exaggeration and 3D skin recovery to produce an exaggerated version of Person B's scanned face. (a) Original model. (b) 75% exaggerated reconstruction.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

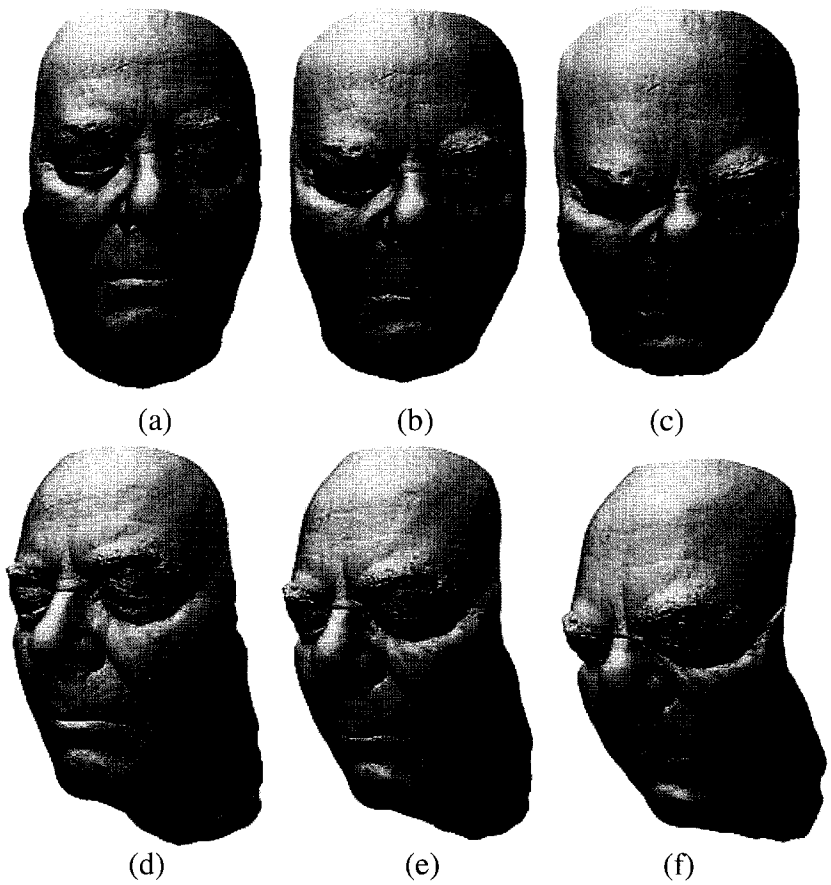


Figure 5.11: Examples of using low resolution shape exaggeration and 3D skin recovery to produce exaggerated versions of Person A's scanned face. (a),(d) Original scanned model. (b),(e) 60% exaggerated reconstruction. (c),(f) 120% exaggerated reconstruction.

The images in Figure 5.10 and Figure 5.11 show some results obtained using the 3D skin recovery idea. Figure 5.10 compares Person B's original scanned face with a 75% exaggerated reconstruction. Figure 5.11 compares Person A's original scanned face with 60% and 120% exaggerated versions of it.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5.2.4 Medium and High Resolution Shape Exaggeration

