

Acquisition of Human Faces Using a Measurement-Based Skin Reflectance Model

A dissertation submitted to
ETH Zurich

for the degree of
Doctor of Sciences

presented by

Tim Alexander Weyrich

Dipl.-Inform., Universität Karlsruhe (TU)

born September 26, 1974

citizen of Germany

accepted on the recommendation of

Prof. Markus Gross, examiner

Prof. Henrik Wann Jensen, co-examiner

Dr. Hanspeter Pfister, co-examiner

2006

Abstract

The scope of this thesis is threefold: design and construction of a 3D facial scanning system, development of a practical reflection model describing reflectance properties of human skin, and an analysis of facial appearance of multiple subjects, leading to intuitive editing operations for the skin model.

The scanning system consists of two devices that separately acquire surface reflectance and subsurface scattering properties, respectively. Together they capture all appearance parameters of an individual face needed to generate novel, synthetic images of the face. The former device consists of sixteen cameras and 150 light-sources attached to a hemisphere surrounding the scanned subject. In order to acquire the *reflectance field* of the face, the light-sources are sequentially flashed, while all cameras simultaneously capture the lit face. The latter device is a contact device that uses a bundle of optical fibers to feed light into the skin of the subject, and to collect exiting radiance at different distances to the source fiber.

The second integral part of the thesis has been to develop an appropriate reflectance model that allows a compact representation, as well as an analysis across multiple subjects with varying skin properties. The skin model decomposes the high-dimensional bidirectional scattering surface reflectance distribution function (BSSRDF) of skin into components that can be estimated from the measured data. The model is intuitive, amenable to interactive rendering, and easy to edit.

Using the custom-built acquisition devices, a large group of people has been scanned, providing 3D face geometry, skin reflectance, and subsurface scattering for each subject. The data has been fitted to the model, and an analysis of the reflectance data across the datasets revealed variations according to subject age, skin type, gender, and external factors (heat, cold, makeup, etc.). We use two alternative instantiations of our model. In one case, we derive a low-dimensional model using non-negative matrix factorization (NMF) that spans the space of skin reflectance in our database. In the other case, histogram matching techniques are used to transfer skin characteristics between subjects. Both techniques allow a user to control skin appearance by meaningful parameters—such as skin type, gender, and age—and change the overall appearance of a person (e.g., making a Caucasian face look more Asian) or change local features (e.g., adding moles, freckles, or hair follicles).

These controls make our model an intuitively editable skin reflectance representation that is suited for animation and photo-realistic rendering of human faces, preserving the natural appearance of skin.

Zusammenfassung

Die vorliegende Arbeit befasst sich im wesentlichen mit drei Zielen: Entwurf und Konstruktion eines 3D-Gesichtsscanners, Entwicklung eines geeigneten Reflexionsmodells zur Beschreibung der Reflexionseigenschaften menschlicher Haut und der Analyse der Aussehens der Gesichter mehrerer Probanden, die letztlich zu intuitiven Operationen führen soll, um die Erscheinung des Hautmodells zu verändern.

Das Scanner-System besteht aus zwei Geräten, die unabhängig voneinander Oberflächenreflektivität und Streueffekte in tieferen Hautschichten messen. Gemeinsam erfassen sie damit alle wesentlichen Parameter des Aussehens eines individuellen Gesichts, die benötigt sind, um neue Ansichten des Gesichts künstlich zu generieren. Das erste Gerät besteht aus sechzehn Kameras und 150 Lichtquellen in hemisphärischer Anordnung um den zu scannenden Probanden. Um das *reflectance field* eines Gesichts aufzunehmen werden die Lichtquellen der Reihe nach durchgeschaltet, während die Kameras simultan das beleuchtete Gesicht aufnehmen. Das zweite Gerät ist ein Kontaktgerät, das ein Bündel optischer Fasern verwendet, um Licht in die Haut des Probanden zu leiten und um die wiederaustretende Radianz an verschiedenen Entfernungen zur Quellfaser abzugreifen.

Der zweite Bestandteil der Arbeit befasst sich mit der Entwicklung eines geeigneten Reflexionsmodells, das eine kompakte Darstellung erlaubt und es

ermöglicht, eine Analyse über verschiedene Personen mit unterschiedlichen Hauteigenschaften durchzuführen. Das Hautmodell zerlegt die hochdimensionale *bidirectional scattering surface reflectance distribution function* (BSSRDF) der Haut in Komponenten, die von Messdaten geschätzt werden können. Das Modell ist intuitiv, geeignet fuer interaktive Darstellungen und ist einfach zu editieren.

Mit Hilfe dieser beiden Spezialanfertigungen wurde eine grössere Gruppe von Personen eingescannt. Die Scans umfassen jeweils die 3D-Geometrie des Gesichts, Hautreflektivität und subkutane Streueigenschaften. Das Modell wurde jeweils an die Daten angepasst, und eine Analyse der Reflexionseigenschaften über mehrere Datensätze lieferte Unterschiede je nach Alter, Hauttyp und Geschlecht der Person, sowie in Abhängigkeit von externen Faktoren wie Hitze, Kälte, Make-up, etc. Wir benutzen zwei alternative Instanzen unseres Modells. Im einen Fall leiten wir ein niederdimensionales Modell mittels nicht-negativer Matrixfaktorisierung (NMF) her, das den Raum der verschiedenen Hautreflektivitäten innerhalb unserer Datenbank aufspannt. Im anderen Fall verwenden wir Histogrammanpassungstechniken, um Hautcharakteristiken zwischen verschiedenen Personen zu übertragen. Beide Techniken erlauben es einem Benutzer, die Erscheinung menschlicher Haut über eingängige Parameter wie zum Beispiel Hauttyp, Geschlecht und Alter, zu kontrollieren und die generelle Erscheinung einer Person zu verändern (z.B. indem ein kaukasisches Gesicht eine asiatischere Haut erhält) oder einzelne Merkmale abzuändern (z.B. Leberflecken, Sommersprossen oder Haarfollikel).

Diese Kontrollmöglichkeiten machen unser Modell zu einer intuitiv editierbaren Reflexionsrepräsentation, die fuer die Animation und Darstellung photorealistischer Gesichter geeignet ist und dabei die natürliche visuelle Erscheinung der Haut erhält.

Acknowledgements

My sincere thanks to my advisor Markus Gross. His advice and guidance have been invaluable, his enthusiasm and scientific knowledge both motivating and inspiring. I am grateful to the people at ETH Zurich for providing such an enjoyable and stimulating working environment.

Many thanks also to my co-advisors: Hanspeter Pfister from the Mitsubishi Electric Research Laboratories (MERL), whose scientific intuition and inexhaustible motivation were an invaluable driving force throughout the project, and Henrik Wann Jensen from the University of California, San Diego, who supported me with insightful discussions and comments and lifted my spirits in difficult project phases.

In addition, my thanks go to everyone at MERL I had the opportunity to work with. Special thanks to Wojciech Matusik, Bernd Bickel, Jinho Lee, Addy Ngan, and Craig Donner, without whom this project would not have been possible. Many thanks go also to Jason Lawrence, Diego Nehab, Chien Tu, Jeroen van Baar, Agata Lukasinska, Nathalie Charton, John Barnwell, William "Crash" Yerazunis, Freddy Bürki, Cliff Forlines, Simon Heinzle, and, last not least, Janet McAndless, all of them significantly contributing to the success of the face scanner project.

Many thanks go to my friends and colleagues at ETH, who supported me during my time in Zurich: Christoph Niederberger, Matthias Zwicker, Mark Pauly, Oscar Chinellato, Matthias Teschner, Stephan Würmlin, Christian Sigg, Richard Keiser, Martin Wicke, Michael Waschbüsch, Reto Lütolf, and last not least Remo Ziegler.

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Principal Contributions	3
1.3	Organization	5
2	Fundamentals and Related Work	7
2.1	Radiometry	7
2.2	Appearance Modeling	8
2.2.1	BRDF Models	9
2.2.2	The BSSRDF	11
2.2.3	3D Photography	13
2.2.4	Hybrid Appearance Models	14
2.3	Human Skin Rendering	15
2.3.1	Analytic Face Appearance Models	15
2.3.2	Image-Based Face Models	16
2.3.3	Skin Color and Texture Models	17
2.4	Point-Based Graphics	18
2.4.1	Surfel Representation	18
2.5	Unstructured Lumigraph Interpolation	19

3	Geometry Processing	21
3.1	Motivation	21
3.2	An Interactive Approach	23
3.3	Underlying Techniques	25
3.3.1	Search Data Structures	25
3.3.2	MLS Projection	26
3.3.3	Point Relaxation	26
3.4	Outlier Removal	27
3.4.1	Outlier Criteria	28
3.4.2	Classification	30
3.4.3	Performance Evaluation	30
3.5	Hole Filling and Smoothing	31
3.5.1	MLS Smoother	32
3.5.2	Point Relaxation	33
3.5.3	MLS Spray Can	33
3.5.4	Volumetric Diffusion	34
3.5.5	Performance Evaluation	35
3.6	Discussion	36
4	Surface Reflectance Field Evaluation	39
4.1	Previous Work	40
4.2	Overview	41
4.3	Deformable Surface Reflectance Fields	42
4.3.1	Approximate BRDF Preservation	44
4.3.2	Local Impostor Parameterization	46
4.3.3	Shading	49
4.4	Implementation	50
4.4.1	Surfel Representation	51
4.4.2	Reflectance Image Interpolation	52
4.4.3	Deferred Evaluation	53
4.5	Test Results	54
4.6	Reflectance Fields of Human Faces	58
4.7	Discussion	61
5	Facial Reflectance Field Acquisition	65
5.1	Acquisition Device	65
5.1.1	Design Considerations	65
5.1.2	Final Hardware	72
5.1.3	Acquisition Dome Software	73
5.2	Motion Compensation	74
5.2.1	Motion Estimation	76
5.2.2	Trajectory Recovery	79
5.2.3	Image Warp	80
5.3	Discussion	82

6	Subsurface Scattering Acquisition	85
6.1	Problem Statement	85
6.1.1	Requirements	86
6.1.2	Previous Designs	86
6.2	Acquisition Device	88
6.2.1	Preparatory Experiments	89
6.2.2	Design	90
6.2.3	Calibration	92
6.2.4	Measurement Procedure	93
6.3	Representation	93
6.3.1	Model Fit	94
6.3.2	Re-Parameterization	94
6.4	Evaluation	95
6.4.1	Spatial Variance	95
6.4.2	Anisotropy Measurements	96
6.5	Discussion	97
7	Processing Pipeline	99
7.1	Overview	99
7.2	Geometry Processing	100
7.2.1	Mesh Cleaning	102
7.2.2	Up-Sampling	102
7.2.3	Cropping	103
7.2.4	uv Parameterization	103
7.3	Visibility Classification	106
7.4	Sample Generation	108
7.5	Normal Estimation	110
7.5.1	Normals From Specular Direction	110
7.5.2	Helmholtz Stereopsis	111
7.5.3	Standard Photometric Stereo	113
7.5.4	Photometric Stereo With Highlight Removal	113
7.5.5	Comparison	115
7.6	Geometry Refinement	115
7.7	Reflectance Model Fit	117
7.8	Discussion	118
8	Skin Reflectance Model	119
8.1	Constituents	119
8.1.1	Surface Reflectance	120
8.1.2	Diffuse Reflectance	121
8.1.3	Micro-Normals	122
8.2	Model Fit	123
8.2.1	Diffuse Albedo Separation	125
8.2.2	Diffuse Reflectance	126

8.2.3	Surface Reflectance	128
8.2.4	Micro-Normals	133
8.3	Discussion	133
9	Rendering	135
9.1	Implementation	135
9.2	Results	136
9.2.1	Data-Driven Surface BRDF	136
9.2.2	Torrance-Sparrow & Improved Geometry	137
9.3	Discussion	141
10	Appearance Analysis	145
10.1	Analysis	145
10.1.1	Interface Reflectance	147
10.1.2	Translucency Variations	151
10.1.3	Albedo Texture	154
10.1.4	Further Analysis	156
10.2	Face Editing	156
10.2.1	Trait Vectors	156
10.2.2	Histogram Interpolation	158
10.3	Discussion	160
11	Conclusion and Outlook	161
11.1	Conclusion	161
11.2	Discussion and Future Work	162
A	Glossary And Notation	163
A.1	Notation	163
A.2	Radiometric Quantities	164
B	System Calibration	165
B.1	Overview	165
B.2	Camera System	166
B.2.1	Geometric Calibration	166
B.2.2	Photometric Calibration	167
B.3	Light Sources	168
B.3.1	Light Source Positions	169
B.3.2	Radiometric Calibration	169
B.4	3D Scanner Registration	170
C	Color Plates	171
	Curriculum Vitae	193
	List of Figures	195

List of Tables	199
Bibliography	201

C H A P T E R

1

Introduction

This chapter gives a brief introduction in the main goals of this thesis. It summarizes methods and results contributed by the thesis and concludes with a short overview over the organizational structure of this document.

1.1 Motivation

One of the biggest challenges in computer graphics is creating digital faces that are indistinguishable from real ones. Although general rendering quality in graphics often achieves photo-realistic appearance of synthetic objects, rendering of human faces still remains a demanding task. This is not only because of the complexity of facial appearance, but also due to that fact that human observers are experts in judging whether a face is “real” or not. The process of capturing an actor’s performance and likeness has accurately been named “Digital Face Cloning.” Digital face cloning has many applications in movies, games, medicine, cosmetics, computer vision, biometrics, and virtual reality. While recent feature films already show authentic artificial renderings of character faces [BL03, Fee04], the datasets have been especially designed for an appearance in a particular scene. They are either the result of extensive manual editing to make the model appear real under hand-optimized lighting design [BL03], or they merely tabulate an actor’s lighting-dependent appearance from a single viewpoint [HWT⁺04, Fee04]. Up to date, there is no generic procedure that allows for the acquisition of human faces,

1 Introduction

leading to a self-contained face model that can be used for renderings in an arbitrary context. Key requirements of such a model are generality, to allow for a flexible use, and editability, that is, it should be possible to change the face's appearance using intuitive controls. Up to date, editing of face models requires the hand of a skilled artist with a deep technological understanding.

This project deals with the development of a face acquisition pipeline that allows for the automated acquisition of generic face models from individual subjects. This comprises the construction of a respective acquisition hardware, the development of a suited model representation, and the analysis of facial appearance across multiple subjects to derive meaningful controls for editing of the face model.

The acquisition of the entire reflectance properties of a human face can not be achieved by present 3D scanners. Therefore, a new scanner design was needed. The scanner has to be able to acquire geometry and information on color and full reflectance properties for every single point in the face. Reflectance properties are required to render the reconstructed face under arbitrary lighting conditions, preserving the appearance of the subject's individual tissue properties. We developed a respective scanner design that consists of a large number of cameras and light sources that are triggered independently, in order to acquire the view- and illumination-dependent appearance of the face. It is crucial that the scanning procedure is very short, since motion artifacts tend to compromise a highly detailed reconstruction whenever living subjects are scanned.

Although this scanner design is able to acquire directional reflectance properties, it does not capture subsurface effects, such as scattering and absorption within skin tissue, that delocalize radiance transport. These effects have impact on the perceived smoothness of surface shading and become directly visible at shadow boundaries within the face. In order to capture subsurface effects as well, a second device had to be built to acquire internal reflectance properties of skin. Both devices require extensive geometric and radiometric calibrations in order to relate all measurements to physical quantities.

In the context of the thesis, an appropriate appearance representation had to be developed that can be derived from the acquired reflectance data. The representation should be suited for a flexible use in common rendering architectures, including a potential animation of the model. Due to the large number of cameras and light sources, the reflectance data tends to be huge. Therefore, the appearance representation has to be compact compared to the raw scanner data. Our analysis spans *image-based*, or *data-driven* models, respectively, as well as *parametric* or *model-driven* approaches. As skin contributes the largest part to facial appearance, the model should further be optimized for the physical properties of skin. A proper reconstruction of reflectance properties ideally captures the following effects:

Interface reflection. The reflection that takes place directly at the surface of materials, as it can be described by a bidirectional reflectance distribution function (BRDF). It is important to capture the spatial variance of the BRDF.

Subsurface scattering. This effect is very prominent in human skin. It is essential that the new reflectance model is able to capture scattering effects. It is still subject to investigation whether anisotropic and spatially varying aspects of sub-surface scattering are relevant for a skin representation.

Self-Shadowing. If not implicitly covered by the representation, it must be considered at acquisition time. For image-based representations, this effect becomes important as soon as the facial surface is animated.

Inter-Reflection. For inter-reflection, the same as for self-shadowing applies.

After initial experiments with an image-based representation, we followed a model-driven approach to describe skin reflectance.

Based on the novel skin representation, we analyze facial reflectance of over a hundred subjects. From this analysis, intuitive trait parameters of our skin model had to be derived. Using these parameters, it should become possible to change facial appearance based on simple controls, rather than hand-editing every aspect of the face. We believe that such an editable face model in conjunction with the largely automated model creation pipeline represents an important contribution for high-quality easy-to-use digital face cloning.

This project is a joint effort of the Computer Graphics Laboratory of ETH Zurich and the Mitsubishi Electric Research Laboratories (MERL) Cambridge, MA, USA.

1.2 Principal Contributions

In the context of this thesis, the following contributions have been made.

- ▶ Design of a practical skin model, based on the simple separation of surface and subsurface reflection (see Chapter 8). We evaluate two instantiations of the model, showing that it is well capable of representing the typical reflectance properties of human skin. Images from our model come close to reproducing photographs of real faces for arbitrary illumination and pose, see Chapter 9.
- ▶ Retrieval of model parameters from measurements. An important feature of the model is that all its parameters can be robustly estimated from measurements. No model parameters have to be manually set. At the same time, the model reduces the large amount of measured data to a manageable size and facilitates editing. This process is described in Chapter 8.

1 Introduction

- Development of a novel face scanning device consisting of multiple cameras and light sources (Chapter 5). The system makes high demands on its calibration, requiring an extensive calibration procedure to be developed, see Appendix B. In particular, we contribute a novel light source calibration.
- A novel scan cleaning toolbox, using state-of-the-art techniques complementing the existing concept of a point-based creation pipeline, as described in Chapter 3.
- Refinement of an image-based rendering method, allowing to directly render acquired reflectance field datasets under arbitrary deformations (Chapter 4).
- Development and construction of a subsurface scattering measurement device, suited to determine scattering and absorption properties of human skin. The device is described in Chapter 6.
- Analysis of scans of a large group of people. We fit our model to data of a large and diverse group of subjects. The analysis of this data provides insight into the variance of face reflectance parameters with respect to age, gender, skin type, and facial regions (Chapter 10). For the data-driven instance of our model, we derived a lower-dimensional model basis using non-negative matrix factorization (NMF) that spans the space of skin reflectance in our database (Chapter 8). This leads to a reflectance model specific to skin.
- Design of novel editing operations based on the results of an inter-subject analysis of reflectance properties. These operation enable one to change the facial appearance of an acquired datasets by simple controls that rely on appearance traits, such as age, gender, skin type, etc., see Chapter 10.
- Publication of a database with all statistics that have been acquired. We hope that the research community will benefit from the availability of our data, and that our analysis provides an important contribution to face synthesis and analysis in general.

Most of these contributions have been presented in the following publications:

TIM WEYRICH, WOJCIECH MATUSIK, HANSPETER PFISTER, BERND BICKEL, CRAIG DONNER, CHIEN TU, JANET MCANDLESS, JINHO LEE, ADDY NGAN, HENRIK WANN JENSEN, AND MARKUS GROSS. Analysis of human faces using a measurement-based skin reflectance model. *To appear in ACM Transactions on Graphics (SIGGRAPH 2006)*, 2006.

TIM WEYRICH, HANSPETER PFISTER, AND MARKUS GROSS. Rendering deformable surface reflectance fields. *In IEEE Transactions on Computer Graphics and Visualization*, 11:48–58, January–February 2005.

TIM WEYRICH, WOJCIECH MATUSIK, HANSPETER PFISTER, JINHO LEE, ADDY NGAN, HENRIK WANN JENSEN, AND MARKUS GROSS. A measurement-based skin reflectance model for face rendering and editing. Technical report, Mitsubishi Electric Research Laboratories, TR2005-071, December 2005.

FILIP SADLO, TIM WEYRICH, RONNY PEIKERT, AND MARKUS GROSS. A practical structured light acquisition system for point-based geometry and texture. In *To appear in Proceedings of the Eurographics Symposium on Point-Based Graphics 2005*, Stony Brook, USA, June 2005.

DANIEL COTTING, TIM WEYRICH, MARK PAULY, AND MARKUS GROSS. Robust watermarking of point-sampled geometry. In *Proceedings of International Conference on Shape Modeling and Applications 2004*, pages 233–242, Genova, Italy, June 2004. IEEE Computer Society Press.

TIM WEYRICH, MARK PAULY, RICHARD KEISER, SIMON HEINZLE, SASCHA SCANDELLA, AND MARKUS GROSS. Post-processing of scanned 3D surface data. In *Proceedings of Eurographics Symposium on Point-Based Graphics 2004*, pages 85–94, Zurich, Switzerland, June 2004.

1.3 Organization

Chapter 2 presents fundamentals of radiometry, 3D photography, and appearance modeling in general. It provides a short overview over the state of the art in face and skin modeling. Chapter 3 presents our contribution to processing of scanned surface data, a task that occurs for the scanned face datasets as well. In Chapter 4, we present our initial evaluation of image-based reflectance representations, in particular with respect to rendering datasets under deformations. We present the two reflectance acquisition devices in Chapter 5 and Chapter 6, respectively, describing the processing pipeline to refine the acquired data in Chapter 7. Our final modeling approach is presented in Chapter 8, and Chapter 9 discusses our rendering architecture with sample images of our acquired face models. Chapter 10 presents a detailed analysis of face reflectance, based on our model, and Chapter 11 summarizes the thesis and discusses potential directions of future work. Appendix A lists the notations and conventions used in this document. Additional information to the system calibration can be found in Appendix B.

Fundamentals and Related Work

This chapter introduces fundamentals of this thesis and discusses related work. The following section gives a short introduction into radiometry, Section 2.2 presents methodologies and previous work in the field of appearance acquisition. In Section 2.3, an overview of the state of the art in skin rendering is given.

2.1 Radiometry

Radiometry is the science of measuring light. It quantifies energy transport of electromagnetic waves. In computer graphics it provides the system of units for image synthesis, and it is the basis of every measurement of reflectance properties of real objects. In the remainder, we give a short summary over the respective quantities and their units. In addition, Appendix A.2 provides a quick overview over radiometric units in Table A.1.

Radiant Energy. The elementary quantity in radiometry is *radiant energy* Q . It is expressed in *joules* [J]. The *spectral radiant energy* $Q_\lambda = dQ/d\lambda$ is the amount of radiant energy per unit wavelength interval at wavelength λ . As colors are of particular interest for computer graphics, radiometric quantities should always be related to wavelength. However, for simplicity of notation, we omit wavelength in the remainder.

Radiant Flux. The time rate of flow of radiant energy is the *radiant flux* $\Phi = dQ/dt$. It expresses the amount of energy that “flows” through space per unit time. Radiant flux is measured in joules per second, that is, *watts* [W].

Radiant Flux Density. The radiant flux per unit area at a point on a surface is called *radiant flux density* [$W\ m^{-1}$]. Depending on whether the flux is incident or outgoing, this quantity is called *irradiance* E at point x_i (for incident flux),

$$E(x_i) = d\Phi/dA = \int_{\Omega_+} L(x_i, \omega_i) n d\omega_i, \quad (2.1)$$

or *radiant exitance* M at point x_o (for outgoing flux), with

$$M(x_o) = d\Phi/dA_{source} = \int_{\Omega_+} L(x_o, \omega_o) n d\omega_o. \quad (2.2)$$

Both quantities are expressed as integrals over the positive hemisphere Ω_+ over the surface point, integrating radiant flux. Note that the integration rules infer a “cosine weight” $\langle \omega, n \rangle$ with the surface normal n . This fact is often referred to in computer graphics as the *cosine law*.

Radiance. Radiant flux density per unit solid angle is *radiance* L [$W\ m^{-1}\ sr$]:

$$L(x, \omega) = \frac{d^2\Phi}{d\Omega dA_{source}} \quad (2.3)$$

Intuitively, it can be seen as being proportional to the radiant flux density within an infinitesimally thin cone of light, emanating from a surface point. Such a degenerate cone is in fact the definition of a *ray* of light in computer graphics, as it is used in ray tracers. When observing objects with a camera, each camera pixel integrates over the radiance from each surface point of the object.

In contrast to *photometric units*, which are used in lighting engineering, radiometry uses SI units (Le Système International d’Unités). The face scanner processing pipeline strictly follows this convention.

2.2 Appearance Modeling

A variety of different appearance representations exists for the visualization and animation of realistic 3D computer graphics models. We provide an overview over existing modeling techniques and about approaches that capture the appearance of real-world objects.

2.2.1 BRDF Models

An object's appearance is determined by its interaction with light. Thus, in early computer graphics light models have been developed to model different material properties. They describe the portion of light that is reflected toward the view-point when a surface point is lit from a certain direction.

This circumstance is commonly formalized by means of the *bi-directional reflectance distribution function* (BRDF, [NRH⁺77]). A BRDF f_λ describes the ratio of outgoing radiance toward ω_o at a point \mathbf{x}_o and incident radiance from a direction ω_i :

$$f_\lambda(\mathbf{x}_o, \omega_o, \omega_i) = \frac{L_\lambda(\mathbf{x}_o, \omega_o)}{L_\lambda(\mathbf{x}_o, \omega_i)} \quad (2.4)$$

Although a BRDF is dependent on wavelength λ , we will continue to drop λ for the sake of easier notation. Unless differently stated, we use a scalar notation and approximate any spectral distribution by an RGB model of three independent color channels. The two spherical directions ω_o and ω_i correspond to the view vector \mathbf{v} and a lighting direction \mathbf{l} , relative to a local surface frame. They are often expressed in spherical coordinates (ϑ_o, φ_o) and (ϑ_i, φ_i) . If the reflectance is independent from the tangential orientation, that is, from a rotation around the normal, a BRDF is called *isotropic*. In this case, φ_i is set to zero without loss of generality. Also note that a BRDF is potentially dependent on the surface location \mathbf{x}_o . In this case, we refer to a *spatially varying* BRDF (SVBRDF).

The rendering equation can be expressed by means of a BRDF f by integrating over the positive hemisphere of incidence:

$$L(\mathbf{x}_o, \omega_o) = \int_{\Omega_+} f(\mathbf{x}_o, \omega_o, \omega_i) n d\omega_i. \quad (2.5)$$

Several BRDF models have been developed in the past [TS67, Bli77, CT82, HTSG91, War92, MPBM03]. In the remainder we present a selection of models that are relevant for this project.

Blinn-Phong. The Blinn-Phong model [Bli77] is the most widely used analytic reflectance model due to its simplicity. It is not physically-based, but it is very fast to compute and easy to apply. A variant of the Blinn-Phong model is:

$$f_s^{\text{BP}} = \rho_s \frac{n+2}{2\pi} \cos^n \delta. \quad (2.6)$$

Here δ is the angle between the normal \mathbf{n} and the half-way vector \mathbf{h} between the incoming illumination and the camera view direction, $\mathbf{h} := (\mathbf{l} + \mathbf{v}) / \|\mathbf{l} + \mathbf{v}\|$. The specular exponent n controls the shape of the lobe and ρ_s is a scaling factor. The component $(n+2)/(2\pi)$ is for energy normalization so that the cosine lobe always integrates to one.

Torrance-Sparrow. The Torrance-Sparrow model [TS67] is an analytical model that assumes that the surface consists of small, randomly disposed, mirror-like facets. It assumes that specular reflection from these facets plus a diffuse component due to multiple reflections and internal scattering are the basic mechanisms of the reflection process. Ngan et al. [NDM05] analyzed a variety of analytical BRDF models and has shown that the Torrance-Sparrow model is the most accurate isotropic reflectance model among them.

The model can be stated as:

$$f_s^{\text{TS}} = \rho_s \frac{1}{\pi} F_r(\langle \boldsymbol{\omega}_o, \mathbf{h} \rangle) \frac{DG}{\langle \mathbf{n}, \boldsymbol{\omega}_i \rangle \langle \mathbf{n}, \boldsymbol{\omega}_o \rangle}, \quad (2.7)$$

with

$$G = \min\left\{1, \frac{2\langle \mathbf{n}, \mathbf{h} \rangle \langle \mathbf{n}, \boldsymbol{\omega}_o \rangle}{\langle \boldsymbol{\omega}_o, \mathbf{h} \rangle}, \frac{2\langle \mathbf{n}, \mathbf{h} \rangle \langle \mathbf{n}, \boldsymbol{\omega}_i \rangle}{\langle \boldsymbol{\omega}_o, \mathbf{h} \rangle}\right\},$$

$$D = \frac{1}{m^2 \cos^4 \delta} e^{-[(\tan \delta)/m]^2}. \quad (2.8)$$

G is the masking and shadowing factor, also referred to as the geometrical attenuation factor. It takes into account that adjacent facets may obstruct either the flux incident upon a given facet or the flux reflected from it. D is the Beckmann micro-facet distribution and F_r is the reflective Fresnel term. The free variables are a scale parameter σ_s , the roughness parameter m and the refraction index that controls the reflective Fresnel function $F_r(\eta, \langle \mathbf{v}, \mathbf{h} \rangle)$. \mathbf{h} is the normalized half-way vector between the incoming illumination \mathbf{l} and the camera view direction \mathbf{v} , η is the refractive index at the air/material transition.

Data-Driven BRDF Matusik et al. [MPBM03] introduce an innovative representation for BRDFs. Instead of using analytic functions to describe surface reflectance, they measured a large number of real-world material samples, obtaining from each a densely interpolated function table of its BRDF (see Figure 2.1). They use this BRDF database as a linear basis to generate novel BRDFs by interpolating and extrapolating in the space of acquired BRDFs. Let the measured BRDFs be tabulated functions f_i . Then a new BRDF f is synthesized by

$$f(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i) = \sum_{i=1}^n w_i f_i(\boldsymbol{\omega}_o, \boldsymbol{\omega}_i). \quad (2.9)$$

That is, the new BRDF can be represented by a weight vector $\mathbf{w} = (w_1, \dots, w_n)^\top$, denoting a point in the data-driven BRDF space.

They apply both linear and non-linear dimensionality reduction techniques to discover a lower-dimensional representation of their measurements. They use the vectorial representation to navigate the BRDF space. From their analysis, the authors derive that a BRDF basis of comparatively low dimensionality is sufficient

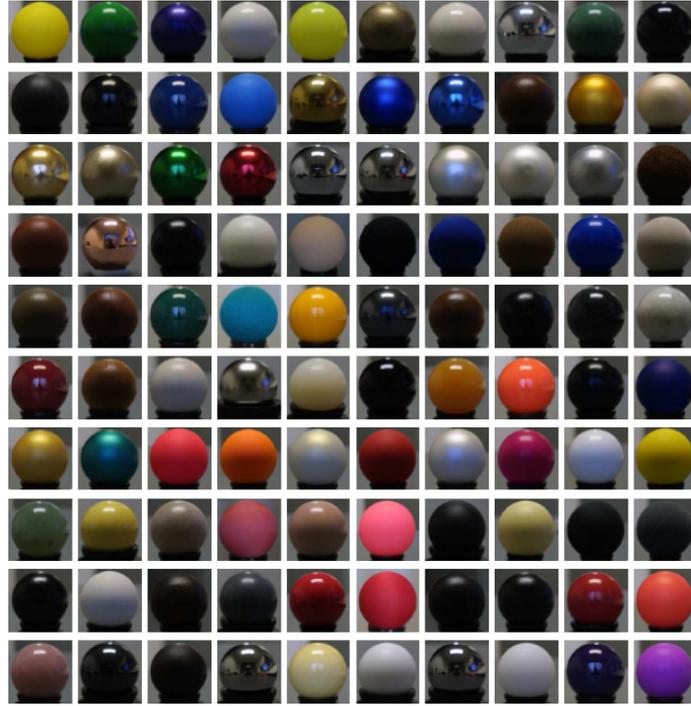


Figure 2.1: Sample basis BRDFs. Depicted are spheres that have been measured to construct the data-driven BRDF basis. (From [MPBM03], by courtesy of W. Matusik.)

for a perceptually faithful reproduction of any physically plausible BRDF. Note that any convex combination of basis BRDFs builds a physically valid BRDF.

As a linear model, the data-driven BRDF is well suited for data fitting. Linearity further simplifies statistical analysis. Accordingly, some of our experiments use the data-driven BRDF within our skin model (see Section 8.2.3).

2.2.2 The BSSRDF

Generally the BRDF approach exhibits a significant simplification: for most real-world materials, the interaction of light and matter does not exclusively take place at the surface. Usually, a fraction of light enters the surface layer and leaves it at different points, after having been scattered multiple times inside the material. This process is commonly called subsurface scattering. A model that also incorporates subsurface light transport is the *bi-directional surface scattering reflectance distribution function* (BSSRDF, [NRH⁺77]) S_λ , expressing the radiance transport between two different surface points \mathbf{x}_i and \mathbf{x}_o :

$$S_\lambda(\mathbf{x}_o, \boldsymbol{\omega}_o; \mathbf{x}_i, \boldsymbol{\omega}_i) = \frac{dL_\lambda(\mathbf{x}_o, \boldsymbol{\omega}_o)}{d\Phi_\lambda(\mathbf{x}_i, \boldsymbol{\omega}_i)} \quad (2.10)$$

Consequently, the respective rendering equation additionally has to integrate over the material surface (again, we omit wavelength):

$$L(\mathbf{x}_o, \boldsymbol{\omega}_o) = \int_A \int_{\Omega_+} S(\mathbf{x}_o, \boldsymbol{\omega}_o; \mathbf{x}_i, \boldsymbol{\omega}_i) \mathbf{n} d\boldsymbol{\omega}_i dx_i. \quad (2.11)$$

Subsurface scattering effects can be simulated by Monte Carlo methods [DEL⁺99, JD99]. A simple analytical model is presented in the following section.

The Diffusion Approximation

For homogeneous, isotropic materials with a high material density, causing penetrating photons to be subject to a sufficiently large number of scattering events before exiting the material, light transport can be approximated by diffusion theory [FPW92, Sta95]. The diffusion process is sufficiently specified by the material's scattering and absorption properties, namely by the reduced scattering coefficient μ'_s and the absorption coefficient μ_a .

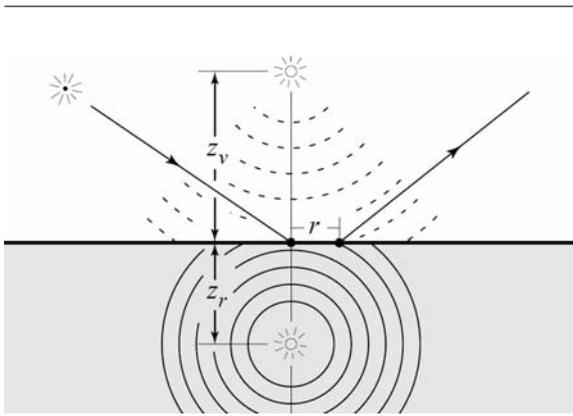


Figure 2.2: *The dipole model.* (Courtesy of H. W. Jensen.)

When observing points at the interface to the material, Fresnel effects at the air/material interface additionally have to be considered. Whereas pure surface reflectance models, such as the Torrance-Sparrow model, only model Fresnel reflection F_r at the surface, the incident flux entering the diffusion process is subject to fresnel transmission $F_t = 1 - F_r$.

Jensen et al. [JMLH01] show how the diffusion model can be used to efficiently render the BSSRDF. They introduce the *dipole model* as an approximation of the diffusion process, and

demonstrate how it can be fitted to measurements of real-world materials, as described in Section 6.1.2. Using the dipole diffusion approximation, rendering times for highly scattering materials drastically decreased. The diffusion model can be written in BSSRDF form:

$$S_d(\mathbf{x}_o, \boldsymbol{\omega}_o; \mathbf{x}_i, \boldsymbol{\omega}_i) = \frac{1}{\pi} F_t(\mathbf{x}_i, \boldsymbol{\omega}_i) R_d(\|\mathbf{x}_o - \mathbf{x}_i\|) F_t(\mathbf{x}_o, \boldsymbol{\omega}_o), \quad (2.12)$$

with a radially symmetric point-spread function $R_d(r)$ and two transmissive Fresnel terms F_t at the points of incidence and exitance. The dipole approximation

uses two virtual point light sources at distances z_v and z_r , see Figure 2.2, one of them of “negative” intensity, to derive a simple formula of R_d :

$$R_d(r) = \frac{\alpha' z_r (1 + \mu_{tr} d_r) e^{-\mu_{tr} d_r}}{4\pi d_r^3} - \frac{\alpha' z_v (1 + \mu_{tr} d_v) e^{-\mu_{tr} d_v}}{4\pi d_v^3}, \quad (2.13)$$

with the reduced extinction coefficient $\mu'_t = \mu_a + \mu'_s$, the reduced albedo $\alpha' = \mu'_s / \mu'_t$, the effective transport coefficient $\mu_{tr} = \sqrt{3\mu_a \mu'_t}$, and with d_r and d_v the distances to the two point light sources.

The dipole model allows for rapid visualization of scattering materials [JB02]. However, when using the model one has to be aware about its approximating assumptions: for diffusion theory to hold, the material has to be highly scattering, homogeneous, and isotropic. The dipole model further assumes that the material fills a semi-infinite half-space. This neglects the influence of variation in object geometry and, more severe, does not properly model effects that occur for thin layers of material. This introduces an error that increases as the real material thickness reaches the same order of magnitude as the mean free photon path length $\ell \approx 1/\mu_{tr}$.

Donner and Jensen [DJ05] present an extension of the dipole approximation to a multi-pole model that remedies this problem, modeling an arbitrary number of thin layers of infinite extent. They show that modeling material thickness is particularly important for human skin where the mean free path ℓ is in the order of one millimeter.

2.2.3 3D Photography

In recent years, many approaches were developed to capture and represent the appearance of real-world objects, allowing for novel renderings of the objects. These techniques are summarized by the term *3D photography*. It loosely describes methods that capture object shape and appearance from images.

Image-based and related techniques aim to reproduce the object’s appearance directly from these images with little or no knowledge of its shape. They simulate its visual appearance by interpolating camera images of the object. Representatives of this techniques are the light field approaches [LH96, GGSC96, MRP98, WAA⁺00], storing view-dependent image data of an object. They re-sample input images to achieve a ray-centered parameterization. These approaches can be seen as purely *data-driven*, as they visualize objects by a smart recombination of the original sensor data, without making any assumptions about the actual object properties.

Other methods require a detailed model of the object and its material properties to predict its appearance under arbitrary illumination [SWI97, YDMH99, KM99,

HS99, MAA01, LKG⁺01]. The object model includes a geometric description as well as an analytic definition of its surface reflectance properties, often using BRDF representations. These techniques are called *parametric* or *model-driven* approaches, as they are based on parameterizable physical models

The image-based methods capture a large variety of effects, including refraction, subsurface scattering, self-shadowing, and inter-reflection. However, they show limitations when objects are observed from close distance or extreme angles, where the input data has to be interpolated or extrapolated, respectively. The model-driven class of techniques in principle leads to renderings of arbitrary fidelity. Nevertheless, these methods can only display effects that have been explicitly modeled, which in particular poses heavy requirements at the acquisition procedure.

Consequently, hybrid approaches have become very popular, combining image-based techniques with explicit modeling. They parameterize an image-based model on an impostor geometry, which generally improves the conditioning of the image reconstruction step.

2.2.4 Hybrid Appearance Models

The lumigraph [GGSC96] is the first method to combine an impostor geometry (a visual hull of the object) with an image-based representation. Images of the object are stored as a dense light field [LH96], and the impostor geometry is used to improve light field interpolation, i.e., to minimize blooming and ghosting. To reduce the amount of image data, view-dependent texture mapping [PCD⁺97, DYB98, DTM96] uses simple geometry and sparse texture data. This method is extremely effective despite the approximate 3D shape, but it has some limitations for highly specular surfaces due to the relatively small number of textures.

Surface light fields [MRP98, WAA⁺00, CBCG02] are a more general and efficient representation because they parameterize the image data onto the object surface. Surface light fields can either be stored on accurate high-density geometry [WAA⁺00] or on coarse triangular meshes for objects with low geometric complexity [NSI99a]. Some techniques [NSI99b, CBCG02] aggressively compress the data such that the models can be rendered in real-time on modern graphics hardware. To improve the appearance of complex object silhouettes, surface light fields can be combined with view-dependent opacity data into *opacity light fields* [VPM⁺03]. Unstructured lumigraph rendering [BBM⁺01] is a very effective method for rendering both surface and opacity light fields.

Although surface light fields are capable of reproducing important global effects such as inter-reflections and self-shadowing, they only show the object under fixed lighting. To overcome this limitation, recent approaches have used surface reflectance fields. The reflectance field of an object is the radiant light from a

surface under every possible incident field of illumination. In practice, the reflectance field is sampled sparsely and interpolated during rendering. To further reduce the amount of data, most reflectance fields are acquired for a single view [DHT⁺00, HCD01, KBMK01]. For approximately planar geometry and diffuse surfaces the data can be compressed further by fitting a parametric function to the reflectance field [MGW01].