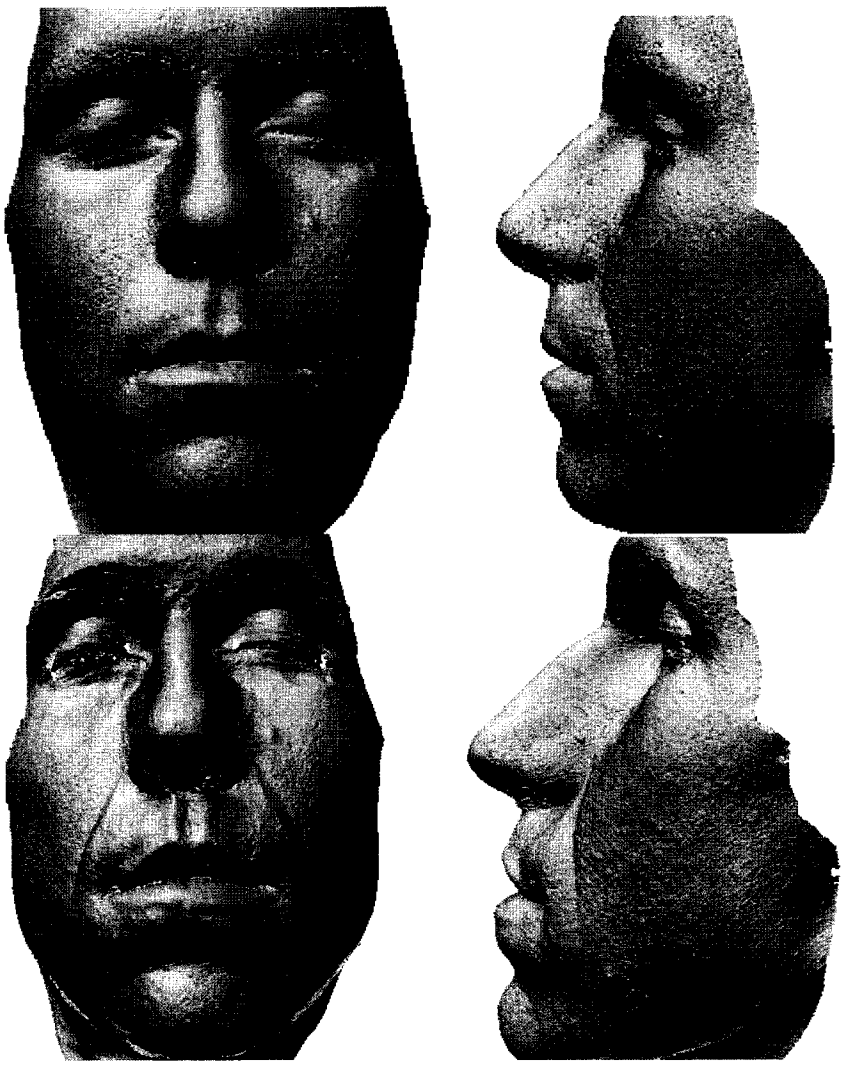


Figure 5.12: Example of using medium resolution shape exaggeration and 3D skin recovery to produce an exaggerated version of Person B's scanned face. Top row: Original scanned model. Bottom row: 100% exaggerated model reconstruction.

The vector-based exaggeration approach may also be applied to the medium and high resolution models to accomplish medium and high resolution shape exaggeration, respectively. Here, exaggeration affects not only the shape, but also some of the skin detail as both medium and high resolution shapes contain skin detail. The remainder of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the discussion of exaggeration in this section focuses on medium resolution shape, but the extension to high resolution shape is straightforward.

Like low resolution shape exaggeration, medium resolution shape exaggeration can be paired with 3D skin to obtain an exaggerated version of a scanned model. For this type of exaggeration, Equation (5.9) becomes

$$M_{MO}^{exaggeratedPerson} = M_{M}^{exaggeratedPerson} +_{pseudo} D_{O-M} \cdot \tag{5.10}$$

Figure 5.12 shows an example of medium resolution shape exaggeration. The top row of images show Person B's original scanned model and the bottom row of images depict the face exaggerated by 100%. Skin detail is also exaggerated as a byproduct because medium resolution shape contains some skin detail.

5.3 Multi-resolution Detail Exaggeration

The aim of detail exaggeration is to only alter the skin-level detail of a model without modifying the global shape. The discussion in this section demonstrates how a high resolution model is recomposed from low resolution shape and exaggerated D_{H-L}, but the formulation for other detail levels and models follows the same process and can be easily obtained.

5.3.1 Multi-resolution Center Detail

Like shape exaggeration, the first step is to define some notion of the center or average detail. The center detail is calculated from a set of common-resolution skin detail instances D. Technically speaking, detail is a set of offset vectors (see Section 3.3.3 for the definition). However, to facilitate discussion, detail will be represented by only one

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

vector from the set since descriptions of sets of vectors can be cumbersome and harder to comprehend.

As a vector, detail D can then be decomposed into two components: a direction component UD and a magnitude $\|D\|$ such that $D = \|D\| UD$. The direction component is a unit vector defined as

$$UD = \frac{D}{\|D\|}. \tag{5.11}$$

Detail is decomposed in this manner so that the magnitude component can be manipulated during exaggeration without affecting the direction component. Unlike an individual person's detail, the center detail, D^{center}, is a scalar which is calculated as

$$D^{center} = \frac{1}{K}\sum_{k=1}^{K}\|D^k\|, \tag{5.12}$$

where D^k is the detail for person k and K is the number of instances used to calculate the average.

5.3.2 Exaggeration using Scalar-based Scaling

Unlike shape exaggeration, skin detail exaggeration is achieved by scalar adjustment rather than vector adjustment. The reason for this is that skin detail is defined as the result of performing shape refinement using cylindrical projection as stated in Equation (3.10). Figure 5.13 is used as reference for the proceeding text. Shape refinement, as discussed in Section 3.3.2, operates by extending projection vectors that pass through vertices in the preliminary model (e.g., P) and finding the intersection of these vectors with the scanned surface (e.g., Q). Thus, the act of adding skin detail is

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

simply to displace each vertex along its corresponding projection vector to the intersection point. The way to exaggerate skin detail then is to either increase or decrease the effect of shape refinement depending on the average, i.e., displace each vertex along the offset direction. Since the offset vectors calculated for skin detail are in the same direction as the projection vectors, exaggerated detail can be computed by changing the magnitude of the offset vectors. In Figure 5.13, the exaggerated detail vector originates at P and terminates at R.

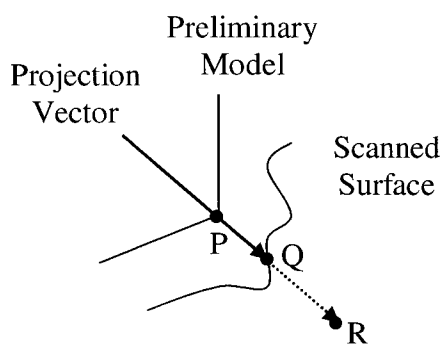


Figure 5.13: Illustration of skin detail exaggeration.

Therefore, the expression for skin detail exaggeration is

$$D^{exaggeratedPerson} = D^{person} + ef\left(\left\|D^{person}\right\| - D^{center}\right)UD^{person}, \tag{5.13}$$

where the + operator indicates vector addition and *ef* is a constant exaggeration factor. Then,

$$M^{exaggeratedPerson} = M^{person} +_{pseudo} D^{exaggeratedPerson}, \tag{5.14}$$

where the $+_{pseudo}$ operator is used in the same sense as in Equation (3.12), $M^{exaggeratedPerson}$ is endowed with the exaggerated detail and M^{person} is subdivided appropriately according to Table 3.2 so that the operation is sensible. The meaning of Equation (5.14) is to apply

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the exaggerated detail $D^{exaggeratedPerson}$ (offset vectors) to the appropriate vertices of M^{person} to arrive at the model $M^{exaggeratedPerson}$ with exaggerated skin detail. In other words, shape is combined with exaggerated skin detail to construct a higher resolution model.

5.3.3 High-Low Skin Detail Exaggeration

For high-low skin detail exaggeration, Equations (5.13) and (5.14) become

$$D_{H-L}^{exaggeratedPerson} = D_{H-L}^{person} + ef\left(\left\|D_{H-L}^{person}\right\| - D_{H-L}^{center}\right)UD_{H-L}^{person} \tag{5.15}$$

and

$$M_{H}^{exaggeratedPerson} = M_{L(S4)}^{person} +_{pseudo} D_{H-L}^{exaggeratedPerson} \tag{5.16}$$

respectively, where the *(S4)* subscript indicates that the low resolution model M_{L}^{person} is subdivided four times prior to adding $D_{H-L}^{exaggeratedPerson}$, i.e., $M_{L(S4)}^{person}$ denotes the preliminary high resolution model.

Figure 5.14 and Figure 5.15 show some examples of D_{H-L} detail exaggeration using several different exaggeration factors and with different low-resolution models acting as M_{L}^{person}. Figure 5.14(a) and Figure 5.14(b) show the result of exaggerating Person A's D_{H-L} by 30% and 90%, respectively while Figure 5.15(a) and Figure 5.15(b) show the same type of exaggeration performed on Person B for factors of 60% and 90%, respectively.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(a) (b)

Figure 5.14: Example of using low resolution shape and exaggerated high-low skin detail to recompose an exaggerated high resolution model of Person A. (a) 30% exaggeration. (b) 90% exaggeration.

(a) (b)

Figure 5.15: Example of using low resolution shape and exaggerated high-low skin detail to recompose an exaggerated high resolution model of Person B. (a) 60% exaggeration. (b) 120% exaggeration.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter 6 Conclusions

The preceding chapters present a multi-resolution hierarchy for representing extremely highly detailed 3D scanned faces. The presented hierarchy consists of three levels of shape and detail information. With this representation, shape and detail can be acted upon separately from one another. Additionally, it includes encodings of the original scanned surface's vertices which can be used to reconstruct said surface's geometry using any shape resolution in the hierarchy. Each resolution in the hierarchy is equipped with its own standardized structure which is imbedded into every face. This homogeneous structure establishes correspondence between models of a common resolution and allows geometry operations affecting shape and detail to be defined in terms of the structure rather than an individual face. Example applications that utilize the proposed representation are also given.

This chapter reports the key contributions of the research documented in the preceding chapters. Section 6.1 lists the contributions to new knowledge made by this thesis. Section 6.2 discusses the generality, the extendibility and the limitations of the proposed multi-resolution scheme. Section 6.3 outlines areas of future research.

6.1 Contributions

The following contributions to new knowledge were made as a result of performing the research described in this thesis:

1. Developed a scheme for representing extremely high resolution 3D scanned faces in several resolutions of structurally-consistent shape and detail, allowing shape-based operations to take place separately from detail-based

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

operations. Lower order shape and detail components can also be used to recompose higher order shapes.