Goesele et al. [GLL⁺04] introduce a novel method that models volumetric radiance transport within objects using a hierarchical representation. They acquire scattering properties of an object using laser illumination. Tong et al. [TWL⁺05] separate surface reflectance into local and global effects, modeling local effects by data-driven modulation textures on incident and outgoing illumination, and describing global light transport by a fitted dipole diffusion approximation. They also use controlled laser illumination for acquisition. See Chapter 6 for more details on acquisition of subsurface scattering properties.

2.3 Human Skin Rendering

Properties of human skin have been measured and studied extensively in the computer graphics, biomedical, and computer vision communities. An excellent survey on the physiological and anatomical properties of skin and the state of the art in skin appearance modeling has been published by Igarashi et al. [INN05]. In this section we provide an overview of the relevant work in the area of face appearance modeling in computer graphics.

2.3.1 Analytic Face Appearance Models

The simplest reflectance model for skin is the Bidirectional Reflectance Distribution Function (BRDF) [NRH⁺77]. Dana et al. [DvGNK99] measured the BRDF of skin and made the data available in the Columbia-Utrecht Reflectance and Texture (CuReT) database (<http://www.cs.columbia.edu/CAVE/curet/>). Because the data was not collected *in vivo* it does not accurately reflect the properties of living skin. Marschner et al. [MWL⁺99, MWLT00] measured spatially uniform BRDFs on the foreheads of four subjects of different gender, race, and age. They observed that the BRDF of skin exhibits strong forward scattering in off-specular direction at grazing angles. In later work they augmented the uniform specular BRDF with a spatially-varying albedo texture [MGR00]. Their BRDF model was used in the human face project at Walt Disney Studios [Wil05]. That project, which had a key influence in the historical development of human face models (e.g., it was influential on the process developed for the *Matrix* sequels), was completed in 2001. Other researchers have estimated parameters of uniform analytic

BRDF models from photographs with static [BV99, PSQ03] or varying illumination [Geo03, FLS05]. Fuchs et al. [FLS05] cluster different parts of the face, such as cheek, forehead, and lips, into regions. They estimate a spatially-varying diffuse albedo but assume specularities to be constant within each region. The fundamental limitation of all these models is that they ignore subsurface scattering of light that is largely responsible for the soft appearance of facial skin.

Hanrahan and Krueger [HK93] were the first to model subsurface scattering in skin using a diffuse constant and a lighting-dependent single scattering term. They described scattering and absorption parameters of skin using the Henyey-Greenstein function and compute scattering with Monte Carlo simulations. A similar approach was used by Ng and Li [NL01]. To avoid expensive Monte Carlo simulations, Stam [Sta01] developed an analytic approximation to multiple subsurface scattering for skin with a rough surface. All of these models describe skin reflectance using a BRDF and ignore light scattering between different incoming and outgoing locations on the surface.

To describe the full effect of light scattering between two points on the surface one can use the Bidirectional Surface-Scattering Distribution Function (BSSRDF). Krishnaswamy et al. [KB04b] introduced a BSSRDF model for skin with spectral and biophysically-based parameters. The BSSRDF is eight-dimensional, making it very difficult to measure and represent.

Jensen et al. [JMLH01] estimated model parameters for the dipole approximation by measuring scattering of light in skin on a forearm *in vivo* using a white focused beam of light and a camera. We use their approximation for the subsurface scattering term of our skin model. To estimate the parameters of their model we developed a measurement device that can safely be used in faces (see Chapter 6).

Jensen and Buhler [JB02] develop fast rendering methods for this approximation, and Héry implemented the technique for use in feature films [Hér03, Hér05]. Both address the derivation of spatially varying subsurface scattering parameters from an albedo map. We use this technique in one instantiation of our general skin model presented in Section 8.1.2. Alternatively, we do not use spatially varying subsurface scattering coefficients and instead use a constant scattering term, modulated by a modulation texture.

2.3.2 Image-Based Face Models

Image-based models have provided highly realistic representations for human faces because they implicitly capture effects such as self-shadowing, inter-reflections, and subsurface scattering. Pighin et al. [PHL⁺98] use view-dependent texture mapping [DTM96] to reproduce faces under static illumination conditions. More recent efforts allow variations in lighting for static faces [GBK99, DHT⁺00], expressions [HWT⁺04], and real-time performances [WGT⁺05]. Debevec et

al. [DHT⁺00] present a process for creating realistic, relightable 3D face models by mapping image-based reflectance characteristics onto 3D-scanned geometry. In order to change the viewpoint, they use color-space analysis to separate the image data into specular and diffuse components that can be extrapolated to new viewpoints. While their method does consider the aggregate behavior of subsurface scattering, they do not model a specific diffusion parameter. Thus, unlike our approach, their method cannot produce correct subsurface scattering effects for closeup light sources or high-frequency spatially-varying illumination. Borshukov and Lewis [BL03] combine an image-based model, an analytic surface BRDF, and an image-space approximation for subsurface scattering to create highly realistic face models for the movie industry. Sander et al. [SGM04] developed a variant of this method for real-time face rendering on modern graphics hardware.

Image-based models come at the price of extensive data storage. They are difficult to edit, and the measurements are not related to any physical model of light-skin interaction. It is very difficult to simulate local light sources and to reproduce spatially-varying illumination [JGB⁺05], especially if shadows from a distant light disagree with shadows from a local light. An example of this would be a light source close to the nose – even though the skin is illuminated, an image-based method would predict zero reflection since a distant light from that direction would be blocked by the nose. In contrast to image-based methods, we measure a full BRDF at each surface point using a face-scanning dome similar to the LightStage by Debevec et al. [DWT⁺02], but with more cameras (16). The resulting parametric face reflectance model has lower storage requirements than a full surface reflectance field. In addition it is physically plausible, easy to render, and intuitive to edit.

2.3.3 Skin Color and Texture Models

Biophysical studies show that skin appearance is largely dependent on wavelength [INN05]. Angelopoulou et al. [AMD01] measured skin BRDF as a function of wavelength for a sample population of 22 people and noted that five Gaussian basis functions reproduce the data well. The study was limited to 0° incident and 4° reflection angle.

Tsumura et al. [TOS⁺03] used Independent Component Analysis (ICA) in color space to decompose images of faces into melanin and hemoglobin layers. To change skin color and transfer skin texture they adapted the texture synthesis method of Heeger and Bergen [HB95]. We use a similar approach to transfer texture (see Section 10.2). Cula et al. [CDMR05, CDMR04] collected the Bidirectional Texture Function (BTF) for different body regions of a sample population (http://www.caip.rutgers.edu/rutgers_texture/cvg/). They analyzed

and classified skin texture using histograms of image-pyramid features [CD02]. We use a similar analysis technique for albedo maps (see Section 10.2).

2.4 Point-Based Graphics

Many parts of this thesis employ point-based representations to describe geometric surfaces. In contrast to the traditional triangle-based surface representation in computer graphics, point-based graphics follow a sample-based modeling paradigm. Instead of approximating a surface by a continuous piece-wise linear triangle mesh, objects are represented by a dense set of unorganized surface samples. Each sample encodes position and surface attributes associated with a respective surface point. This representation is particularly suited when dealing with scanned surface data, as will be elaborated in Chapter 3. The lack of an explicit topological encoding, as present in a triangular representation, required the development of new algorithms for geometry processing and rendering in the past. See [KB04a] and [GP06] for an overview over the field of point-based graphics.

2.4.1 Surfel Representation

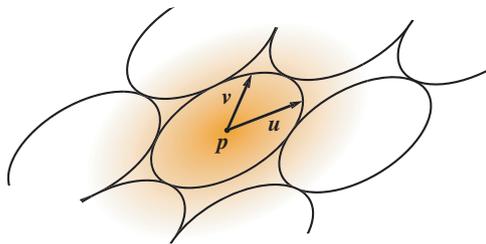


Figure 2.3: A surfel is defined by its position p and its tangential system (u, v) . u and v span a skewed coordinate system that defines the surfel's elliptical region of influence.

In this thesis, we use a *surfel* representation to describe surface samples. Surfels, as introduced by [PZvBG00], are points in \mathbb{R}^3 , augmented by additional attributes, such as normal, radius, and some color properties. In their representation, the normal indicates the surface orientation, whereas the radius can be understood as an approximate, circular region of influence.

Subsequently, [ZPBG01, ZPvBG02] presented a framework to render surfels based on Heckbert's elliptical weighted average (EWA) texture filtering [Hec89].

EWA surface splatting is a forward-mapping algorithm that allows for high quality, aliasing-free rendering of point sampled geometry. As an extension of the initial representation, [ZPvBG02] propose elliptical surface splats. In contrast to circular splats that use a position/normal parametrization, they are defined by a position p and two tangential vectors u and v , spanning a tangential coordinate system. u and v are explicitly allowed to be non-orthonormal. The

surfel is associated with an elliptical region of influence defined by the tangential system (see Fig. 2.3).

An implementation of the EWA surface splatting algorithm supporting elliptical surface splats is included in the publicly available point editing application Pointshop3D [ZPKG02, Poi]. Our surfel rendering implementations have been derived from this source code.

2.5 Unstructured Lumigraph Interpolation

Unstructured lumigraph rendering (ULR) is an image-based rendering algorithm introduced by [BBM⁺01]. It uses an impostor geometry and can therefore be accounted as a hybrid appearance model, as described in Section 2.2.4. For each point on the impostor geometry, ULR determines a weight vector w_i with blending weights for a set of input images associated with certain camera views. The weights are determined based on penalties P_i assigned to each input view. In [BBM⁺01], the penalties are a function of camera resolution and how close the camera is to the desired view. The algorithm is parameterized by an integer constant k with $k > 3$. This constant controls how many available camera views are actually blended together. The algorithm generates $k - 1$ positive w_i , all other weights are set to zero.

In this thesis, we repeatedly use ULR interpolation, not only to blend camera views but also as a general scattered data interpolation. We use a simplified reformulation of ULR that uses relative weights r_i , $0 \leq r_i < 1$, rather than penalties. Without loss of generality, let r_k be the k -th largest r_i . The algorithm determines intermediate weights

$$\tilde{w}_i = \begin{cases} \frac{r_i - r_k}{1 - r_i}, & r_i > r_k \\ 0, & r_i \leq r_k \end{cases}. \quad (2.14)$$

Note how the k -th largest relative weight is mapped to a \tilde{w}_i of zero. Normalization yields the final weights:

$$w_i = \frac{\tilde{w}_i}{\sum_i \tilde{w}_i}. \quad (2.15)$$

A key property of ULR interpolation is that for smoothly changing input weights r_i , the resulting weight vector changes smoothly as well. In the case of camera views to be blended across a surface, this particularly means that camera views blend in and out smoothly, as long as there are no discontinuities in the variation of r_i .

Geometry Processing

As our face model requires an explicit representation of the facial geometry, we investigated in suited 3D reconstruction techniques [Hub02, Sch03, SWPG05]. As shall be motivated in the remainder, raw 3D geometry acquired from real objects generally has to be cleaned before further usage. We developed a respective scan cleaning toolbox for point-sampled geometry that is described in this chapter [Hei04, WPK⁺04].

3.1 Motivation

With growing demand for realism in computer graphics and interactive techniques, we experience a steady increase in the geometric complexity of digital 3D surface models. Ab initio design of such shapes thus becomes increasingly time consuming and expensive. Most designers therefore rely on 3D scanning devices to acquire complex digital models from real-world objects. Accurate 3D acquisition also plays an important role in reverse engineering, rapid prototyping, bio-medicine, architecture, cultural heritage acquisition, or entertainment industry.

This diversity in application fields is reflected in a great variety of 3D imaging techniques: CT and MRI scanners are widely used in medical and engineering applications to acquire volumetric representations of real-world objects. Optical

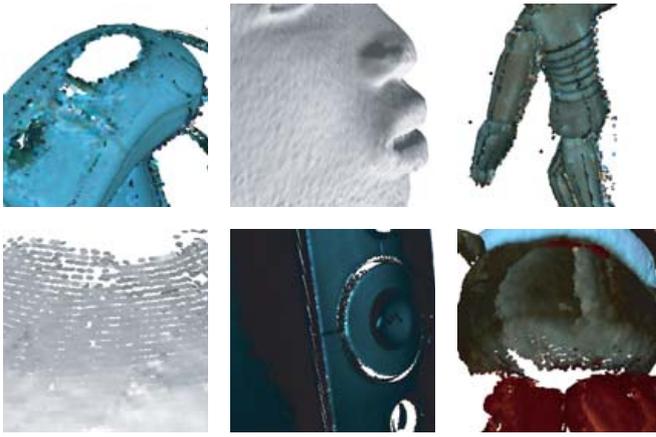


Figure 3.1: Typical artifacts of raw scanner data. Top Row: Holes due to sensor restrictions, noise, outliers. Bottom Row: Low sampling density due to gracing sensor views, low sampling density at delicate surface details, and holes due to critical reflectance properties.

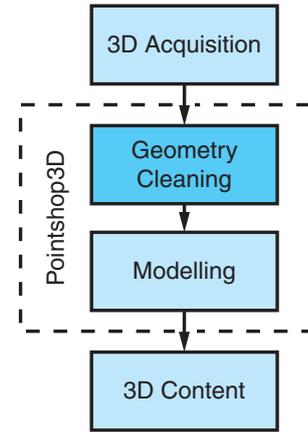


Figure 3.2: Geometry cleaning bridges the gap between 3D acquisition and higher-level modeling. We present an interactive cleaning toolbox implemented within Pointshop3D.

devices, such as laser-range scanners or structured light scanners, are primarily employed for surface and appearance acquisition.

This latter class of scanning devices typically produces a dense set of surface points, where each point samples a 3D position and possible additional attributes such as normal information, color, or material properties, which corresponds to a point-based representation as described in Section 2.4. Depending on the specific acquisition method, a number of scanning artifacts can occur as illustrated in Figure 3.1:

- ▮ Physical limitations of the sensor lead to noise in the acquired data set. Sample points can also be corrupted by quantization or motion artifacts. The latter occur when the scanned object moves during the acquisition process, a common problem when scanning humans or animals.
- ▮ Multiple reflections and heavy noise can produce off-surface points (outliers).
- ▮ Holes and under-sampling in the model surface occur due to occlusion, critical reflectance properties, constraints in the scanning path, or limited sensor resolution.
- ▮ Many scanners tend to create displaced geometry when the scanned object is textured.

The raw point cloud data produced by the scanner thus needs to be processed before subsequent modeling operations can be performed. Commercial scanners

are usually equipped with rudimentary scan cleaning software that uses built-in heuristics for outlier removal and noise reduction. These are often difficult to control as they are optimized for the specific scanner configuration. More sophisticated data processing can only be applied by exporting the acquired surface model from the proprietary scanner software, typically in the form of a triangle mesh. However, if the aforementioned data imperfections have not been successfully removed from the data set, the meshing process itself is fragile and can even introduce further artifacts. Therefore, post-processing of scanned data should be performed directly on the acquired point cloud, before sophisticated surface reconstruction algorithms or advanced modeling operations are applied.

To this end, we propose a purely point-based scan cleaning toolbox, consisting of a selection of user-guided tools that address the different scanning artifacts mentioned above. These include an eraser tool, low-pass filters for noise removal, a set of outlier detection methods, and various re-sampling and hole-filling tools. Since many scan artifacts are strongly coupled, these tools should be applied in an interleaved fashion. Identification of artifacts is difficult and often requires human interpretation. Therefore, user guidance is a necessary prerequisite to achieve optimal results. All algorithms are specifically designed to support rapid feedback during an interactive scan cleaning session.

The toolbox has been integrated as a plug-in into Pointshop3D, an open-source 3D editing tool for point-sampled surfaces [ZPKG02, Poi]. In combination with a 3D scanning front end, the plug-in bridges the gap between 3D acquisition and high-level shape and appearance modeling, thus providing in a single application a complete point-based content creation pipeline (Figure 3.2).

The remainder describes the toolbox and its constituents. Section 3.2 discusses design criteria of the toolbox, while Section 3.3 presents the underlying techniques.

Sections 3.4 and 3.5 present the resulting set of interactive tools. It is explained how the underlying techniques have been extended and combined to realize the different tools. The integration of the tools in a common user interface is presented.

After a scanned object has sufficiently been cleaned, automatic processing techniques can be applied, without suffering from scanning artifacts.

3.2 An Interactive Approach

The central motivation of the toolbox has been to open up modeling techniques to be used for the cleaning of raw scan data. The modeling tools make extensive use of basic techniques (Section 3.3), which are well-known in point graphics community or adapted from triangle based graphics, respectively.

The toolbox was designed to allow for the removal of typical scan artifacts, as depicted in Figure 3.1. In order to support an efficient scan cleaning process, three design goals have been pursued:

Predictability In order to allow a rapid workflow, it is important that each tool's effect is predictable under most circumstances. That is, if the user chooses a tool for a certain purpose, the outcome should meet the user's expectations.

Controllability The range of application must be well-controllable. Where possible, each tool should provide a set of parameters to tune its behaviour.

Intuitive Handling The tools should rest upon intuitive editing metaphors. Any parameters should correspond to meaningful traits.

Following those criteria, the goal has been to make the tools as powerful as possible. However, making a tool powerful usually implies the use of higher-level automatism which are likely to fail when applied to raw scanner data. This would contradict predictability. Increasing the number of parameters to make the outcome more controllable would lead to an unintuitive handling.

Accordingly, the final set of tools comprises operations of different complexity (see Figure 3.3). Simpler, more robust, tools allow for direct editing, especially in

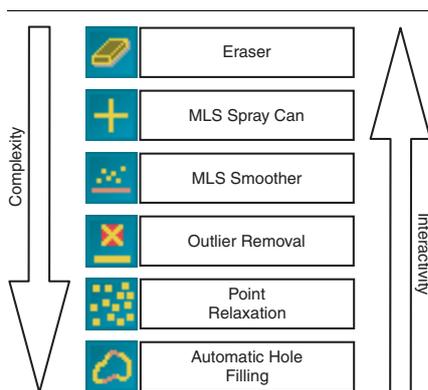


Figure 3.3: The toolbox contains tools of different complexity. Higher complexity goes with less interaction.

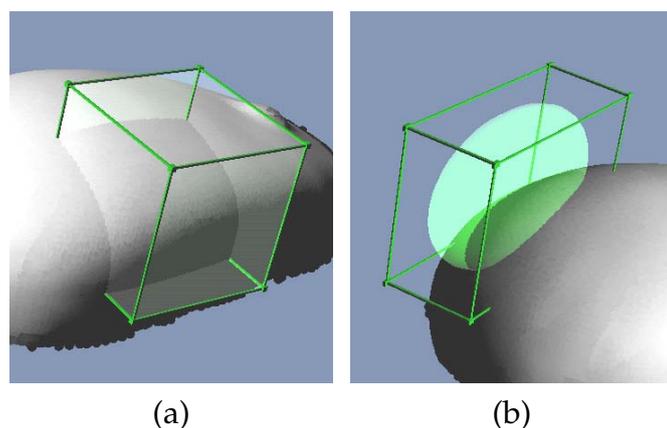


Figure 3.4: The volumetric brush. (a) A box shaped selector, following the object surface. (b) An ellipsoidal selector, freely positioned in space.

the presence of severe scanning artifacts. More complex and powerful tools can be applied at a later point in the scan cleaning process, when a certain sampling quality has already been achieved.

In order to address controllability, all tools provide an exhaustive set of parameters that can be set using the user interface. Each tool comes with a set of reasonable default parameters.

For most of the tools it makes sense to apply them locally. Consequently, they are defined to work on a set of selected surfels.

Pointshop3D provides a selection mechanism. However, the Pointshop3D selection tool requires a well-sampled surface and can not, e.g., select scattered points, as they frequently appear in real-world scans. Hence, a volumetric brush has been introduced to facilitate the selection of surfels in areas where no properly sampled surface exists. The brush, box shaped or ellipsoidal, can freely be moved in space, or alternatively follows the object surface (see Figure 3.4). By resizing and rotating the brush, its shape can be adapted to the local object geometry.

The brush is designed to follow the object surface even in poorly sampled regions. This is achieved by analyzing the depth values of all surfels visible around the mouse pointer. The brush's depth is set to a robust mean of the different depth values.

All tools that support the volumetric selection can be applied to the set of selected surfels. Alternatively, they are simultaneously applied to all points within the volumetric brush during navigation.