2. Proved that the adjacency information of the scanned model is critical to its appearance by demonstrating that displacement mapping does not produce a faithful reconstruction of the original scanned surface.
3. Demonstrated the use of model parameterization using point-to-surface mapping to reconstruct geometry.
4. Proposed a technique to augment a model's resolution using the presented multi-resolution idea in conjunction with 3D skin transfer.

6.2 Discussion

The multi-resolution modeling scheme described in this thesis allows an extremely high resolution 3D scanned face to be represented as a hierarchy of shape and details. General-purpose and face-specific techniques are employed to establish the hierarchy. Feature-point correspondence described in Section 3.2.1 is specialized for human faces in the sense that the feature points that are used were carefully defined for the purpose of identifying an individual. The use of cylindrical projection for shape refinement discussed in Section 3.3.2 is often used for human heads because of their somewhat cylindrical shape. [The human head, however, is not fully cylindrical and a limitation of using this type of projection is discussed later in this section.] Surface approximation using RBF networks (Section 3.2.2), multi-resolution shape construction (Section 3.3.1) and model parameterization and reconstruction (Sections 4.3 and 4.4) are general-purpose ideas and techniques that either can be or have already been used for other classes of objects.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The exaggeration method described in Section 5.2 revolves around a center face that is used to identify and strengthen distinctive characteristics of particular individuals. This procedure is mechanical and its results do not have the same artistic touch as the work of a professional caricaturist. The topic of producing artistic 3D caricatures has not been explored to date. However, there are many documented methods for producing artistic (2D) caricature images by analyzing the existing work of a professional. Some of these known methods are mentioned in Section 2.5. Thus, further work is needed to extend the knowledge developed for images to the field of 3D models before stylized 3D exaggerations can be produced.

Generally speaking, the multi-resolution models can be animated as long as the generic model from which they are derived is properly equipped to support animation. Works like [Lee 93] and [Lee 95] show the use of a physics-based generic model to generate individualized faces for animation whereas [Lee 97], [Lee 98] and [Lee W 00] make use of an animation structure embedded in the generic model. Live performances by subjects can also be recorded using video or marker-based motion capture for data-driven facial animation. The scanned models and the generic model, in their present state, impose some limitations on the ability to produce high quality animation and the limitations are discussed in a later paragraph.

The proposed multi-resolution hierarchy possesses many strengths, including the ability to operate on different levels of shape and skin detail, the ability to reconstitute higher order shapes using lower order shapes and details and the ability to reconstruct the original model topology using any resolution shape in the hierarchy. However, it also possesses limitations as well.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The proposed method for building the multi-resolution hierarchy uses techniques which are suitable for the face region, but not necessarily for the full region of the head. Refinement by cylindrical projection is not an appropriate method for producing proper results for ears. Many aberrations can surface from improper projection (e.g., back of ear gets projected onto the front). In fact, the quality of a coarsely-approximated ear is more likely to degrade than improve if cylindrical projection is applied.

There are two limitations on the direct use of the reconstructed models in animation. Firstly, the source data only captures the face region of subjects and the reconstruction algorithm reproduces the face accordingly. However, full heads are needed in animation to achieve more visually attractive results. The second reason for the limitation is because of inadequate lip and eye treatment. The models are laser scans of plaster cast molds made from actual human faces and the subjects need to close their eyes and mouths while the molds are being made. Without performing any pre-processing on the scanned models, their eyes and mouths do not open properly. The generic model itself also does not have sufficient numbers of triangles for the lips and the eyelids, making it difficult to reconstruct error-free eyes and mouths. The lips and eyes also do not have natural shapes.

When reconstituting a given resolution shape and an associated level of skin detail, the shape must be in its original state to ensure that the recomposed model is error-free. This essentially means that combining results of skin detail manipulation and shape manipulation together does not guarantee meaningful results. The same conclusion applies for making two resolution transitions, i.e., from low to medium, then from medium to high.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

6.3 Future Research

The research documented in this thesis lays the groundwork for many areas of future research. The multi-resolution shape and detail representation along with the 3D skin concept are the fundamental components arising from this work and many topics can be pursued to extend this work.

- More generally suitable refinement algorithms should be explored not only to address the limitation with ears mentioned in the previous section, but to possibly improve the overall quality of the face as well.
- Future research can investigate the construction of full heads. The principle of transferring components can perhaps be applied in that work as well. A limited number of high quality ear and skull samples can be compiled and then transferred to other models in the same manner as skin. Alternatively, an individual's high resolution skin detail may be integrated into a generic head structure (of arbitrary resolution) to produce a functional head model which perhaps has an imbedded animation structure.
- Image analysis algorithms can be employed to automatically define more accurate lip and eyelid contact lines in scanned models. These areas can then be incised to create the separation between the upper and lower entities required for opening during animation. More advanced correspondence tracking during model parameterization can ensure that the lips and eyelids are reconstructed flawlessly. Additionally, improving

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the lips and eyelids of the generic model by altering their shape and increasing triangle counts in those areas can have positive effects.