3.3 Underlying Techniques

The presented toolbox internally utilizes a set of basic geometric modeling techniques. This section describes the respective techniques and explains their adaptation to point clouds.

3.3.1 Search Data Structures

Dealing with point clouds, there is no explicit connectivity information. This means that all computations are based on spatial proximity between point samples instead of geodesic proximity between mesh vertices. In this section we present two data structures for fast nearest neighbor searches and range queries.

A very well known search data structure is the k -d tree (e.g. [Ben85, FBF77, AM93]). A k -d tree can be searched efficiently in $O(\log n)$, while it takes time $O(n \log n)$ to build it. Therefore, and because it is costly to maintain a k -d tree after an insertion, deletion or displacement of points, it is suitable for static data only. If the same point is queried more than once, it might be useful to cache the neighbors. In this case, a nearest-neighbor graph is built, storing the nearest neighbors for each point.

For querying dynamic data a hash data structure similar to [THM⁺03] has been used. The coordinates of an arbitrary point in space are mapped to a cell. If the

cell size is chosen smaller or equal than the maximal query range, all points within this range can be found by searching the adjacent cells to a query point, i.e. 27 cells have to be queried. Note that also k -nearest neighbor queries can be performed efficiently if a maximal range can be given. However, while insertion of a point can be done in $O(1)$, querying takes $O(q)$, where q is the maximum number of points in a cell. In practice, with a sufficient number of cells q will be small.

3.3.2 MLS Projection

To compute a smooth surface that approximates a set of scattered data points, Levin [Lev01] introduced a projection operator based on Moving Least Squares (MLS) optimization. Using this projection procedure, Alexa et al. [ABCO⁺01] presented a high quality rendering algorithm for point set surfaces. Because the MLS method is crucial for the following algorithms, we will briefly review it here.

Let P be an unstructured set of sample points. The MLS projection as used in the cleaning toolbox takes a point x in space and projects it onto a polynomial that locally approximates the underlying surface in the vicinity of x . This polynomial is computed by first fitting a reference plane H using weighted least squares optimization. The reference plane provides a local parameterization of the sample points, which is used in a second least squares fit to compute a bivariate polynomial approximation.

Both, the computation of the reference plane and the polynomial use a radially symmetric Gaussian weight function $w_i = e^{-\|x_r - p_i\|^2/h^2}$, where x_r is the projected point of x onto H and h is a scaling factor. Since w_i drops quickly with increasing distance, the least squares optimization is typically applied in a local neighborhood around the point of interest. The scaling factor h can either be a global constant or proportional to local sample spacing, estimate from a k -neighborhood as described in [PGK02].

3.3.3 Point Relaxation

In [Tur92], Turk uses particle simulation for resampling polygonal surfaces. Pauly et al. [PGK02] adapted this method to point-sampled surfaces.

To achieve a uniform distribution of the particles, neighbored particles are let repel each other. Every particle p exerts a force $f_i(p)$ on its neighbored particles p_i . The summation of all forces that act on a particle gives the resulting force. Finally, the new positions of the particles are computed by explicit Euler integration.

The presented work uses the same repulsion force f as in [Tur92, PGK02]:

$$f_i(p) = k(r - \|p - p_i\|) \frac{p_i - p}{\|p_i - p\|}, \quad (3.1)$$

where k is a force constant and r is the repulsion radius. For finding the nearest neighbors within the radius r the hash data structure described above is used.

After each iteration, the particles are projected back onto the surface by applying the MLS projection described above. In the cleaning toolbox, the particle simulation is performed locally for a selected region. To ensure that the selected surfels keep within this region, for each selected surfel its n nearest neighbors are computed, and the neighbors which are not selected are added to a list. While these surfels repel the selected surfels, their positions are fixed.

3.4 Outlier Removal

Erroneous points outside the object surface are outliers that have to be removed. However, it is hard to specify a general criterion to detect outliers, if the real object surface is unknown. Noise further complicates the detection of outliers. In many cases, the scan quality has to be judged by the user in order to tell a noisy surface point from an outlier.

We developed an interactive tool for outlier removal incorporating the user into the outlier detection. The tool provides three outlier classification heuristics that have to be weighted by the user to obtain an appropriate classification (see Figure 3.5). Outliers are finally removed by applying a threshold to the resulting outlier classification.

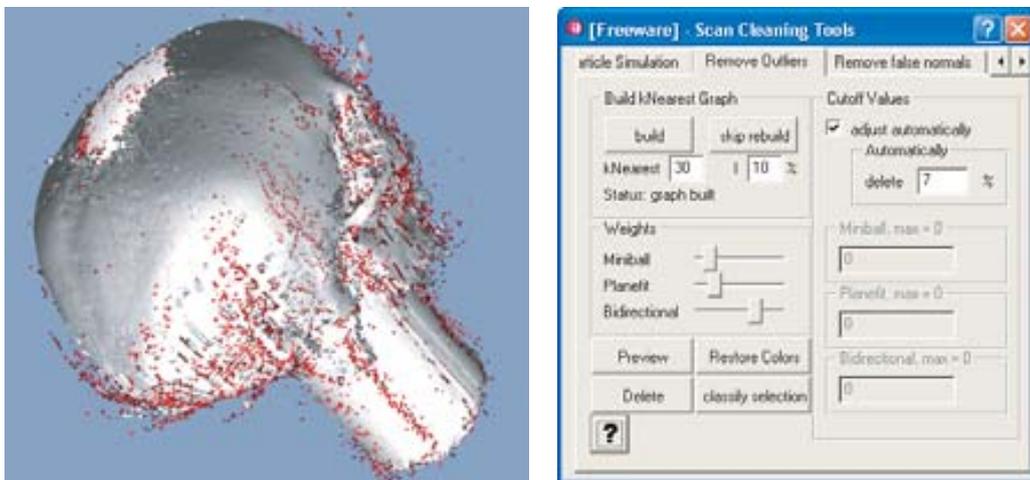


Figure 3.5: Outlier classification. The three classifiers can be weighted using the depicted sliders. Probable outliers, scheduled for removal according to the resulting classification and a given threshold, are rendered in red.

The threshold can be chosen manually. Alternatively, it is automatically set to discard a certain percentage of the points. Outlier classification can be confined to the volumetric brush.

3.4.1 Outlier Criteria

We now present the three underlying outlier criteria. All criteria deliver an estimator $\chi(\mathbf{p}) \in [0, 1]$ assigning the likelihood for a point sample \mathbf{p} to be an outlier. To prevent any bias from an intermediate surface representation, all criteria are based only on analysis of \mathbf{p} 's k nearest neighbors \mathcal{N}_p .

Plane Fit Criterion

An intuitive criterion is the point's deviation from a manifold approximating its neighbors. The *plane fit criterion* considers a plane H that minimizes the squared distances to \mathbf{p} 's neighbors:

$$\min_H \sum_{q \in \mathcal{N}_p} \text{dist}(\mathbf{p}, H)^2 \quad (3.2)$$

(see Figure 3.6). Let d be the distance of \mathbf{p} to H , and \bar{d} the mean distance of points from \mathcal{N}_p to H . The plane fit criterion is defined as

$$\chi_{\text{pl}}(\mathbf{p}) = \frac{d}{d + \bar{d}}. \quad (3.3)$$

Normalization by \bar{d} relates d to possible noise and surface deviations.

Instead of H , it would be possible to use higher-order approximations of \mathcal{N}_p . The plane-fit criterion has been chosen to achieve a maximum of robustness.

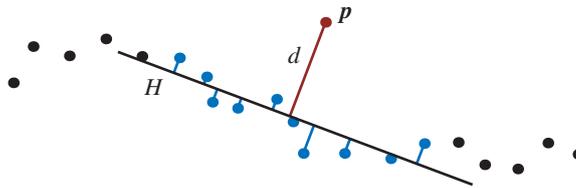


Figure 3.6: The plane fit criterion compares \mathbf{p} 's distance d to a least squares plane H with the average distance of its neighbors to H . \mathbf{p} 's k -neighbors are denoted in blue.

Miniball Criterion

A point comparatively distant to the cluster built by its k nearest neighbors is likely to be an outlier. This observation leads to the following criterion.

For each point p consider the smallest enclosing sphere S around \mathcal{N}_p [Wel91] (see Figure 3.7). S can be seen as an approximation of the k -nearest-neighbor cluster.

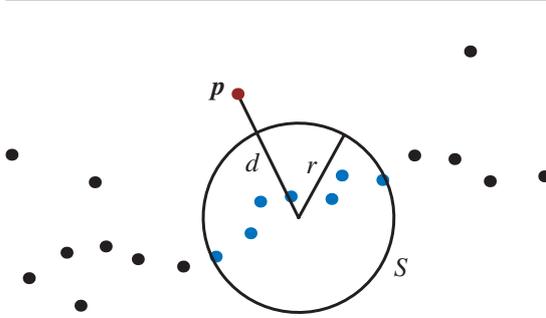


Figure 3.7: *The miniball criterion.* A miniball S approximates the cluster of p 's neighbors. The criterion compares p 's distance to S with the diameter of the sphere.

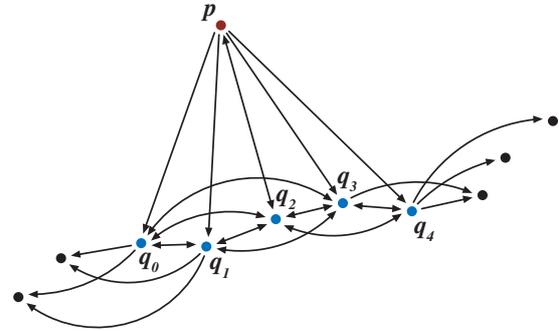


Figure 3.8: *Nearest-neighbor graph.* Depicted are the 5-nearest-neighbor relations for p and its 5-neighbors q_0, \dots, q_4 . Note that only q_2 shares a reciprocal neighbor relationship with p .

Comparing p 's distance d to the center of S with the sphere's diameter yields a measure for p 's likelihood to be an outlier. Consequently, the *miniball criterion* is defined as

$$\chi_{\text{mb}}(p) = \frac{d}{d + 2r/\sqrt{k}}. \quad (3.4)$$

Normalization by \sqrt{k} compensates for the diameter's increase with increasing number of k -neighbors at the object surface.

Nearest-Neighbor Reciprocity Criterion

This criterion is based on the following observation: Potential outliers draw their k -nearest neighbors from a larger vicinity than points in a well-sampled environment. In particular, a "valid" point sample q may be in the k -neighborhood of an outlier, but the outlier will most likely not be part of q 's k -neighborhood.

This relationship can be expressed by means of a directed graph G of k -neighbor relationships (see Figure 3.8). Outliers are assumed to have a high number of uni-directional exitant edges, i.e., asymmetric neighbor relationships. Consequently the criterion considers the ratio between uni-directional and bi-directional exitant edges in G .

The uni-directional neighbors are defined as $\mathcal{N}_{\text{uni}}(\mathbf{p}) = \{\mathbf{q} \mid \mathbf{q} \in \mathcal{N}_{\mathbf{p}}, \mathbf{p} \notin \mathcal{N}_{\mathbf{q}}\}$, while the bi-directional neighbors build a set $\mathcal{N}_{\text{bi}}(\mathbf{p}) = \{\mathbf{q} \mid \mathbf{q} \in \mathcal{N}_{\mathbf{p}}, \mathbf{p} \in \mathcal{N}_{\mathbf{q}}\}$. The classifier is expressed as follows:

$$\chi_{\text{bi}}(\mathbf{p}) = \frac{\|\mathcal{N}_{\text{uni}}(\mathbf{p})\|}{\|\mathcal{N}_{\text{bi}}(\mathbf{p})\| + \|\mathcal{N}_{\text{uni}}(\mathbf{p})\|} = \frac{\|\mathcal{N}_{\text{uni}}(\mathbf{p})\|}{k}. \quad (3.5)$$

3.4.2 Classification

The final outlier classification is computed using weights w_1, \dots, w_3 , $\sum_i w_i = 1$, interactively defined by the user:

$$\chi(\mathbf{p}) = w_1\chi_{\text{pl}}(\mathbf{p}) + w_2\chi_{\text{mb}}(\mathbf{p}) + w_3\chi_{\text{bi}}(\mathbf{p}). \quad (3.6)$$

As all outlier criteria are based on the k -nearest-neighbor graph, χ_{pl} , χ_{mb} , and χ_{bi} are computed once and cached during the computation of χ .

Depending on the scanning technique, outliers may occur in small clusters. In this case, χ_{pl} and χ_{mb} tend to fail to detect the clustered outliers correctly. In order to make them suitable for clustered outliers, a maximum cluster size l can be defined by the user. Subsequently, all k -nearest-neighbor queries will discard the first l neighbors, returning the $(l + 1)$ st to $(l + k)$ th neighbor instead. This effectively increases the robustness against clustered outliers while maintaining the basic functionality of the outlier criteria.

3.4.3 Performance Evaluation

When applying the outlier removal tool, the three different elementary outlier criteria show to be differently suited depending on the situation (see Figure 3.9). The plane fit criterion is best suited to detect outliers in a noisy reconstruction of a smooth surface. It produces poor results around small features and creases, as the orientation of the fitted plane becomes instable. The miniball criterion proved to be more robust, even around high-frequency details, but in contrast to the plane fit criterion it shows a poor outlier detection for points that hover close to a smooth surface.

In comparison with the previous two, the criterion based on nearest-neighbor reciprocity shows the most robust outlier classification. It is equally sensitive around smooth and detailed regions. However, it consistently yields erroneous outlier classifications around manifold borders (see Figure 3.10). Obviously, each criterion is advantageous in different situations. The outlier removal tool allows to confine the outlier detection to certain areas for the model and to weight the criteria according to the local situation.

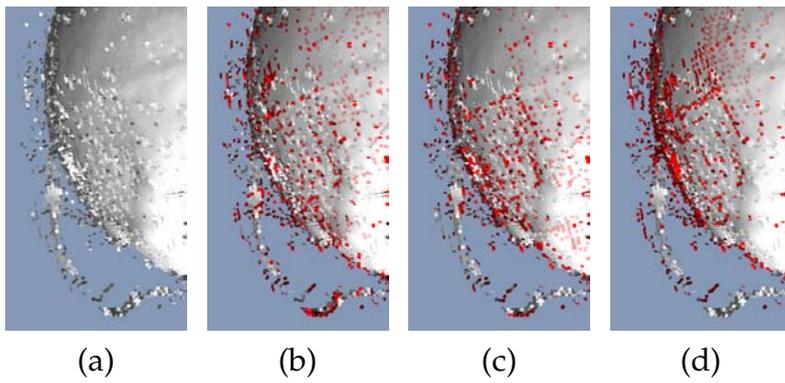


Figure 3.9: Three different outlier classifiers. Potential outliers marked in red. (a) Raw scanned geometry. (b) Classification using the miniball criterion. (c) Plane-fit criterion. (d) k -nearest-neighbor graph criterion. All criteria were thresholded to classify 7% of the surfels as outliers.

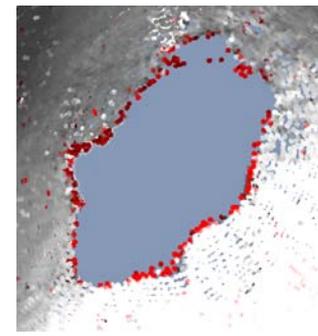


Figure 3.10: The nearest-neighbor outlier criterion performs poor around manifold borders.

3.5 Hole Filling and Smoothing

After outliers have been removed, there is usually still inherent noise in the surface samples. Additional holes further complicate surface analysis and reconstruction.

Noise can be removed by applying a spatial low-pass filter to the 3D point data. Alternatively, noise can implicitly be handled during a surface reconstruction stage: [ABCO⁺01, CBC⁺01, OBA⁺03] yield a smooth surface by approximating the sample points. However, most automatic surface reconstruction algorithms fail in the presence of severe noise.

In the past, various hole-filling techniques have been proposed to address this problem. These methods mostly use implicit representations to define the underlying surface. Verdera et al. [VCBS03] extend image inpainting techniques to 3D surfaces by solving anisotropic partial differential equations defined on the surface. Carr et al. [CBC⁺01] and Ohtake et al. [OBA⁺03] exploit the extrapolation properties of radial basis functions to fill regions of sparse sampling. Davis et al. [DMGL02] propose a method that applies a diffusion operator on the signed distance field of an incomplete triangle mesh.

The remainder of this section presents noise removal and hole filling tools of the presented toolbox.

3.5.1 MLS Smoother

Smoothing is an elementary editing operation. It can be used for noise reduction, to smooth-out high-frequency details, such as small artifacts like spikes and ripples, or to soften creases created during the editing process. Various smoothing operators have been proposed, partly with feature-preserving properties.

Given the unpredictable quality of input data, it had been decided against locally adapting filters, as they still tend to amplify scanning artifacts. Instead, a simple, more robust, filter based on MLS projection has been implemented, leaving the treatment of features to the user's control by confining the operation to the volumetric brush selection.

The *MLS smoother tool* works by shifting point positions towards the corresponding MLS surface. For each point p , its MLS projection p' is computed. A user-adjustable blending parameter α defines how far p is to be moved towards its "smoothed" position p' . The point is finally set to

$$p_{\text{smoothed}} = (1 - \alpha)p + \alpha p'. \quad (3.7)$$

An associated normal is filtered analogously, blending the original normal with

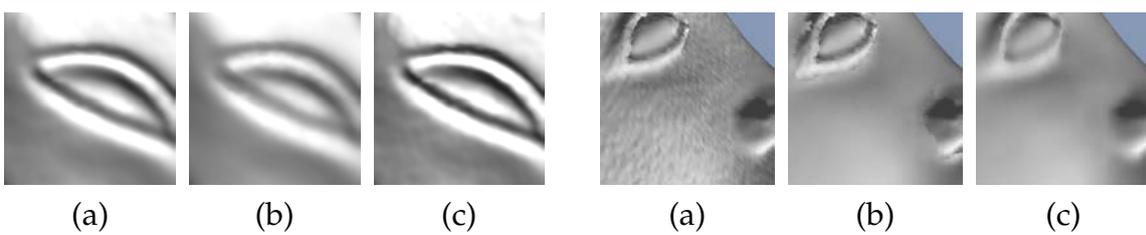


Figure 3.11: The *MLS smoother tool*. (a) Fine surface details. (b) Smoothing with $\alpha = 0.8$. (c) Detail enhancement for $\alpha = -0.75$.

Figure 3.12: Selective noise removal using the *MLS smoother*. (a) Noisy input surface. (b) Smoothing of a sub-set of surfels, excluding high-frequency details. (c) Subsequently, global smoothing of the model.

the normal of the MLS surface. Parameterization of the MLS kernel function, as described in Section 3.3.2, allows the user to adjust the low-pass characteristic of the MLS projection.

An additional user parameter D allows to attenuate the tool's effect towards the selection border. Within distance D to the border, α is weighted by a blending polynomial to vanish at the border. A point's distance to the border is defined as the distance to its nearest neighbor outside the selection.

α is usually set to values within $[0, 1]$, corresponding to strong, or no smoothing, respectively (see Figure 3.11(b)). Alternatively, following the concept of USM filtering [Jai89], one may set α to negative values, corresponding to a detail (and

noise) enhancement (see Figure 3.11(c)). This is a useful feature, however, for larger absolute values of α , surface self-penetration can occur.

3.5.2 Point Relaxation

Scanned models may contain regions of uneven point distribution. While some editing operations may change the point distribution directly, raw scan data will be unevenly sampled wherever point samples are missing due to scanning artifacts. Merging of depth-maps also produces an uneven point distribution. However, a uniform distribution of the surfels is often required to guarantee a hole-free rendering of the surface.

To achieve an even distribution of the surfels a particle simulation as described in Section 3.3.3 is employed. The attributes of the relaxed surfels, such as the color, are interpolated from the attributes of the k -nearest original surfel neighbors.

The particle simulation can also be used to close small holes, as the repelling force will distribute the surfels over uncovered areas.

3.5.3 MLS Spray Can

Complementary to the eraser, the *MLS spray can tool* was introduced in order to fill small holes in the geometry. It randomly creates points inside the brush volume and projects them onto the MLS surface in the brush's vicinity.

A projected point p is added to the surface whenever the surrounding splat coverage is below a certain threshold. The local coverage is estimated by determining the ratio between the average distance \bar{d} of p to its k -neighbors and the mean splat radius \bar{r} of its neighbors. p is added if

$$\frac{\bar{r}}{\bar{d}} < 1. \quad (3.8)$$

Consequently, the MLS spray relies on valid splat radii. When importing a model, initial splat radii are computed using a local surface analysis as proposed in [Pau03], based on a Voronoi diagram of the point cloud.