- The scanned models currently do not have colour information. Some work can be done in the area of texture mapping to add surface colour for better rendering. Compiling a bank of textures may be a worthwhile endeavour.
- This research can be extended to encompass the entire body. Representing a body in multiple resolutions is equally as important as it is for faces in computer graphics. With bodies in full correspondence, animations can be encoded with respect to structure to compile a database of animations. This is particularly beneficial for virtual worlds that require a large population of humans to be animated.
- More work should be done to address the limitation of skin detail. New ways to compute skin detail effectively should also be explored.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

References

[Blanz 99] Blanz, V. and Vetter, T. 1999. A morphable model for the synthesis of 3D faces. In *Proceedings of the 26th Annual Conference on Computer Graphics and interactive Techniques* International Conference on Computer Graphics and Interactive Techniques. ACM Press/Addison-Wesley Publishing Co., New York, NY, 187-194.

[Blinn 78] Blinn, J. F. 1978. Simulation of wrinkled surfaces. In *Proceedings of the 5th Annual Conference on Computer Graphics and interactive Techniques* (August 23 - 25, 1978). SIGGRAPH '78. ACM Press, New York, NY, 286-292.

[Borshukov 03] Borshukov, G. and Lewis, J. P. 2003. Realistic human face rendering for "The Matrix Reloaded". In *Proceedings of the SIGGRAPH 2003 Conference on Sketches &Amp; Applications: in Conjunction with the 30th Annual Conference on Computer Graphics and interactive Techniques* (San Diego, California, July 27 - 31, 2003). SIGGRAPH '03. ACM Press, New York, NY, 1-1.

[Brennan 85] Brennan, S. E. 1985. The caricature generator. *Leonardo*. 18, 170-178.

[Bui 03] Bui, T. D., Poel, M., Heylen, D. and Nijholt, A. 2003. Automatic face morphing for transferring facial animation. In *Proceedings of the 6th IASTED International Conference on Computers, Graphics and Imaging* (Honolulu, Hawaii, USA, August 13, 2003). 19-23.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[Chiang 05] Chiang, P.-Y., Liao, W.-H. and Li, T.-Y. 2005. Automatic caricature generation by analyzing facial features. In *Proceedings of Asian Conference on Computer Vision* (Korea, 2005).

[Cohen 96] Cohen, J., Varshney, A., Manocha, D., Turk, G., Weber, H., Agarwal, P., Brooks, F., and Wright, W. 1996. Simplification envelopes. In *Proceedings of the 23rd Annual Conference on Computer Graphics and interactive Techniques* SIGGRAPH '96. ACM Press, New York, NY, 119-128.

[Cook 84] Cook, R. L. 1984. Shade trees. In *Proceedings of the 11th Annual Conference on Computer Graphics and interactive Techniques* H. Christiansen, Ed. SIGGRAPH '84. ACM Press, New York, NY, 223-231.

[Coquillart 90] Coquillart, S. 1990. Extended free-form deformation: a sculpturing tool for 3D geometric modeling. In *Proceedings of the 17th Annual Conference on Computer Graphics and interactive Techniques* (Dallas, TX, USA). SIGGRAPH '90. ACM Press, New York, NY, 187-196.

[DeRose 98] DeRose, T., Kass, M., and Truong, T. 1998. Subdivision surfaces in character animation. In *Proceedings of the 25th Annual Conference on Computer Graphics and interactive Techniques* SIGGRAPH '98. ACM Press, New York, NY, 85-94.

[Eck 91] Eck, M. 1991. Interpolation methods for reconstruction of 3D surfaces from sequences of planar slices. *CAD und Computergraphik.* 13(5), 109 – 120.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[Fidaleo 00] Fidaleo, D., Noh, J-Y., Kim, T., Enciso, R. and Neumann, U. 2000. Classification and volume morphing for performance-driven facial animation. *International Workshop on Digital and Computational Video.*

[Fordham 03] Fordham, J.. 2003. Middle-earth Strikes Back. *CineFex* number 92, 70-142.

[Fujiwara 01] Fujiwara, T., Koshimizu, H., Fujimura, K., Fujita, G., Noguchi, Y. and Ishikawa, N. 2001. 3D modeling system of human face and full 3D facial caricaturing. In *Proceedings of the Seventh International Conference on Virtual Systems and Multimedia.* 625-633.

[Godin 01] Godin, G., Rioux, M., Beraldin, J.-A., Levoy, M., Cournoyer, L. and Blais, F., 2001. An assessment of laser range measurement on marble surfaces. In *Proceedings of the 5th Conference on Optical 3D Measurement Techniques* (Vienna, Austria, Oct. 1 – 4, 2001).