If a new point is added to the surface, its normal is adopted from the MLS surface. All other surfel attributes, e.g. color and reflectance properties, are determined by interpolating attributes from neighboring surfels.

Application of the spray can tool may result in a roughly uniform point distribution (see Figure 3.14). Eventually, the point distribution has to be relaxed using the point relaxation tool.

3.5.4 Volumetric Diffusion

While the MLS spray can tool introduced above is very effective for filling small holes, it still remains a tedious process to create a complete watertight model when larger and more complex holes occur in the acquired point cloud. This is frequently the case, however, as line-of-sight constraints, difficult surface reflectance properties, or extensive noise- and outlier removal, can lead to a highly incomplete representation of the model surface (see also Figure 3.16). As previously stated, many automatic hole filling algorithms exist.

We extend the volumetric diffusion method by [DMGL02] to point-sampled models by replacing the distance field estimation of [DMGL02] by an MLS projection step as proposed in [PKKG03]. The distance field is computed on a regular 3D grid that encloses the model surface (see Figure 3.13 (a)). At each grid point we

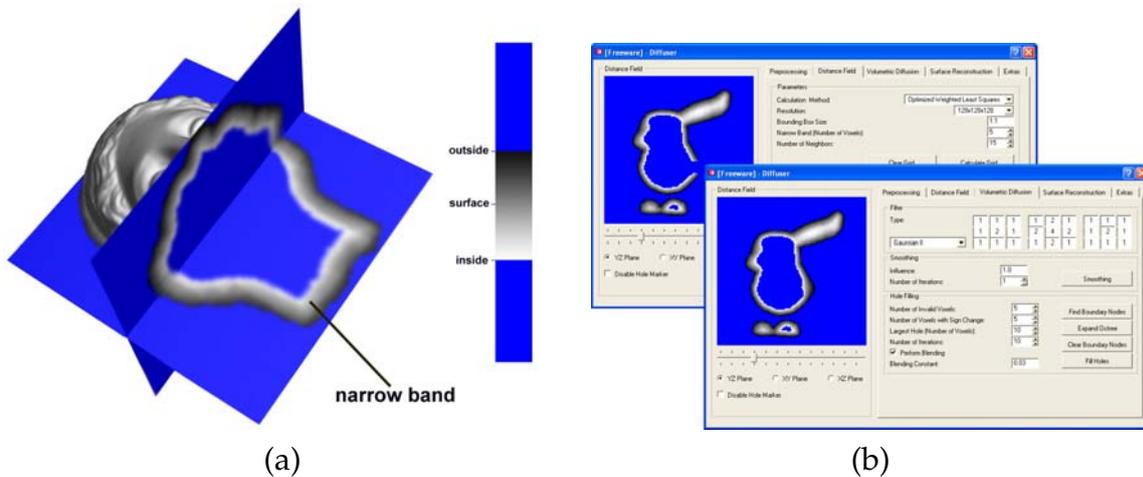


Figure 3.13: *Volumetric diffusion. (a) Slices of the distance volume reveal the narrow band. (b) The user interface of the automatic hole filling tool allows to fine-tune the algorithm. The volumetric representation can be previewed before surface reconstruction.*

compute the signed distance to the MLS surface defined by the given input point set. To efficiently represent this volumetric grid an octree data structure similar to [FP03] is used. This method makes use of binary location codes to address octree cells, allowing for fast point location and efficient neighborhood queries. A similar representation, particularly suited for GPU implementations, would have been the hashed octree [Sig06], which is particularly suited for GPU implementations.

Memory and computation costs are further reduced by only representing the distance field in a narrow band around the surface, similar to level set methods [OS88]. Holes in the distance field are detected using the classification method of [DMGL02]. Distance values on the boundary of holes can then be extrapolated

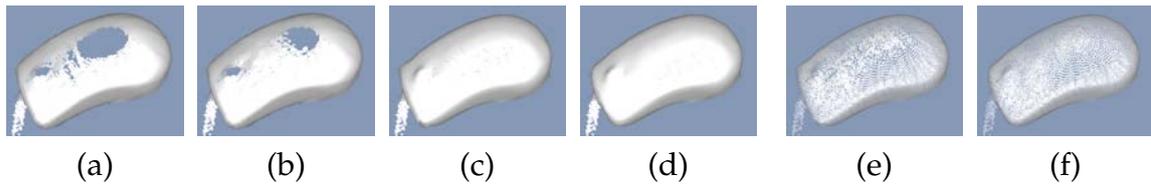


Figure 3.14: *Manual hole filling using the MLS spray can tool. (a) A poor scan of a computer mouse, containing a hole in the surface. (b,c) Gradually filling the hole using the MLS spray can. (d) Point relaxation improves the point distribution. (e,f) Versions of (c) and (d) with reduced splat radii to reveal the point distribution.*

by applying an iterative convolution operator until all holes of a user-specified size are filled. More details on this diffusion process can be found in [DMGL02].

To convert the distance field back to an explicit point-sampled representation the tool either applies a contouring method similar to marching cubes [LC87], or uses a particle simulation as described in Section 3.3.3. In the latter case, the MLS projection that keeps the particles on the surface is replaced by a projection based on gradient descent that moves particles to the zero set of the signed distance field. Normals of the newly generated points can also be directly estimated from the distance field gradient.

The user interface supports fine-tuning of the algorithm (see Figure 3.13 (b)). Though, using the default parameters the automatic hole filling tool is robust and easy to use.

3.5.5 Performance Evaluation

The hole filling and smoothing tools have extensively been used to clean various models acquired with different scanning technologies. They have been used with models acquired by a CyberWare[®] laser range scanner, a single-shot structured light scanner by 3Q Technologies Ltd., and a phase-shift structured light scanner similar to the system presented in [SWPG05]. This section shows some exemplary situations as they occur during the model cleaning process.

The general experience is that the simpler, more interactive tools are typically used at the beginning of the cleaning process, whereas the more complex, semi-automatic tools are applied toward the end of the procedure.

It turns out that the simpler tools are often used in combinations to achieve a desired effect. Figure 3.14 shows how the MLS spray can tool and point relaxation are used to manually fill a hole in a surface.

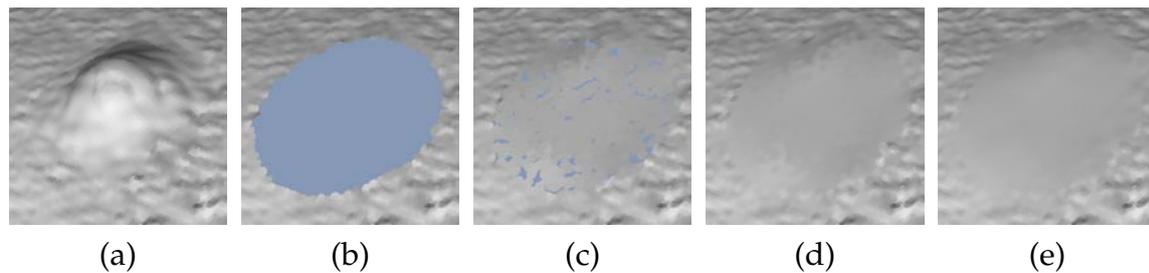


Figure 3.15: *Removal of an undesired bump. (a) Close-up of the original data. (b) The eraser is used to stamp out a hole. (c) Using the MLS spray can, the hole is filled. (d) Point relaxation redistributes points. (e) Locally applying the MLS smoother, attenuating its strength towards the border of the hole. Note the smooth transition of the novel surface to the noisy surrounding.*

A similar combination can be used to remove undesired bumps from a surface. Figure 3.15 shows how the eraser, the MLS spray, point relaxation, and the MLS smoother work together to remove a bump from a surface.

In combination with the selection tool, the MLS smoother can also be used to smooth selected surface parts while preserving details, see Figure 3.12.

The robustness of the automatic hole filling tool has been tested using a structured light scan of a furry toy reindeer (see Figure 3.16). Fur is one of the most difficult materials to be scanned with optical methods. Consequently, the scan shows severe noise and a lot of outliers. Outlier removal leads to a very sparse object reconstruction. However, as shown in Figure 3.16, the hole filling tool is still capable of producing a water-tight model. Only above the top of the model, the volumetric diffusion had to be constrained in order to get a closed surface.

Figure 3.17 shows an application of the automatic hole filling tool to a scan with largely varying sampling density. The model has been scanned by INSIGHT [INS] at the British Museum, London.

3.6 Discussion

This chapter presented a cleaning toolkit for the post-processing of raw scanner data. It is entirely based on point-based modeling techniques, which are given at hand in the form of simple, interactively controllable, tools. We introduced the underlying techniques and discussed the design principles leading to the presented set of tools.

The tools include an eraser tool, low-pass filters, and various re-sampling and hole-filling tools. We propose three different outlier criteria that are incorporated

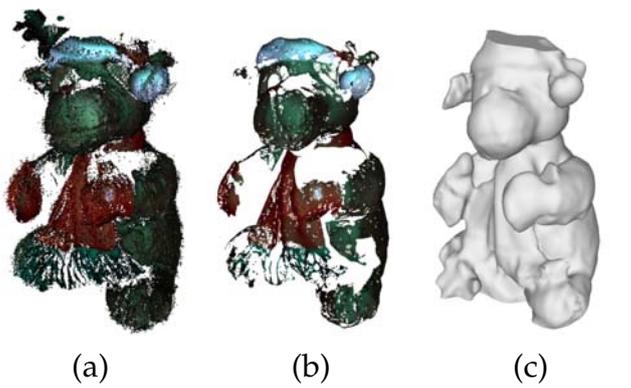


Figure 3.16: *Robustness of the volumetric diffusion tool. (a) The furred object surface produces severe noise and outliers. (b) After the outlier removal, only little object points are left. (c) The volumetric diffusion tool still reconstructs a water-tight model.*

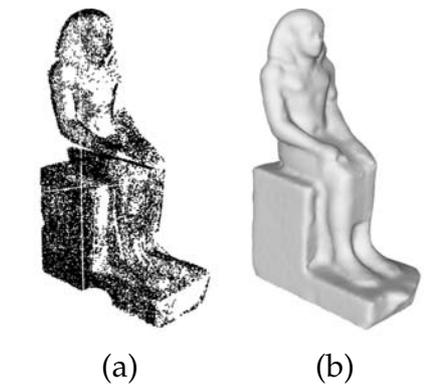


Figure 3.17: *Egyptian sculpture scanned at the British Museum. (a) Input scan with varying sampling density. (b) Application of the volumetric diffusion tool.*

in an outlier detection tool. An adaption of the volumetric diffusion algorithm to point-sampled data is used to build an automatic hole-filling tool.

We evaluated the toolbox, cleaning various objects acquired with different scanner technologies. It proves to be versatile and well-adaptable, as the tools can interactively be re-combined depending on the situation. Most operations are robust against sampling artifacts and do not impose any topological constraints on the data. Future experiences will show whether the toolbox has to be extended. Possible extensions may be additional filter tools, or the integration of texture synthesis into the MLS spray can. As a Pointshop3D plugin the toolbox rounds off a point-based work flow for the processing of scanned 3D surface data. The plugin is publicly available with an increasing number of users.

At an early stage of this project, the cleaning toolbox has been used to post-process acquired face scans. It has been ideally suited, because our initial appearance representation is entirely point-based, as described in Chapter 4. Later on, existing tools asked for an alternative, triangle-based representation. As discussed in Section 7.2, this forced us to drop the purely point-based scan cleaning in favor of a triangle-based cleaning procedure. Note, however, that in contrast to the presented cleaning toolbox, the triangle-based procedure described in 7.2 is less versatile and requires substantial pre-filtering by the acquisition software of the employed 3D scanner.

Surface Reflectance Field Evaluation

A central part of this thesis is the choice of a suited appearance representation, cf. Section 2.2. Independent from the choice of an image-based or an analytic representation, we are targeting an acquisition device that acquires the reflectance field of a face, and performs a 3D reconstruction of the facial surface. (The device is described in the next chapter.) In fact, this kind of data already represents a *surface reflectance field*, as introduced in Section 2.2.3. Consequently, the question arises whether this representation is directly suited for our purposes, in particular, for editing and animation of a human face model. However, there has been no previous work that successfully animated surface reflectance fields.

In general, animating an image-based or hybrid appearance representation requires a scheme to evaluate the image-based data set to simulate varying object deformations. A key question is how to preserve the visual appearance of the deformed object surface, that is, how to preserve the perceptual impression of material properties under different lighting conditions.

We developed a method to render surface reflectance fields under arbitrary deformations, which is presented in this chapter. We develop a shading scheme that aims to preserve the appearance of object materials during deformation. Our method uses a local parameterization of the impostor geometry that enables arbitrary warps. Our shading method uses this local parameterization to approximately preserve the spatially varying BRDFs of the undeformed object. We present a cache-optimized shading strategy to minimize computation time. The

technique is applicable to other hybrid representations that contain an impostor geometry.

4.1 Previous Work

Although it is an important aspect for many practical applications, the animation of image-based data has received very little attention in the literature. Wood et al. [WAA⁺00] describe arbitrary deformations on a surface light field and produce plausible renderings of the deformed model. However, their method does not deal properly with the diffuse component of the surface color, and it only works for purely reflective isotropic BRDFs. Feature-based light field morphing [ZWGS02] morphs two light fields into each other, based on the concept of ray-correspondencies. The method requires substantial user input to specify corresponding feature polygons between the two objects, and it is not applicable to the general animation setting. Both methods work only for static illumination. Furukawa et al. [FKIS02] presented a scanning system to capture objects and spatially varying BRDFs, also called Bidirectional Texture Functions (BTFs) [DvGNK99]. They use tensor product expansion to compress the BTF data and show results with surface deformations. Their system is the first to render deformations of a relightable, image-based object representation. However, they rely on a tight impostor geometry to support the BTF representation. Thus, the object geometry has to be acquired with a range scanning device. Moreover, their paper does not explicitly address appearance preservation under non-uniform, skewed deformations.

Our approach supports the animation of surface reflectance fields mapped onto *approximate geometry*. Our method allows to place the objects in new environments with arbitrary illumination, including dynamically changing lights. We carefully analyze the conditions that have to be met to preserve the appearance of the object (Section 4.3.1), and we present a novel method to approximately preserve spatially varying BRDFs during deformations (Section 4.3.2). We discuss the limitations of our approach (Section 4.7) and show results using objects that are difficult to handle for image-based approaches, including objects with specularities, transparency, and self-shadowing (Section 4.5).

Prior to building our own acquisition device, we performed first experiments using the surface reflectance field data acquired by the 3D photography system of Matusik et al. [MPN⁺02]. The system acquires reflectance fields for over 400 views using cameras, turntables, and a rotating array of lights. The acquired data also includes view-dependent opacity information, called the *opacity reflectance field*, which has not been part of our evaluation. The impostor geometry is the visual hull of the objects.

4.2 Overview



Figure 4.1: *Example reflectance image.*

A surface reflectance field consists of a large set of image-based reflectance data of an object, in conjunction with an impostor geometry used for rendering. The reflectance data is given by a collection of *reflectance images*. A reflectance image is a stack of camera images showing the object from the same viewpoint under varying directional illumination (Figure 4.1). A surface reflectance

field consists of reflectance images from many viewpoints around the object. Our test datasets contain approximately 400 reflectance images, each with 60 high dynamic range images of the object under different directional illumination.

Rendering a surface reflectance field can be understood as a two step procedure. First the reflectance images are used to compute an image of the object under the new illumination for each viewpoint (Figure 4.2). These images are then

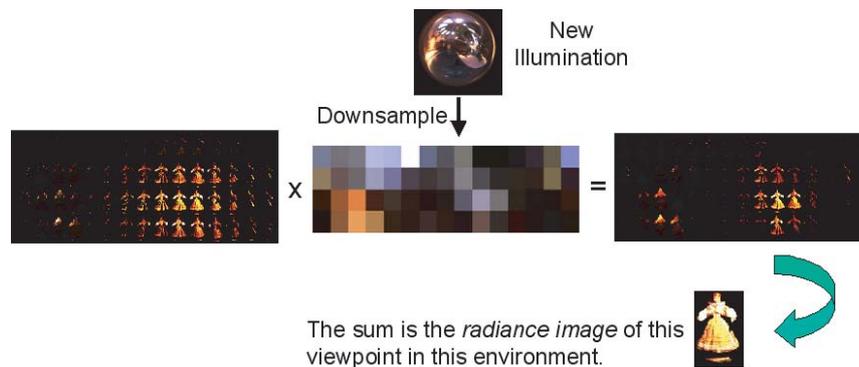


Figure 4.2: *Relighting using a reflectance image. The environment map of the new illumination is downsampled to the resolution of the reflectance image. Images are multiplied with the corresponding environment map color and added up to yield an image of the object under new illumination.*

rendered together using the impostor geometry with unstructured lumigraph interpolation [BBM⁺01], see Section 2.5. Our impostor geometry is the visual hull of the object that can easily be determined from observed silhouette images. However, our method is independent of the choice of impostor geometry.

We animate and deform a surface reflectance field (SRF) by first applying a 3D warp to the impostor geometry. In order to preserve the warped object’s appear-

ance, it becomes necessary to blend the reflectance images individually for each point of the impostor geometry. We developed a new look-up function to evaluate the reflectance images of the warped SRF. This function depends on the warp and is used to shade the warped impostor geometry during image generation (Figure 4.3). Shading a point on the impostor surface requires to apply this look-up

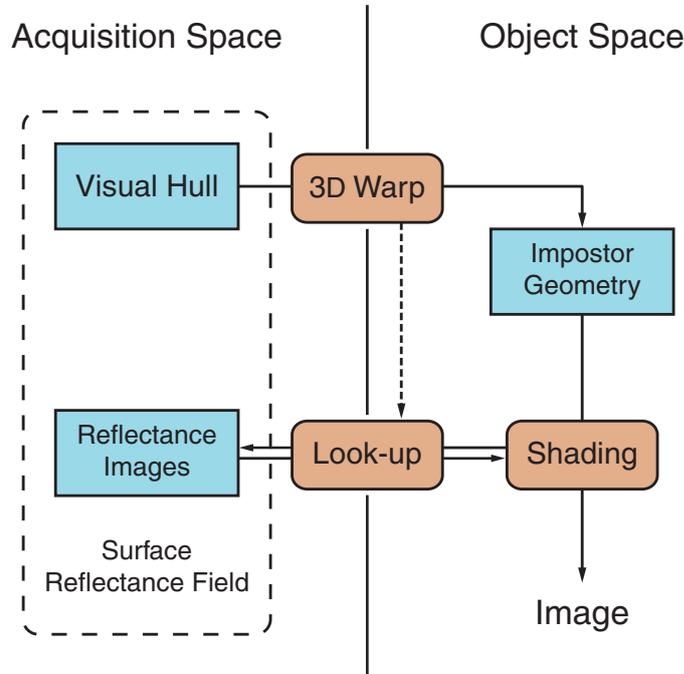


Figure 4.3: Rendering a deformed surface reflectance field requires switching between reference frames. The acquired impostor geometry is warped into object space. During the shading process, queries to the reflectance images must be transformed back into the acquisition frame using a local parameterization of the impostor geometry.

function to the surface point, the viewing ray, and the incident light direction. Once warped back into acquisition frame, the point can be shaded by blending the reflectance images according to the mapped light direction and viewing ray.

In the remainder of this paper we develop this look-up scheme. We discuss some shading issues and present our implementation based on a point-based impostor geometry. Section 4.5 shows some examples of deformed and animated surface reflectance fields.

4.3 Deformable Surface Reflectance Fields

Neglecting global light transport, a surface reflectance field can be understood as a discrete sampling of the object’s BRDFs, knowing neither the exact location of the

surface, nor its normals. Ignoring wavelength and time, a BRDF is a scalar function $\text{BRDF}(\mathbf{l}, \mathbf{v})$, describing the fraction of light that is reflected in direction \mathbf{v} as the surface is lit from direction \mathbf{l} . Similar to the BRDF notation, we use $\text{SRF}(\mathbf{p}, \mathbf{l}, \mathbf{v})$ to denote the reflectance of the SRF for an impostor point \mathbf{p} , a light direction \mathbf{l} , and a viewing direction \mathbf{v} . This relationship is fundamental for our analysis of appearance preservation during surface reflectance field animation.