[Goto 02] Goto, T., Lee, W.-S. and Magnenat-Thalmann, N. 2002. Facial feature extraction for quick 3D face modeling. *Signal Processing: Image Communication.* 17(3), 243-259.

[Guenter 98] Guenter, B., Grimm, C., Wood, D., Malvar, H., and Pighin, F. 1998. Making faces. In *Proceedings of the 25th Annual Conference on Computer Graphics and interactive Techniques* SIGGRAPH '98. ACM Press, New York, NY, 55-66.

[Guskov 00] Guskov, I., Vidimče, K., Sweldens, W., and Schröder, P. 2000. Normal meshes. In *Proceedings of the 27th Annual Conference on Computer Graphics and interactive Techniques* International Conference on Computer Graphics and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Interactive Techniques. ACM Press/Addison-Wesley Publishing Co., New York, NY, 95-102.

[Hilton 02] Hilton, A,. Starck, J. and Collins, G. 2002. From 3d shape capture to animated models. *Proceedings of the 1st International Symposium on 3D Data Processing, Visualization and Transmission* (Padova, Italy, June 19 - 21, 2002). 246-255.

[Ip 96] Ip, H. H.S. and Yin, L. 1996. Constructing a 3D individual head model from two orthogonal views. *The Visual Computer.* 12(5), 254-266.

[Kurihara 91] Kurihara T. and Arai K. 1991. A transformation method for modeling and animation of the human face from photographs. In *Proceedings of Computer Animation '91* (Tokyo, Japan, 1991) N. Magnenat-Thalmann and D. Thalmann. Eds. Computer Animation '91. Springer-Verlag, 45-58.

[Lasseter 87] Lasseter, J. 1987. Principles of traditional animation applied to 3D computer animation. In *Proceedings of the 14th Annual Conference on Computer Graphics and interactive Techniques* M. C. Stone, Ed. SIGGRAPH '87. ACM Press, New York, NY, 35-44.

[LeBlanc 91] LeBlanc, A., Kalra, P., Magnenat-Thalmann, N. and Thalmann, D. 1991. Sculpting with the 'ball & mouse' metaphor. In *Proceedings of Graphics Interface'91* (Calgary, Canada). 152-159.

[Lee A 00] Lee, A., Moreton, H., and Hoppe, H. 2000. Displaced subdivision surfaces. In *Proceedings of the 27th Annual Conference on Computer Graphics and interactive*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Techniques International Conference on Computer Graphics and Interactive Techniques. ACM Press/Addison-Wesley Publishing Co., New York, NY, 85-94.

[Lee 97] Lee, W.-S., Kalra, P. and Magnenat-Thalmann, N. 1997. Model based face reconstruction for animation. In *Proceedings of MMM'97* (Singapore). World Scientific Press, 323-338.

[Lee 98] Lee, W.-S. and Magnenat-Thalmann, N. 1998. Head modeling from pictures and morphing in 3D with image metamorphosis based on triangulation. In *Proceedings of the international Workshop on Modeling and Motion Capture Techniques For Virtual Environments* N. Magnenat-Thalmann and D. Thalmann, Eds. Lecture Notes In Computer Science, vol. 1537, Springer-Verlag, London, 254-267.

[Lee 99] Lee, W.-S., Beylot, P., Sankoff, D., and Magnenat-Thalmann, N. 1999. Generating 3D virtual populations from pictures of a few individuals. In *Proceedings of the 6th international Workshop on Algorithms and Data Structures* (August 11 - 14, 1999). F. K. Dehne, A. Gupta, J. Sack, and R. Tamassia, Eds. Lecture Notes In Computer Science, vol. 1663, Springer-Verlag, London, 134-144.

[Lee W 00] Lee, W.-S. and Magnenat-Thalmann, N. 2000. Fast head modeling for animation. *Journal Image and Vision Computing* (Elsevier, March 1, 2000). 18(4), 355-364.

[Lee 02] Lee, W.-S. 2002. Making one structured body. *VSMM 2002* (GyeongJu, Korea, September 25 - 27, 2002). 558-567.

[Lee 93] Lee, Y. C., Terzopoulos, D. and Waters, K. 1993. Constructing physics-based facial models of individuals. In *Proceedings of Graphics Interface '93*, 1--8,.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[Lee 95] Lee, Y. C., Terzopoulos, D., and Waters, K. 1995. Realistic face modeling for animation. In *Proceedings of SIGGRAPH 95*, 55-62

[Levoy 00] Levoy, M., Pulli, K., Curless, B., Rusinkiewicz, S., Koller, D., Pereira, L., Ginzton, M., Anderson, S., Davis, J., Ginsberg, J., Shade, J., and Fulk, D. 2000. The digital Michelangelo project: 3D scanning of large statues. In *Proceedings of the 27th Annual Conference on Computer Graphics and interactive Techniques* International Conference on Computer Graphics and Interactive Techniques. ACM Press/Addison-Wesley Publishing Co., New York, NY, 131-144.

[Liang 02] Liang, L., Chen, H., Xu, Y., and Shum, H. 2002. Example-based caricature generation with exaggeration. In *Proceedings of the 10th Pacific Conference on Computer Graphics and Applications* (Beijing, China, October 9 - 11, 2002). 386-393.

[Loop 87] Loop, C. 1987. *Smooth subdivision surfaces based on triangles*. Master's thesis, University of Utah, Department of Mathematics.

[Magnenat-Thalmann 87] Magnenat-Thalmann, N. and Thalmann, D. 1987. The direction of synthetic actors in the film Rendez-vous à Montrèal. *IEEE Computer Graphics and Applications*. 7(12), 9-19.

[Mo 03] Mo, Z., Lewis, J. P. and Neumann, U. 2003. Improved automatic caricature by feature normalization and exaggeration. *SIGGRAPH 2003 Sketches and Applications* (San Diego, CA): ACM SIGGRAPH, July (2004).

[Noh 00] Noh, J., Fidaleo, D., and Neumann, U. 2000. Animated deformations with radial basis functions. In *Proceedings of the ACM Symposium on Virtual Reality*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Software and Technology (Seoul, Korea, October 22-25, 2000). VRST '00. ACM Press, New York, NY, 166-174.

[Noh 01] Noh, J. and Neumann, U. 2001. Expression cloning. In *Proceedings of the 28th Annual Conference on Computer Graphics and interactive Techniques* SIGGRAPH '01. ACM Press, New York, NY, 277-288.

[Peercy 97] Peercy, M., Airey, J., and Cabral, B. 1997. Efficient bump mapping hardware. In *Proceedings of the 24th Annual Conference on Computer Graphics and interactive Techniques* International Conference on Computer Graphics and Interactive Techniques. ACM Press/Addison-Wesley Publishing Co., New York, NY, 303-306.

[Pötzsch 96] Pötzsch, M., Maurer, T., Wiskott, L. and von der Malsburg, C. Reconstruction from graphs labeled with responses of Gabor filters. In *Proceedings of the ICANN 1996*. 845-850.

[Powell 87] Powell, M. J. D. 1987. Radial basis functions for multivariate interpolation: a review. In *Algorithms for Approximation*, Clarendon Press, Oxford.

[Praun 01] Praun, E., Sweldens, W., and Schröder, P. 2001. Consistent mesh parameterizations. In *Proceedings of the 28th Annual Conference on Computer Graphics and interactive Techniques* SIGGRAPH '01. ACM Press, New York, NY, 179-184.

[Proesmans 97] Proesmans, M. and Van Gool, L. 1997. Reading between the lines—a method for extracting dynamic 3D with texture. In *Proceedings of the ACM*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Symposium on Virtual Reality Software and Technology (Lausanne, Switzerland). VRST '97. ACM Press, New York, NY, 95-102.

[Sannier 97] Sannier, G. and Magnenat-Thalmann, N. 1997. A user-friendly texture-fitting methodology for virtual humans. *Proceedings of Computer Graphics International '97* (Los Angeles, California, USA, 1997). 167-176.

[Sederberg 86] Sederberg, T. W. and Parry, S. R. 1986. Free-form deformation of solid geometric models. In *Proceedings of the 13th Annual Conference on Computer Graphics and interactive Techniques* D. C. Evans and R. J. Athay, Eds. SIGGRAPH '86. ACM Press, New York, NY, 151-160.

[Shan 04] Shan, Y., Lin, Z. and Zhang, Z. 2004. Image-based surface detail transfer. *IEEE Computer Graphics and Applications* (May/Jun, 2004). 24(3), 30-35.

[Soon 06 "Exaggeration"] Soon, A. and Lee, W.-S. 2006. Exaggeration of extremely detailed 3D faces. In *Proceedings of The IASTED International Conference on Internet and Multimedia Systems and Applications* (Innsbruck, Austria, February 13-15, 2006).

[Soon 06 "Shape"] Soon, A. and Lee, W.-S. 2006. Shape-based detail-preserving exaggeration of extremely accurate 3D faces. Accepted for publication in *The Visual Computer*.

[Sun 01] Sun, W., Hilton, A., Smith, R. and Illingworth, J. 2001. Layered animation of captured data. *The Visual Computer: International Journal of Computer Graphics*. 17(8), 457-474.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

[Taylor 02] Taylor, J., Beraldin, J.-A., Godin, G., Cournoyer, L., Rioux, M. and Domey, J. 2002. NRC 3D imaging technology for museums & heritage. In *Proceedings of The First International Workshop on 3D Virtual Heritage* (Geneva, Switzerland, 2002). 70-75.