The initial step in animating an SRF is to deform the impostor geometry. In this section, we assume the impostor warp is defined by a differentiable warp function

$$\Psi: \mathbb{R}^3 \rightarrow \mathbb{R}^3. \quad (4.1)$$

Shading an object includes queries to the surface reflectance field to determine the object's local reflectance. If the object is warped, the shading operation has to

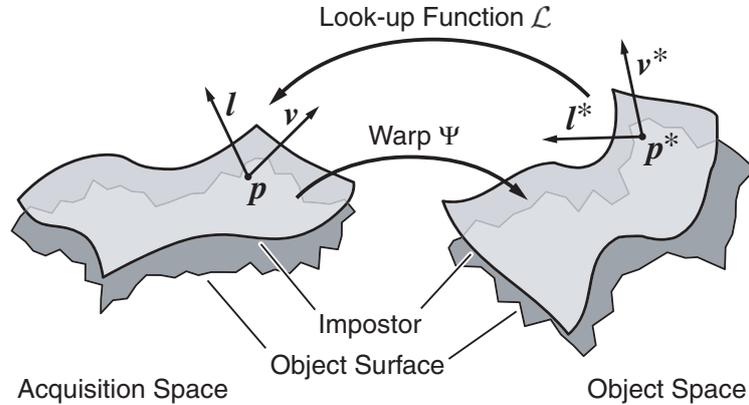


Figure 4.4: During rendering, reflectance queries must be mapped into acquisition space. Given lighting and viewing direction \mathbf{l}^* and \mathbf{v}^* for an impostor point \mathbf{p}^* , an appropriate surface reflectance field query $\text{SRF}(\mathbf{p}, \mathbf{l}, \mathbf{v})$ with \mathbf{p} , \mathbf{l} , and \mathbf{v} in acquisition space has to be found. The look-up function $\mathcal{L}: (\mathbf{p}^*, \mathbf{l}^*, \mathbf{v}^*) \mapsto (\mathbf{p}, \mathbf{l}, \mathbf{v})$ should be appearance preserving.

map a query $(\mathbf{p}^*, \mathbf{l}^*, \mathbf{v}^*)$ in object space to a query $(\mathbf{p}, \mathbf{l}, \mathbf{v})$ in the original acquisition space (Figure 4.4). By intuition, the required mapping approximately follows the inverse warp Ψ^{-1} .

However, applying Ψ^{-1} to the lighting and viewing directions is not necessarily appearance preserving, especially in the case of non-uniform deformations. Section 4.3.1 develops a mapping \mathcal{L} to perform this operation, leading to a new impostor parameterization presented in 4.3.2. The final shading process is discussed in section 4.3.3.

4.3.1 Approximate BRDF Preservation

In general, a mapping from object space to SRF acquisition space that preserves all aspects of the object’s appearance can not exist. This is due to the lack of exact object geometry and material properties in the SRF representation. Both would be needed to allow a prediction of complex non-local effects such as self-shadowing and inter-reflections.

Thus, we make the simplifying assumption, that the observed object can be described completely by its local BRDFs. We develop a look-up function that tries to preserve the main characteristics of the original object BRDFs. We do not aim at modeling changes in the BRDF due to the deformation of the material’s microstructure. This would require precise knowledge of the structure, e.g. its microfacet distribution. Instead, we want to map the original BRDF to the deformed surface.

The desired look-up function is a mapping

$$\mathcal{L}: (\mathbf{p}^*, \mathbf{l}^*, \mathbf{v}^*) \mapsto (\mathbf{p}, \mathbf{l}, \mathbf{v}). \quad (4.2)$$

We enforce appearance preservation by imposing three conditions:

1. Suppose the viewing ray in object space $\mathbf{p}^* + s\mathbf{v}^*$, $s \in \mathbb{R}$ intersects the warped object geometry at a point \mathbf{q}^* (see Figure 4.5). Then the viewing

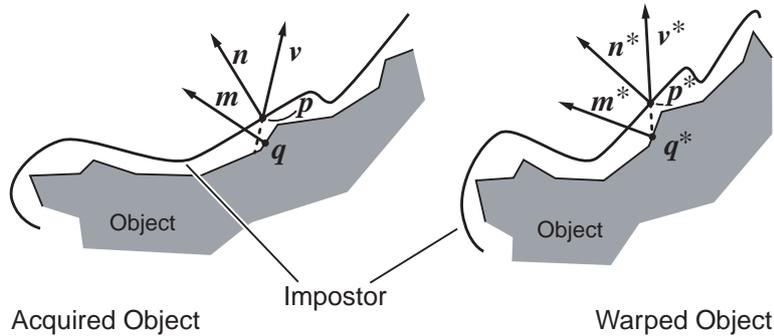


Figure 4.5: A viewing ray $\mathbf{p}^* + s\mathbf{v}^*$, $s \in \mathbb{R}$, in object space intersects the warped object at \mathbf{q}^* . The corresponding viewing ray in the surface reflectance field’s acquisition frame should intersect the object at the original object point \mathbf{q} . However, the impostor geometry of the SRF is typically not the same as the actual object geometry, so \mathbf{q}^* and \mathbf{q} remain unknown. Instead, we are using $\mathbf{p} = \Psi^{-1}(\mathbf{p}^*)$ as the origin of the viewing ray $\mathbf{p} + t\mathbf{v}$, $t \in \mathbb{R}$ in acquisition space.

ray in acquisition space $\mathbf{p} + t\mathbf{v}$, $t \in \mathbb{R}$ should intersect the original object at the corresponding point $\mathbf{q} = \Psi^{-1}(\mathbf{q}^*)$ to ensure the reflected light to originate from the same surface point with the same BRDF. Unfortunately, this

condition can not be guaranteed, since the true object geometry is typically not known, i.e., the exact location of the points q and q^* can not be determined. However, it is a reasonable approximation to force the two viewing rays to intersect corresponding points of the impostor geometry, since the impostor point p is very likely close to q . This can be achieved by choosing

$$p := \Psi^{-1}(p^*). \quad (4.3)$$

2. l and the surface normal m in q have to enclose the same angle as l^* and the normal m^* in q^* . This is a necessary condition to retain the reflectance characteristics of the object, as for example, the shape of the reflectance lobes of the corresponding BRDF. The same condition applies for v and v^* . Preserving the shape of the reflectance lobes requires to preserve the angle between l^* and v^* as well (see Figure 4.6).

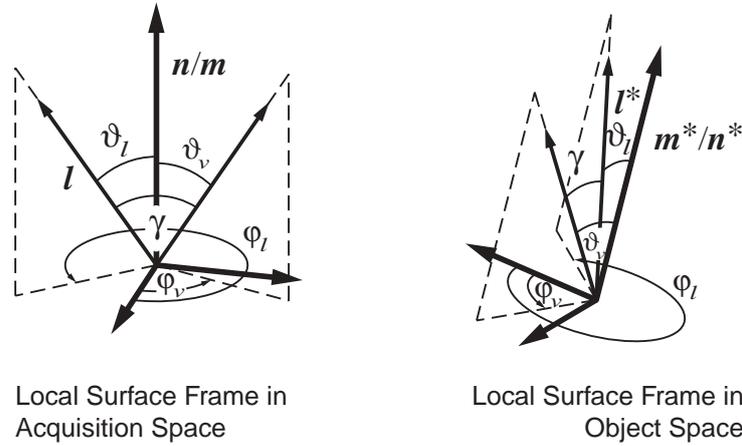


Figure 4.6: The bi-directional reflection distribution function (BRDF) describes the fraction of light from l that is reflected toward v . The spatial characteristics of a BRDF can be preserved by preserving the angles ϑ_l , ϑ_v , and γ when changing to acquisition space. Anisotropic materials require the preservation of the azimuthal orientations φ_l and φ_v relative to the surface frame. ϑ_l and ϑ_v are relative to the surface normal m . In an SRF, m is not exactly known. Instead, the impostor normal n is taken as an approximation.

3. In order to preserve the effect of anisotropic object BRDFs, l and v should have the same azimuthal orientation relative to the object surface as l^* and v^* on the warped object. (See Figure 4.6.)

As p immediately follows from (4.3), the mapping $(p^*, l^*, v^*) \mapsto (l, v)$ remains to be found. This can be rewritten as a locally affine mapping

$$\mathcal{L}_{p^*}: (l^*, v^*) \mapsto (l, v) \quad (4.4)$$

of lighting and viewing directions in the vicinity of \mathbf{p}^* . Note, that $\mathcal{L}_{\mathbf{p}^*}$ is a function of \mathbf{p}^* .

According to the second condition, $\mathcal{L}_{\mathbf{p}^*}$ has to be angle preserving. By convention, \mathbf{l} , \mathbf{v} , \mathbf{l}^* , and \mathbf{v}^* are unit vectors, so, $\mathcal{L}_{\mathbf{p}^*}$ needs to be length preserving as well. This implies that $\mathcal{L}_{\mathbf{p}^*}$ is an isometry, which means that the effect of \mathcal{L} on \mathbf{l}^* and \mathbf{v}^* can be described by a rotation or a reflection, respectively, as a function of location \mathbf{p}^* . In the remainder of this paper, we assume $\mathcal{L}_{\mathbf{p}^*}$ to be a rotation. The special case of a reflection can be handled similarly and is left out for simplicity.

Using this observation and interpreting (4.3) as a translation of \mathbf{p}^* by $\Psi^{-1}(\mathbf{p}^*) - \mathbf{p}^*$, the total effect of \mathcal{L} on a local setting around \mathbf{p}^* can be expressed as a *rigid transformation*, translating the point \mathbf{p}^* onto \mathbf{p} while rotating lighting and viewing directions.

Condition 2 further restricts \mathcal{L} to rigid transformations that map the warped object normal at \mathbf{q}^* onto the normal at the original point \mathbf{q} . We do not know the exact object geometry. Instead, we refer to the impostor normal in \mathbf{p} as an approximation of the true surface normal. Consequently, the searched \mathcal{L} maps the tangential plane of the warped impostor point \mathbf{p}^* onto the tangential plane in \mathbf{p} .

According to condition 3, \mathcal{L} should be chosen to preserve \mathbf{l}^* 's and \mathbf{v}^* 's orientation inside the tangential plane. The following section presents a local impostor parameterization that enables one to track the transformation of the local tangential frame in order to find a mapping \mathcal{L} that fulfills the conditions.

4.3.2 Local Impostor Parameterization

We augment the impostor geometry with a local parameterization that allows us to determine the look-up function \mathcal{L} at each impostor point \mathbf{p}^* . The parameterization is independent from the geometric representation and can be applied to triangular meshes as well as to point sampled geometry. Depending on the representation, \mathbf{p}^* can be a mesh vertex or a non-connected surface point, respectively.

This section develops the set of parameters that are stored at every point \mathbf{p}^* (see Table 4.1).

Given an arbitrary warp $\Psi: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, finding an explicit inverse function Ψ^{-1} is impossible in general. Constraining the SRF deformation to cases where an explicit inverse warp function exists would be too limiting. Consequently, another design goal for the parameterization was to determine \mathcal{L} without using a closed form of Ψ^{-1} .

As shown above, \mathcal{L} can be decomposed into a translational part and a rotation. The translation moves the warped impostor point back to its original position (see (4.3)). Storing the original position \mathbf{p}_0 in each impostor point \mathbf{p}^* allows the application of the translation without using Ψ^{-1} .

<i>Parameter</i>	<i>Description</i>
\mathbf{p}_0	Original position of the impostor point \mathbf{p}^* in acquisition frame.
\mathbf{r}_0	Orientation R_0 of the tangential system of \mathbf{p}^* in the acquisition frame, stored as a Rodrigues vector.
(\mathbf{u}, \mathbf{v})	Local tangential frame of \mathbf{p}^* .

Table 4.1: The local impostor parameterization allows for the determination of the look-up function \mathcal{L} during rendering. The parameterization consists of a set of vectors $(\mathbf{p}_0, \mathbf{r}_0, \mathbf{u}, \mathbf{v})$ in \mathbb{R}^3 , stored at each point of the discretized impostor geometry. \mathbf{p}_0 and \mathbf{r}_0 are fixed, while \mathbf{u} and \mathbf{v} have to be adapted whenever the geometry is deformed.

The rotational part $\mathcal{L}_{\mathbf{p}^*}$ (see (4.4)) aligns the warped tangential frame of the impostor point with the corresponding tangential system in the acquisition frame (see Figure 4.7). To avoid the application of Ψ^{-1} , the tangential orientation R_0 in the acquisition frame must explicitly be stored in \mathbf{p}^* . R_0 is defined by the rotation matrix $(\mathbf{u}_0, \mathbf{v}_0, \mathbf{n}_0)$, built by the tangential system in acquisition space. \mathbf{u}_0 and \mathbf{v}_0 must be orthogonal and can be arbitrarily chosen. To minimize its memory footprint, R_0 can be stored as a Rodrigues vector

$$\mathbf{r}_0 = k \tan \frac{\vartheta}{2}, \quad (4.5)$$

given R_0 as a rotation of angle ϑ around an axis \mathbf{k} . This definition provides a minimal (i.e., three-dimensional) parameterization of R_0 . Rodrigues vectors are closely related to quaternions [Alt89]. The main difference to quaternions is that Rodrigues vectors contain a singularity for 180° rotations, which can be avoided as R_0 can be arbitrarily chosen in our context.

Additionally, a tangential coordinate system (\mathbf{u}, \mathbf{v}) is attached to \mathbf{p}^* . While \mathbf{p}_0 and R_0 remain unchanged during the deformation, (\mathbf{u}, \mathbf{v}) is subject to the same warp as the geometry: After a deformation by Ψ , (\mathbf{u}, \mathbf{v}) is set to $\Psi((\mathbf{u}_0, \mathbf{v}_0))$. Applying the warp function Ψ to a local tangential system is a standard problem in differential geometry. The tangential system (\mathbf{u}, \mathbf{v}) has to be mapped by taking directional derivatives of Ψ . In our implementation, we use central differences to approximate

$$\mathbf{u} = \frac{\partial \Psi}{\partial \mathbf{u}_0} \quad \text{and} \quad \mathbf{v} = \frac{\partial \Psi}{\partial \mathbf{v}_0} \quad (4.6)$$

by

$$\begin{aligned} \mathbf{u} &= \frac{1}{2\varepsilon} (\Psi(\mathbf{p}_0 + \varepsilon \mathbf{u}_0) - \Psi(\mathbf{p}_0 - \varepsilon \mathbf{u}_0)), \\ \mathbf{v} &= \frac{1}{2\varepsilon} (\Psi(\mathbf{p}_0 + \varepsilon \mathbf{v}_0) - \Psi(\mathbf{p}_0 - \varepsilon \mathbf{v}_0)). \end{aligned} \quad (4.7)$$

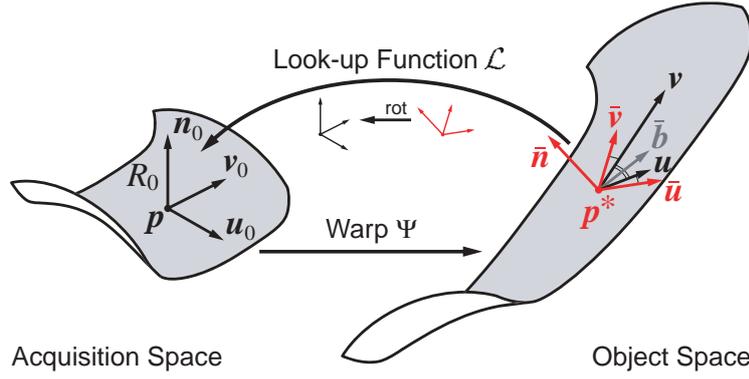


Figure 4.7: For each point on the impostor geometry, the local impostor parameterization provides two coordinate systems: The orthogonal tangential system R_0 in \mathbf{p}_0 in the acquisition frame and its warped counter-part on the rendered impostor geometry in object space, defined by \mathbf{p}^* and (\mathbf{u}, \mathbf{v}) . During rendering, \mathbf{p}_0 , R_0 , and the orthogonalized tangential system $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ are used to determine the appearance preserving back-projection of the SRF query.

for a small $\varepsilon > 0$. In particular, this only restricts Ψ to be differentiable in a vicinity of the impostor geometry.

Knowing R_0 and (\mathbf{u}, \mathbf{v}) , the rotation $\mathcal{L}_{\mathbf{p}^*}$ can easily be reconstructed during rendering. Let

$$\bar{\mathbf{n}} = \frac{\mathbf{u} \times \mathbf{v}}{\|\mathbf{u} \times \mathbf{v}\|} \quad (4.8)$$

and

$$\bar{\mathbf{b}} = \frac{\mathbf{b}}{\|\mathbf{b}\|} \quad \text{with} \quad \mathbf{b} = \frac{\mathbf{u}}{\|\mathbf{u}\|} + \frac{\mathbf{v}}{\|\mathbf{v}\|} \quad (4.9)$$

be the normal and the normalized bisecting vector of the warped tangential system spanned by \mathbf{u} and \mathbf{v} , respectively (see Figure 4.7). Then,

$$\begin{aligned} \bar{\mathbf{u}} &= \frac{1}{\sqrt{2}}(\bar{\mathbf{b}} + \bar{\mathbf{b}} \times \bar{\mathbf{n}}) & \text{and} \\ \bar{\mathbf{v}} &= \frac{1}{\sqrt{2}}(\bar{\mathbf{b}} + \bar{\mathbf{n}} \times \bar{\mathbf{b}}) \end{aligned} \quad (4.10)$$

build an orthogonalized tangential system that minimizes the squared angular differences between corresponding basis vectors of (\mathbf{u}, \mathbf{v}) and $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$.

Using this orthogonalized system, the rotation $\mathcal{L}_{\mathbf{p}^*}$ is given by

$$\mathcal{L}_{\mathbf{p}^*}(\mathbf{x}) = R_0 (\bar{\mathbf{u}}, \bar{\mathbf{v}}, \bar{\mathbf{n}})^\top \mathbf{x}. \quad (4.11)$$

Note that by choosing $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ as proposed, $\mathcal{L}_{\mathbf{p}^*}$ is an approximation of tangential orientation preservation in the sense of condition 3 of the previous section.

This approximation may introduce a slight rotation of the principal axis of an anisotropic BRDF when the surface is sheared. The rotation may be counter-intuitive depending on the direction of the shear. In general, the effect of a shear

on a real-world material effectively changes the original BRDF, depending on its microstructure. As the microstructure is not known, we decided to use the original BRDF. Thus, a tangential rotation of certain characteristics of the anisotropic BRDF can not be avoided.

It may not be surprising that \mathcal{L} turned out to be a rigid transformation. For rigid object transformations and cases where impostor geometry and object surface coincide, rotating the reflectance data according to the inverse object transformation is the appropriate choice. Consequently, this technique is used in many applications, e.g., for image-based BRDF measurements [War92, MWLT00, LKG⁺01], for bump mapping, for BTFs [DvGNK99, FKIS02] and for various other texturing techniques. All these techniques assume relatively accurate impostor geometry.

We assume arbitrary, in particular non-uniform, skewed deformations of an approximate impostor geometry that is different from the real object surface. The central result of our analysis is that in these cases again a rigid transformation of the impostor surface frame meets the requirements of appearance preservation best. Our framework allows the derivation of \mathcal{L} for arbitrary deformations, using the local impostor parameterization. As the resulting transformation is rigid, material properties—including anisotropic BRDFs—stay the same, even for skewed object transformations.

4.3.3 Shading

Once \mathcal{L} was determined, an impostor point can be shaded as described in Section 4.2. However, there are some issues that should be considered when lighting a deformed surface reflectance field.

When using an environment map to light an SRF, it needs to be filtered according to the spatial resolution of the reflectance images representing the SRF (see Figure 4.1). Using an unfiltered environment map may lead to aliasing artifacts, if the map contains details that are finer than the spacing of the light sources used to acquire the SRF.

During rendering of a deformed SRF, \mathcal{L}_{p^*} is applied to all lighting directions incident to p^* . This corresponds to a rotation of the environment map before evaluating the SRF with that environment. Thus, the rotated environment map needs to be re-filtered according to the reflectance field sampling. This operation would have to be performed for every surface point, which is impractical as filtering is an expensive operation for a non-uniform reflectance field sampling.

We present an alternative way to light a deformed surface reflectance field that comes without the need of re-filtering. Although an SRF is acquired using directional lighting, we can simulate lighting by *point light sources* with only little artifacts.

Lighting an impostor point \mathbf{p}^* by a point light source starts with the reflectance query $(\mathbf{p}^*, \mathbf{l}^*, \mathbf{v}^*)$ describing the viewing ray to \mathbf{p}^* , and the direction \mathbf{l}^* to the point light-source as seen from \mathbf{p}^* . Applying the look-up scheme $(\mathbf{p}, \mathbf{l}, \mathbf{v}) = \mathcal{L}(\mathbf{p}^*, \mathbf{l}^*, \mathbf{v}^*)$ transforms the query into acquisition space. $\text{SRF}(\mathbf{p}, \mathbf{l}, \mathbf{v})$ yields a reflectance coefficient that can be used to shade the impostor point \mathbf{p} with a color I_p :

$$I_p = \text{SRF}(\mathbf{p}, \mathbf{l}, \mathbf{v})A(\|\mathbf{p} - \mathbf{p}_l\|)I_l, \quad (4.12)$$

where I_l is the color of the point light source at \mathbf{p}_l , and $A(d)$ is the light attenuation factor depending on the distance to the light source. Using the impostor point \mathbf{p}^* instead of a point on the real object surface introduces an error in the incident lighting direction. But usually this error is small compared to the resolution of the SRF.

In contrast to environment mapping, this technique introduces SRF queries for lighting directions that are not present in the reflectance images. We rely on the SRF implementation to properly interpolate novel lighting directions, as discussed in Section 4.4.2.

The proposed approximation of point light-sources allows for dynamic lighting effects when animating surface reflectance fields. However, as our look-up scheme is ignoring global illumination effects, these effects may appear wrong on deformed objects.

Point light sources can also be used to implement environment mapping without the need to refilter the map for every impostor point. This can be done by properly sub-sampling the environment once for a dense set of directions. Then point light sources are defined at these directions, colored by the corresponding values of the environment. Provided the sub-sampling is dense enough, lighting the scene with these light sources leads to an appropriate reconstruction of the lighting environment. Note that light direction interpolation implicitly acts as a reconstruction filter. Thus, adaptively filtering the environment map is traded for interpolation. This alternative scheme is easier to implement and fits naturally into the framework of point light sources. For more flexible lighting effects, environment maps and point light sources can be combined.

4.4 Implementation

We implemented an animation system for surface reflectance fields similar to the system presented by Matusik et al. [MPN⁺02]. We improved the impostor representation, its rendering, and the reflectance field interpolation. Finally, we present a new shading scheme that speeds-up the rendering of warped SRFs.

4.4.1 Surfel Representation

Previous work on surface reflectance fields [MPN⁺02] used a point-based representation of the impostor geometry. In this approach the impostor was densely sampled by *surfels*, cf. Section 2.4.1. Although featuring high image quality, the surfel renderer as used by [MPN⁺02] is not suited for our purposes. It organizes the surfels in a static layered depth cube (LDC) tree, a fixed, axis aligned spatial data structure, that can not be warped. Instead, we use a variant of the Pointshop3D renderer [ZPKG02] which is an advanced version of EWA surface splatting [ZPBG01]. The Pointshop3D renderer is able to render unorganized point sets and thus can cope with dynamically changing geometry.

Deformations of the impostor geometry may lead to visible holes in the point-based reconstruction. For reasonable deformations, this can be avoided by deforming the splat geometry appropriately. We are using elliptical splats as described in Section 2.4.1, adapting their shape to the geometric deformation. Figure 4.8 shows the effect on a deformed patch of a surfel geometry.

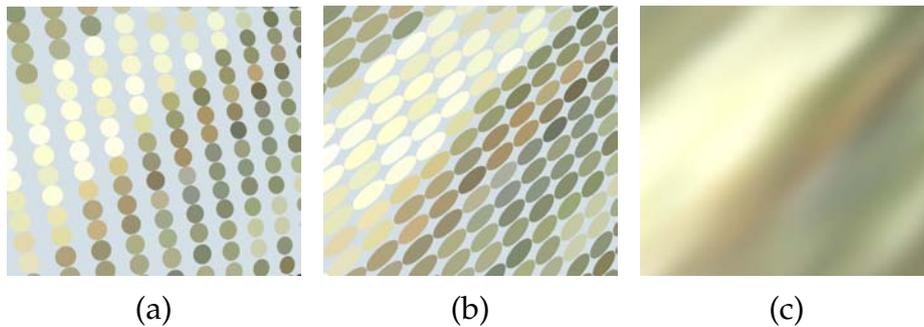


Figure 4.8: Close-ups of a surfel geometry. The surfels are placed on a square grid that was anisotropically stretched. (a) and (b) depict surface splats with reduced support to show the surfel distribution. (a) Circular splats can not follow the deformation. Thus, holes in the surface may appear. (b) Adaptively deformed elliptical splats preserve a closed surface and provide anisotropic texture filtering. (c) Blending the elliptical splats at their original size produces a smooth texturing.

When warping the impostor geometry, the warp has to be applied to each surfel’s tangential system by using an affine approximation of the local distortion. In our implementation, this approximation is determined similarly to (\mathbf{u}, \mathbf{v}) in section 4.3.2. This ensures surface coverage as long as the warp’s first derivative does not change excessively (see Figure 4.8).

The surfel’s tangential vectors \mathbf{u} and \mathbf{v} provide us directly with the corresponding parameters of our local impostor parameterization. By adding the surfel’s

original position p_0 and orientation r_0 to the set of surfel attributes, all impostor parameters required for the warped SRF rendering are encoded in the surfels.

4.4.2 Reflectance Image Interpolation

In general, each surface reflectance field look-up $\text{SRF}(p, l, v)$, must be interpolated from the fixed set of reflectance images. Depending on the viewing ray $p + sv$, $s \in \mathbb{R}$, the query must be handled by interpolating between the observing camera views of the SRF. View interpolation is a common problem in image-based rendering. Like Matusik et al. [MPN⁺02], we use unstructured lumigraph interpolation, see Section 2.5.

ULR interpolation has proven to be a very flexible interpolation scheme that can be controlled by an arbitrary penalty function, rating each camera's appropriateness. In our implementation, we chose the same penalty as in [MPN⁺02].

In previous work, the queried lighting direction l was always chosen from the set of acquired illuminations, sampling a stationary environment map. However, since our point light-source implementation produces arbitrary lighting directions, we have to interpolate reflectance queries between the acquired lighting directions as well.

This problem has been addressed by Koudelka et al. [KBMK01] who intersected the queried light ray with a triangulated hull of the acquisition light source positions. The intersected triangle yielded the three closest light sources that were consequently used to interpolate between three different lighting conditions. However, this does not work for arbitrary positions of the acquisition light sources. We also found that for slim triangles in the hull light source interpolation was not smooth enough.

Instead, we are also using ULR to interpolate between light directions. Treating the light sources of the acquisition stage as cameras at infinite distance enables us to use ULR for light source interpolation. In our implementation, this is done by using the same penalty function as for camera interpolation. Camera interpolation uses $k = 4$ to minimize ghosting, and light source interpolation uses a slightly larger $k = 5$ for smoother shading transitions.

Using ULR for camera selection affords consideration of occlusions in the scene. If the observed object point is not visible from a given camera, this camera should not be considered for interpolation. When interpolating between light sources, there are no visibility constraints, since shadowing effects are already captured in the acquired reflectance images. Moreover, taking occlusion into account would change the appearance of translucent or refractive materials. Visibility is determined using the method described in Section 7.3 on page 106.

4.4.3 Deferred Evaluation

Compared to the shading of static objects, substantially more reflectance images are involved during shading of a deformed model. This is because the rendered viewing rays tend to diverge in acquisition space when the impostor is warped (see Figure 4.9).

The number of reflectance images involved in the shading process increases with deformations. Since SRF data usually exceeds the conventional main memory size, most SRF look-ups must be read from disk. The time needed for disk access exceeds all other computation times by orders of magnitude. This makes SRF look-up performance a critical issue, especially when producing animated SRFs.

<i>Entry</i>	<i>Description</i>
$\text{REF}(S)$	Reference to the target surfel.
(i_c, i_l)	Camera index and light source index for the reflectance image look-up.
(x_c, y_c)	Look-up position on the camera plane.
(w_r, w_g, w_b)	Color weights.

Table 4.2: A shading operation represents a basic reflectance field look-up. It contains the depicted entries that are used to accumulate reflectance values in the surfel colors. A shading operation is executed by looking-up the reflectance image (i_c, i_l) at position (x_c, y_c) . The resulting rgb reflectance value is weighted by (w_r, w_g, w_b) and added to the color of surfel S .

We address this problem with two techniques: A smaller cache holds the most recently used blocks of data. Unfortunately, naively shading the geometry shows adverse cache-coherency. Thus, as a second technique, all shading operations are

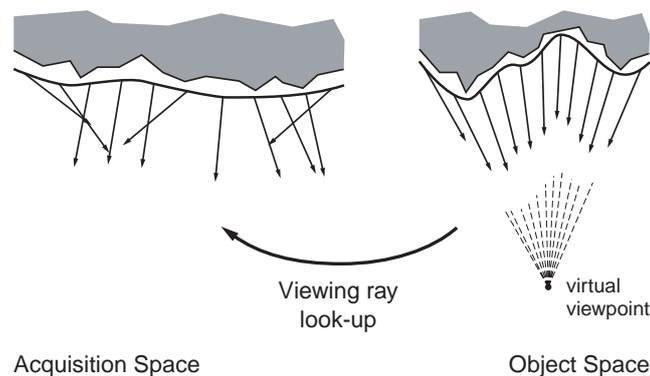


Figure 4.9: Viewing ray divergence due to back-projection.

decomposed into minimal *shading operations*, each containing a single SRF look-up. Table 4.2 shows their detailed structure. The execution of the shading operation is deferred until a larger number of operations has been collected. Sorting them before execution in a cache-optimal way decreases shading times by an order of magnitude. Moreover, due to the the cache coherent shading execution, the cache can be kept very small. In our system, it was sufficient to dimension the cache large enough to hold three reflectance images at the time. To facilitate the deferred shading, we accumulate all surfel colors first before rendering them to frame buffer.

4.5 Test Results

For our experiments we use the two models shown in Figure 4.10 that were scanned by Matusik et al. [MPN⁺02, MPZ⁺02]. The doll SRF is well suited for



Model	Viewpoints	Lights	Surfels	PCA Data
Doll	432 (360°)	60	532 935	4.4 GB
Mug	216 (180°)	44	757 760	1.8 GB

Figure 4.10: The presented models. Left: The surface reflectance field of a doll. Right: A half filled beer mug. The table shows the complexity of each data set.

the analysis of appearance preservation, as it contains many different materials. The mug SRF was chosen to show the limitations of our approach, as it contains refractive effects that can not be handled correctly by our technique.

The reflectance images of the models were captured using the high dynamic range (HDR) technique of Debevec et al. [DM97]. For each viewpoint and lighting condition, four 1360×1032 pictures with different exposure time were taken. Using up to 432 different viewpoints and 60 light directions, this produced a total of $432 \times 60 \times 4 = 103\,680$ images. Uncompressed this would correspond to 407 GB of data. All reflectance images were compressed using a PCA compression that achieved a reduction of at least a factor of 10. Figure 4.10 lists the compressed size of both models.

With these optimizations, the vast part of the rendering time is spent for the ULR computation. Our current implementation assumes arbitrarily placed cameras and light sources during the acquisition. Using knowledge about the scanner geometry or a spatial data structure to locate the k -nearest light sources or cameras would speed up the rendering times significantly. However, this optimization has not been performed, resulting in rendering times of about 10 minutes per frame for the doll data set with 5 point light sources on a 1 GHz Pentium III.

View extrapolation is an important issue when deforming surface reflectance fields, as parts of the objects may become visible from directions where no data was acquired before. The extrapolation is implicitly covered by the ULR interpolation. Figure 4.11 shows an example.

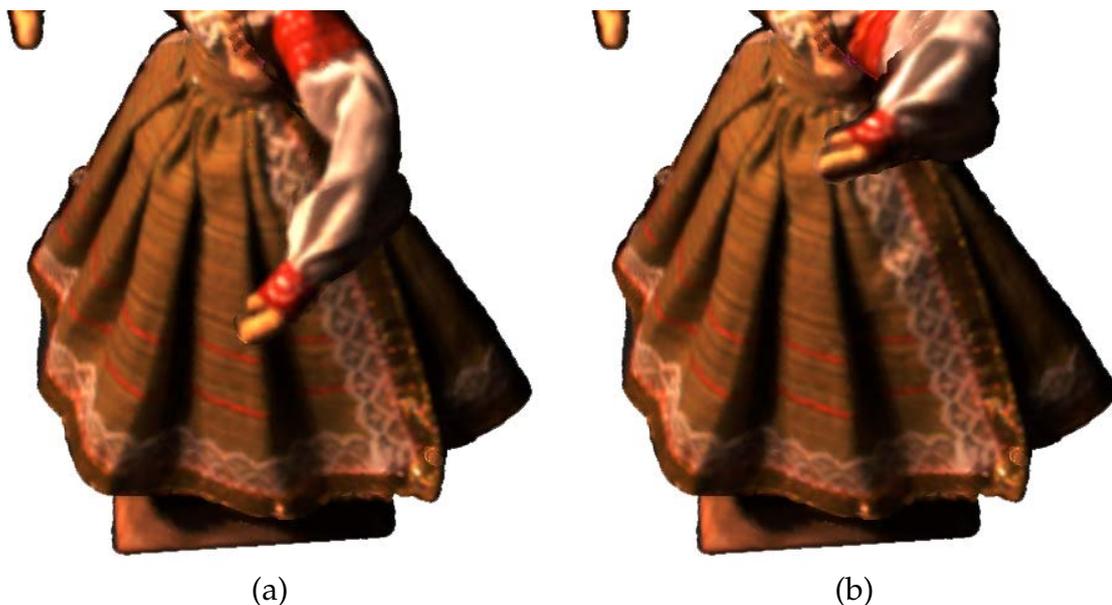


Figure 4.11: *View extrapolation. (a) Original doll data set. (b) Occluded parts of the skirt become visible as one arm is lifted. The model is lit by two point light sources. (a) is chosen to be an original view to force the views used for view extrapolation to be at maximum distance from the current view.*

Figure 4.12 shows the quality of appearance preservation at the example of the doll data set. Note the visual preservation of the different materials of the wooden base, the diffuse fabric, and the specular braid of the skirt. All scenes were lit by three colored point light sources (*red, green, and blue*) from static positions.



Figure 4.12: *Preservation of material properties* Left: *The original doll surface reflectance field.* Center, Right: *A sinusoidal warp is applied to the data set.*

Figure 4.13 shows various deformations of the beer mug model. Although our approach is not able to preserve refractive effects, the results appear realistic for reasonable deformations.

Our surfel renderer distorts surface splats in object space according to the local warp. This allows for rendering large deformations without visible holes (Figure 4.14).

Section 4.4.3 discussed the viewing ray divergence coming with the deformation of a surface reflectance field. Figure 4.15 shows an example of an SRF warp together with its corresponding camera and light source blending fields. Both blending fields were visualized by assigning random colors to cameras and light sources, respectively. Subsequently, the colors were blended together according to the ULR interpolation coefficients. The blending fields clearly show the viewing ray divergence in distorted regions.

Figure 4.16 shows a number of frames of a surface reflectance field animation including large deformations and varying lighting conditions.



Figure 4.13: *Deformation of refractive objects. Different warps were applied to the mug data set. All glasses are rendered from the same viewpoint and lit by five point light sources (red, green, blue, 2×white). The effect of the deformations on specular highlights can easily be observed.*

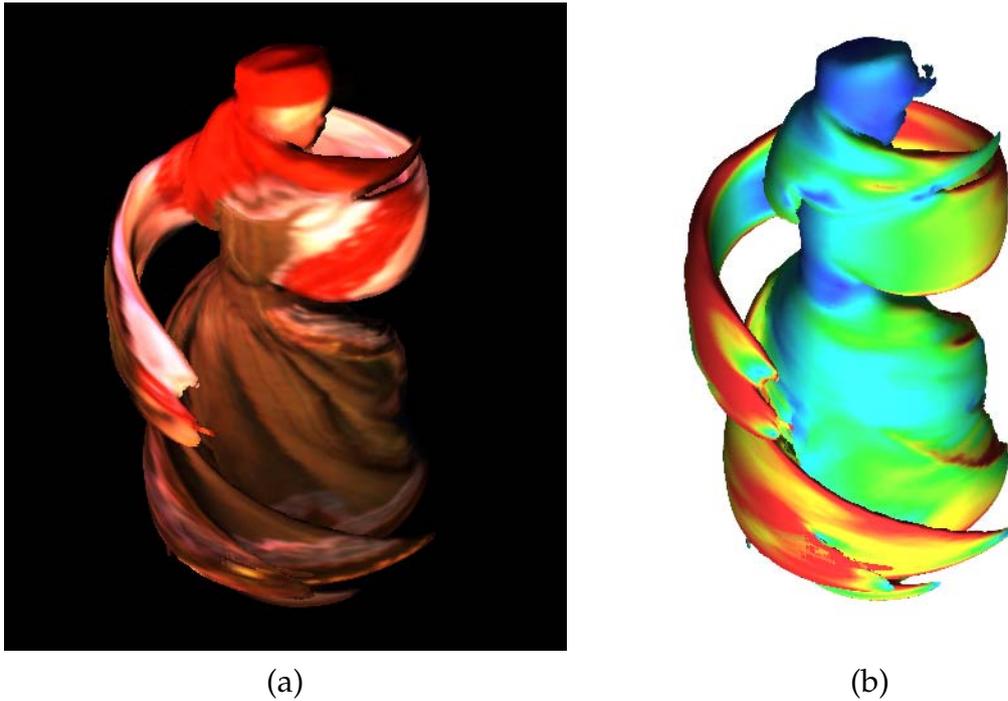


Figure 4.14: *Even under large deformations, the surfel renderer displays a closed surface. (a) The depicted warp anisotropically stretches parts of the doll data (b) The stretch anisotropy is color encoded showing the ratio of the principal axes of the elliptical surface splats. Splats in blue regions remain circular, green areas show a stretch ratio of 1:2, while red denotes a local stretch of more than 1:4.*

4.6 Reflectance Fields of Human Faces

In a preliminary acquisition setup consisting of 11 single-chip color cameras with 768×1024 resolution, we acquired a series of human faces. In contrast to Matusik's system the cameras stayed on tripods at fixed positions to allow for rapid acquisition without having to re-position the cameras. Consequently, the setup provides less viewing directions than Matusik's rig. The goal was to learn more about requirements with respect to:

- ▶ Camera coverage and density required for smooth view interpolation of faces.
- ▶ Precision of geometric and photometric multi-camera calibration.
- ▶ Quality of the impostor geometry.

In the experiment, illumination was kept fixed. We used the visual hull geometry reconstruction by [MPN⁺02] to obtain datasets similar to the previously presented. The required object segmentation has been performed by hand for each of



Figure 4.15: Viewing ray divergence. Top: The doll's skirt was distorted. Left: The corresponding camera blending field displays the viewing ray divergence in distorted regions. Right: The light source blending field.

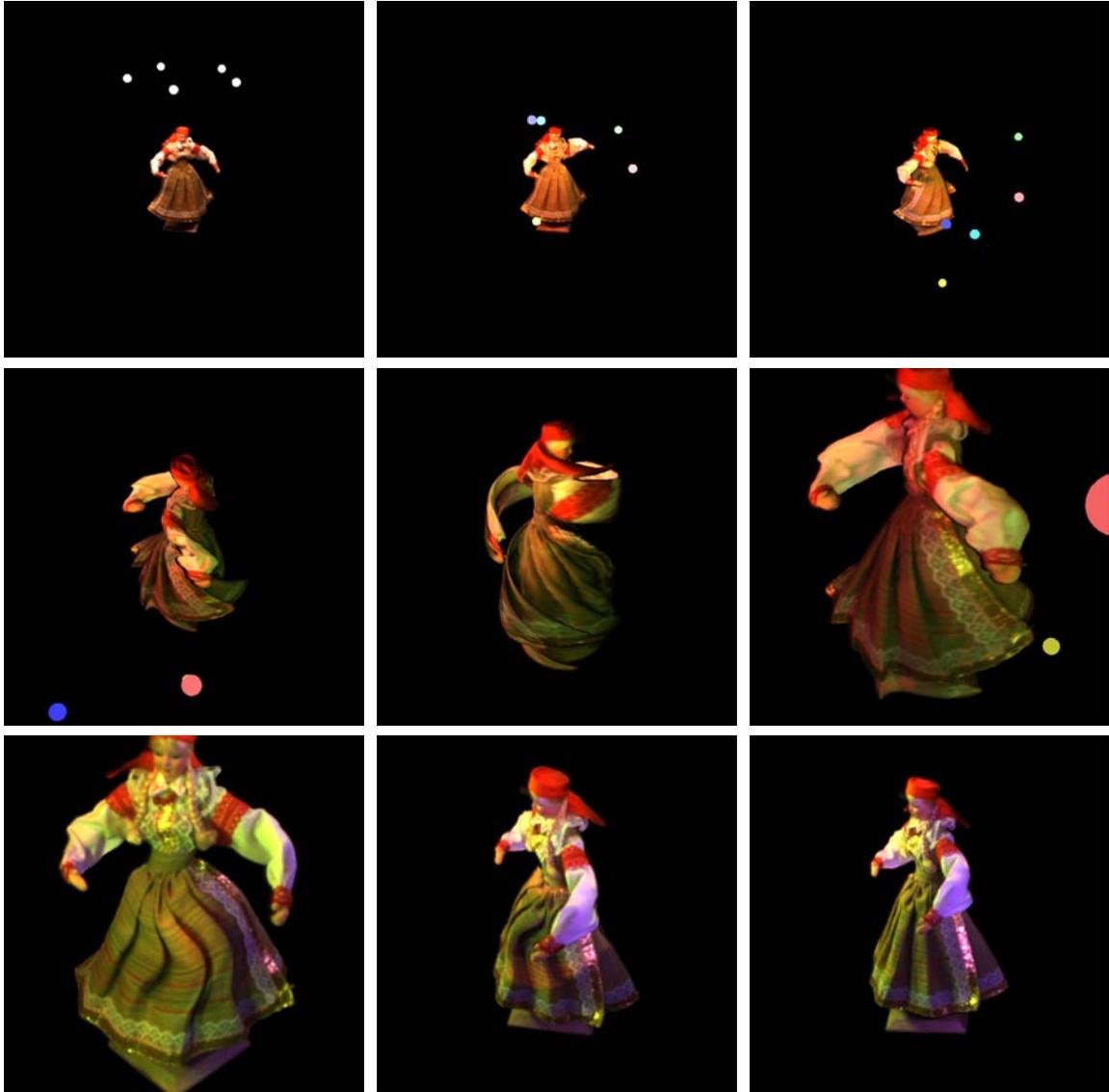


Figure 4.16: Selected frames of a surface reflectance field animation. The model is continuously deformed, while viewpoint and lighting conditions change. The model is lit by colored light sources, denoted by spheres of the light color.