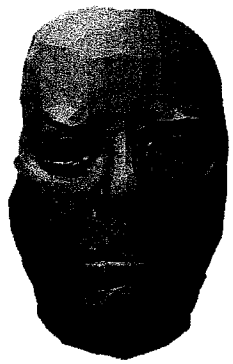
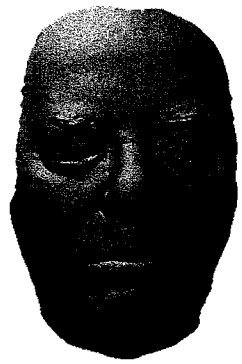
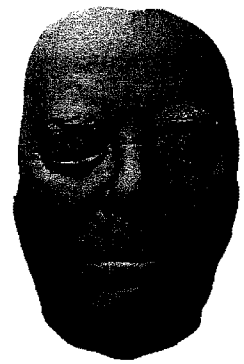
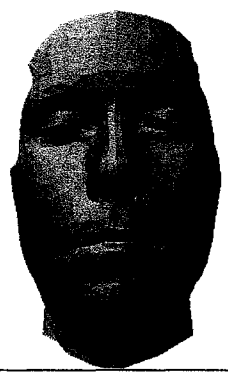
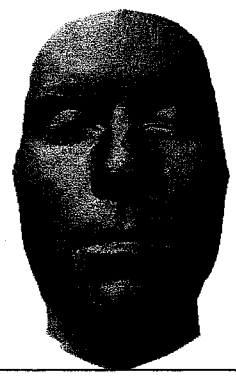
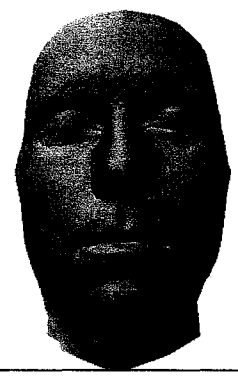
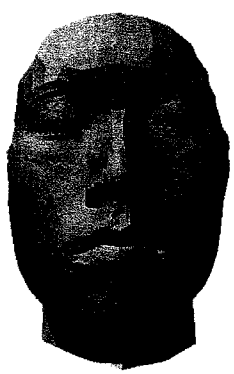
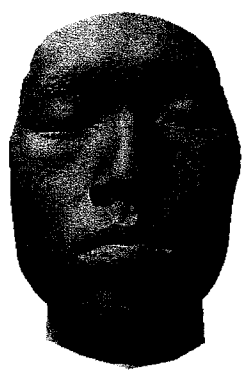
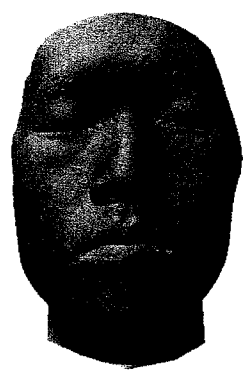
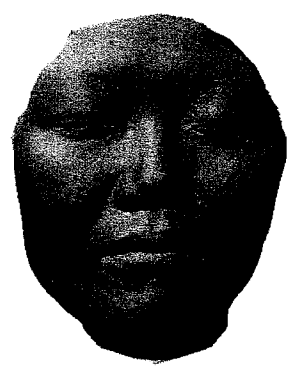
[Zhang 02] Zhang, L., Curless, B. and Seitz, S. M. 2002. Rapid shape acquisition using color structured light and multi-pass dynamic programming. In *Proceedings of the 1st International Symposium on 3D Data Processing, Visualization and Transmission* (Padova, Italy, June 19-21, 2002). 24-36.

[Zhang L 04] Zhang, L., Snavely, N., Curless, B., and Seitz, S. M. 2004. Spacetime faces: high resolution capture for modeling and animation. *ACM Trans. Graph.* 23, 3 (Aug. 2004). 548-558.

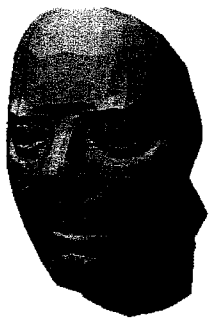
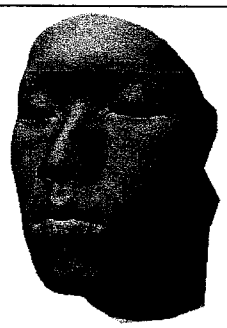
[Zhang Y 04] Zhang, Y., Sim, T. and Tan, C. L. 2004. Adaptation-based individualized face modeling for animation using displacement map. In *Proceedings of Computer Graphics International 2004* (Crete, Greece). 518-521.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Appendix A Multi-resolution Faces

Person	Low	Medium	High
A			
B			
C			
D			

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Person	Low	Medium	High
A			
B			
C			
D			

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.