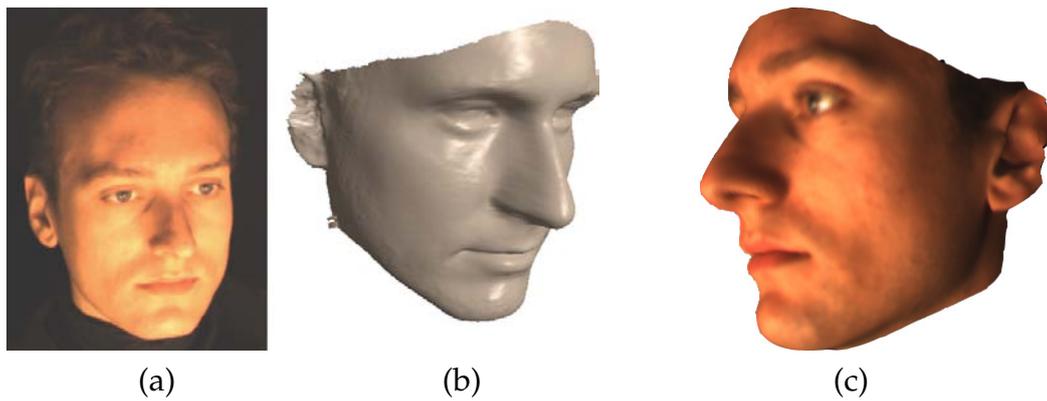


Figure 4.17: Face datasets obtained from silhouette fitting. (a,b) One of the input views and the fitted geometry. (c) A rendering using the deformable surface reflectance field renderer.

the 11 input images, as first attempts of an automatic segmentation using Bayesian Matting [CCSS01] have been unsatisfactory. The calibration already followed approximately the procedure presented in Appendix B. The resulting impostor shows significant deviations from the real geometry around the nose, as the visual hull is not able to fully capture the concavities left and right of the nose. This is the case although great care has been taken to place cameras that capture the silhouette of the nostrils.

In order to obtain a tighter geometry, we used the system presented in [LMPM03] to fit a parametric face model to the observed silhouettes. Figure 4.17 shows a resulting dataset and a rendering from a close-by local view, using the previously described look-up scheme. Animated renderings of the face datasets, however, still show more ghosting artifacts than the doll and mug datasets, which is mainly due to the sparser set of viewing directions.

The experiments indicate that our calibration procedure is suited, and that visual hull geometry would not be sufficient to faithfully render faces if only a limited set of viewing directions exists. Critical observation of the camera images shows that only little more cameras would be required to cover all areas of the facial surface at near-orthogonal angles. The camera resolution of 768×1024 , however, can be considered to represent the lower bound of what is required to capture facial skin detail.

4.7 Discussion

We presented a simple technique to preserve the appearance of a surface reflectance field under arbitrary deformations. However, the look-up scheme con-

tains some approximations affecting the proper reproduction of lighting effects. There are three classes of errors that may occur.

First, our BRDF-based derivation of the look-up function is ignoring non-local effects like inter-object reflection, self-shadowing, refraction, or sub-surface scattering. This may lead to false shadowing and erroneous refractions. However, modeling these effects requires incorporating the exact object geometry. As an SRF contains only approximate geometry information, these global lighting effects can not completely be preserved during the deformation.

The BRDF preserving approach itself contains some approximations. As \mathcal{L}_{p^*} is derived from the local warp around the observed impostor point instead of a point on the real object surface, the warped SRF may show some BRDFs that seem to be rotated relative to the surface frame. The effect grows as the local warp's rotational part around p^* is different from the rotation of the observed, real surface point q^* . Note that the error introduced by choosing the impostor normal to determine \mathcal{L}_{p^*} is comparatively low. In fact, the normal is only used to derive the rotational approximation of the local warp. Evaluating the SRF does not incorporate a normal anymore. In particular, this normal has no impact on the reflectance properties of the reproduced BRDFs.

A third error class affects object/impostor parallax in regions where the impostor geometry is distant to the object surface: For non-uniform stretches, the object texture may appear to be shifted. This happens if the corresponding viewing rays in object and acquisition space do not intersect the object at corresponding points q and q^* . Figure 4.5 shows such an example.

Despite these limitations, the proposed scheme shows significant advantages. Its ability to preserve reflectance properties without explicit knowledge of the real surface normals is crucial for the deformation of surface reflectance fields. Moreover, it makes it suited for many other image-based applications where the exact object normals are not known.

An important property of this look-up scheme is that in the limit the BRDF preservation is exact, as the SRF geometry converges to the real object geometry. The proposed method directly benefits from improvements of the geometric representation, for example, for geometry acquired with laser range or structured light scanning. As such, it could well be applied to our face datasets, provided a 3D reconstruction of sufficient precision.

When increasing the geometric precision, however, an attractive alternative to face modeling using deformable surface reflectance fields would be explicit modeling of facial geometry and its spatially varying reflectance properties. A respective representation would have to be able to geometrically model the skin surface, down to high-frequency normal variations due to pores and follicles. Provided that material properties are explicitly described at each surface point,

higher-order effects such as inter-reflection and self-shadowing can be simulated by a physically-based rendering system.

In our project, we decided to investigate the possibilities that an explicit modeling paradigm provides. That is, we treat facial surface reflectance fields as an intermediate representation that is used as input for a model fit. For more detail on this procedure refer to Chapter 8.

Facial Reflectance Field Acquisition

Independent from the choice of skin representation, it was clear at an early point in time that a full reflectance field of the human face would be a central component of our analysis. The general goal is to acquire enough information per surface point that allows equally for surface reflectance field rendering, as well as for a local BRDF analysis. We describe the development of a respective acquisition device and how we address the inherent issue of motion during acquisition.

5.1 Acquisition Device

Design and construction of the facial reflectance field acquisition device have been a central part of this thesis. This section gives a summary of the underlying design decisions and of the final acquisition system.

5.1.1 Design Considerations

Previous designs to acquire the reflectance field of static objects [MPN⁺02] move cameras and light sources to obtain a high number of viewpoints and illumination directions. This leads to acquisition times in the order of hours, which is not tolerable when scanning human faces. A different approach acquires human faces by keeping the camera fixed and only moving the light source [DHT⁺00], which

still leads to an acquisition time of about one minute. Acquisition time is not only an issue with respect to comfort of the subject. During longer acquisition times, subjects also tend to move, invalidating the reflectance samples of each surface point. Motion during facial reflectance acquisition has previously been addressed using fiducial markers attached to the forehead [MWLT00]. However, our application aims at the acquisition of an unobstructed face where markers would be unwanted occluders. As a consequence, reduction of acquisition time stayed one of the primary goals of our design.

Consequently, we decided for an entirely static design, using a set of fixed cameras and light sources. The lights are subsequently triggered, while the cameras simultaneously capture the different lighting conditions. Multiple-exposure images should allow for high dynamic range (HDR) acquisition. The static layout, however, requires a larger number of physical cameras and light sources, calling for a cost-efficient choice of these components. Furthermore, a suited spatial layout of cameras and lights had to be found, and a 3D reconstruction facility to obtain a three-dimensional representation of the facial surface had to be created.

Light Sources

Depending on the further use of the reflectance field data, there are diametrical requirements with respect to each light source's extent. For BRDF acquisition, infinitesimally small point lights would be the optimal choice. As soon as the light source spans a noticeable solid angle, the BRDF fit has to take the convolution of the BRDF with the light source profile into account. [LKG⁺01] and [MPBM03], for instance, use a Xenon arc lamp as a reasonable approximation of a point light source. For reflectance field acquisition, however, where the data is directly used for rendering, aliasing issues have to be considered. Ideal light sources that effectively suppress aliasing would correspond to the reconstruction kernel that is used during rendering to blend different lighting conditions of the SRF. That is, in the ideal case they were overlapping area light sources with an intensity profile according to the reconstruction kernel. A lighting facility that comes close to this case is presented in [MTB⁺05] and uses a computer-steered moving-head spotlight inside a foam-core enclosure to generate area light sources by indirect illumination. Given that facial BRDFs are rather glossy than specular, making reflectance field aliasing less of an issue, we finally decided not to use SRF-optimized area light sources. Instead, we performed experiments that should show how densely we have to place small light sources to minimize visible ghosting. Figure 5.1 shows an image series where a face is illuminated by differently colored light sources. Below a light source spacing of 10° , shadow boundaries and multiple highlights start to vanish.

In order to meet short acquisition times, the light sources have to be triggerable within fractions of a second. It is important that each light source reaches max-

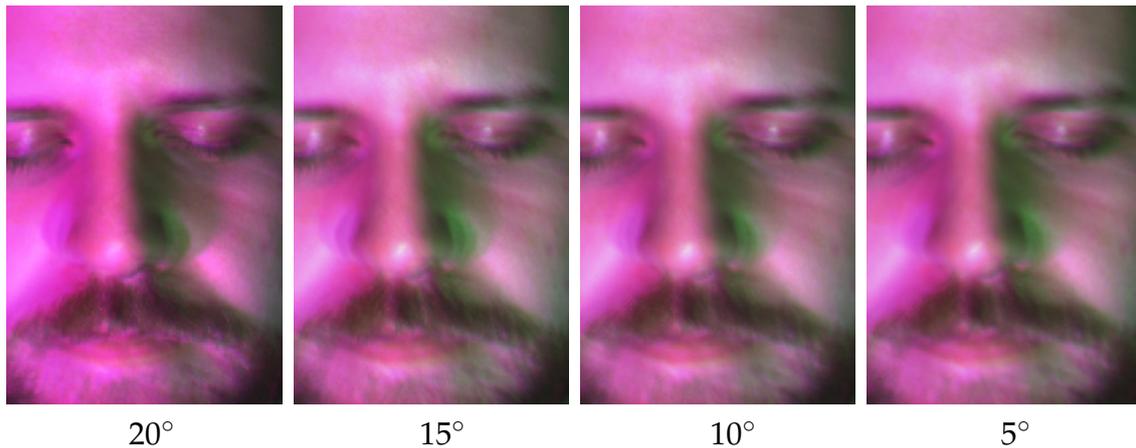


Figure 5.1: Face images under differently spaced light sources. Below a light source spacing of 10° , shadow boundaries and multiple highlights start to vanish.

imum intensity very quickly, and that a potential after-glow is minimal, as each light source has to be completely dark before the next illumination condition can be acquired. Xenon arc lamps used by previous systems are not suited for frequent on/off cycles. An alternative would be flash-lights, which are reasonably small and bright enough to allow for short exposure times. However, as we are targeting multi-exposure acquisition, each light has to be triggered multiple times, making flash-lights with a noticeable recharging period an inferior choice.

In our system, we decided for LED light sources that are quickly triggerable and come at a reasonable price, allowing a larger number of light sources. One downside remains, however, as LEDs do not provide an even spectrum. Wenger et al. discuss possible workarounds using multiple monochromatic LEDs [WHD03].

We initially looked into commercially available LED lighting solutions. One product appeared most appealing in terms of size and brightness: Color Kinetics Inc. offer LED light sources consisting of red, green, and blue LEDs, allowing a 24-bit RGB control of the light color, see Figure 5.2. However, the color control, not of direct use for our purposes, turned out to be problematic. The mixed color is produced by three disjunct clusters of LEDs of the primary colors. The spatial separation of these clusters leads to color fringes at shadow boundaries. In turn, we had to use white LEDs instead. At the time of construction, there were no readily-built triggerable white LED light sources of sufficient brightness on the market. Hence, we decided to let them custom-build.



Figure 5.2: Color Kinetics iColor™ light source.

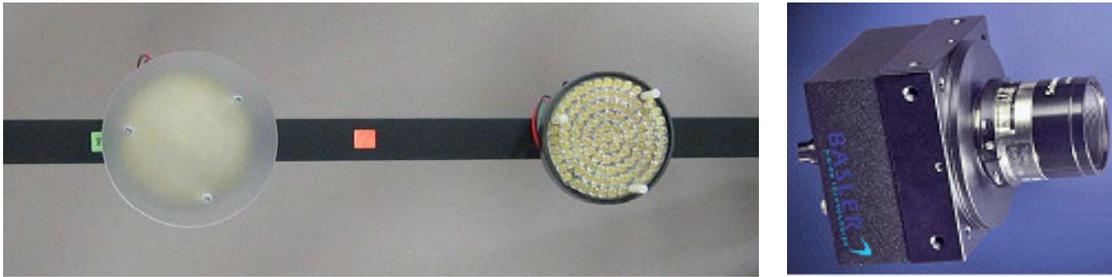


Figure 5.3: *Dome light source and camera type. Left: The custom build LED light sources; one light source has the diffuser removed. Right: We use industrial computer vision cameras.*

The final light sources are disk-shaped assemblies of 150 white LEDs with a reasonably wide spectrum. Although brighter white LEDs just became available (Lumileds Luxeon[®] LEDs), their spectrum was still too irregular. The disk-shape of our light sources has been chosen to allow an easier deconvolution of the BRDF.

During first experiments with a light source prototype, it turned out that the light source shows a considerable intensity fall-off away from its directional axis. The beam structure is rather complex, making it difficult to properly calibrate for this effect. Unmodeled light intensity variations, however, make the fit of a reflectance model impossible. Using opal glass diffusers to obtain a near-Lambertian beam characteristic caused too much intensity loss due to the beam widening. We finally found a suited acrylic diffuser material that solely smoothed the beam structure so that its cross-section could be modeled by a bi-variate second-degree polynomial (see Appendix B.3). Figure 5.3 (a) shows the light source assembled with the acrylic diffuser.

Cameras

From our first experiments presented with facial reflectance fields (Section 4.6) we know that the cameras have to have a resolution beyond 1024×768 to capture sufficient surface detail. They have to be fast enough to acquire and transmit images at a maximum rate possible with the given light source intensity. Possible exposure times also depend from the lenses and aperture settings that are used. Aperture itself has to be chosen to obtain an adequate depth-of-field with a given lens, keeping all facial areas in focus of the camera.

Finding an optimal combination of these parameters, compliant with the financial constraints of the project, was a complex task. Between 2002 and 2003, we experimented with a number of different cameras and lenses. In addition, we developed a camera simulation system that allows to capture reflectance fields of a

face model using arbitrary camera parameters, including the simulation of camera artifacts, such as vignetting, blooming, smear, and Bayer interpolation [Nef02].

The final choice has been the A101fc camera model by Basler Vision Technologies AG, shown in Figure 5.3 (b), with 16mm C-mount lenses. As a 1300×1030 single-chip color camera, it provides sufficient resolution for our purposes. The firewire interface allows to capture this resolution at maximal 12 frames per second. Although the camera is also available with a proprietary interface for a faster transmission rate, 12 fps appear to be the maximum speed that keeps the camera noise within acceptable bounds. Other camera models with a superior signal-noise ratio, however, would have entered a different price region that would not have allowed us to equip the scanner with a high number of them.

The A101fc's sensor characteristic is highly linear. Initial experiments with the camera showed that the A/D converter is precise enough to capture a dynamic range sufficient for facial reflectance using only two different exposure settings. Consequently, the scanner has been designed to acquire each lighting condition under two different exposures.

Layout

An important question is how light sources and cameras should be arranged around the subject. It makes sense to attach them to a spherical dome structure around the head to cover different spatial directions, at the same time keeping a constant distance to the face. Additionally, their spatial distribution has to consider the coverage of the facial surface, as well as the coverage of the angular domain above each surface point.

In accordance with the experiment shown in Figure 5.1, the light sources should maximally be spaced at 10° angles. This naturally guarantees surface coverage by the light sources. For a BRDF fit, it generally makes sense to have a higher light source sampling in directions of expected specular highlights [Rus98, MPBM03]. However, it turned out to be impossible to identify such preferred sampling directions for a full face. Due to the general normal orientations across the facial surface, the best sampling that suits all surface points equally well is an even light source distribution. Consequently, we place the light sources at a regular 10° spacing within the dome structure, leaving out locations where the face is shadowed by the subject. This leads to a total of 150 light sources.

Similar considerations lead to a desired even spacing of the cameras. However, we slightly increase the number of near-frontal views to capture more textural detail from views that are important for the perception of faces. Using a software simulation of the camera placement, we ensured that each surface point is observed by at least two cameras. The final camera layout uses 16 cameras, and it

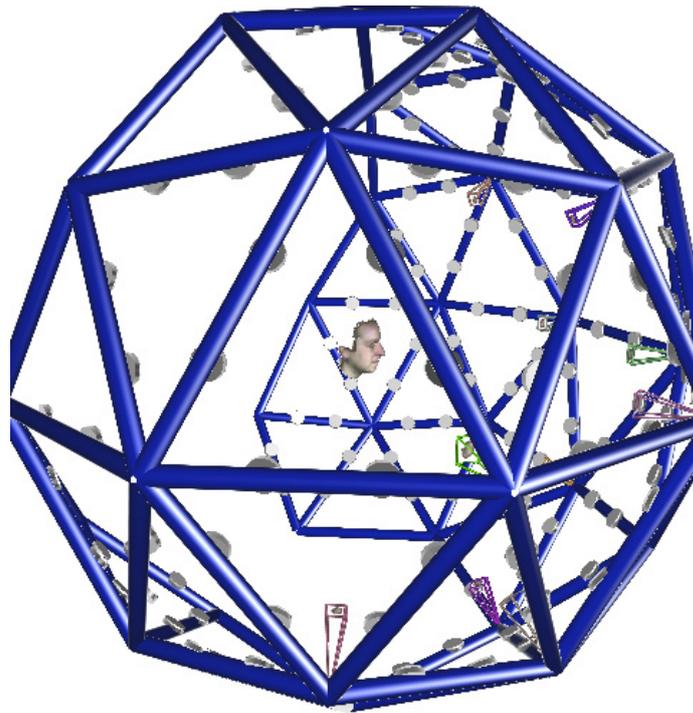


Figure 5.4: Acquisition dome simulator. Light sources are denoted as white disks, cameras appear as view frusta.

is a superset of the camera positions used for the experiment presented in Section 4.6. For some of the cameras, alternative mount positions exist. Replacing those cameras at these positions provides better viewpoints for 3D face recognition from silhouettes [LMPM04].

The easiest way to construct a stable spherical construction is to use a geodesic dome. We meet the requirements described in this section by choosing a *pentakis dodecahedron* (the dual to a truncated icosahedron with unit edge lengths) for the construction with two light sources per edge and camera mount points at the nodes. Such a dome topology is also known as *2V dome*. See Figure 5.4 for the final design within the software simulator used for the preparation of the project.

3D Reconstruction

Last but not least, the system has to include a facility for 3D reconstruction to provide a faithful geometrical representation of the facial surface. As argued in Section 4.6, visual hull geometry is not sufficient for human faces. In the course of this thesis, various 3D reconstruction algorithms have been evaluated using a setup consisting of a projector, multiple light sources and multiple cameras, as depicted in Figure 5.5. In particular, we investigated in active and passive stereo



Figure 5.5: *3D reconstruction testbed. We evaluated photometric and active stereo techniques using three light sources, five cameras, and a structured light projector.*

techniques and their combination with photometric stereo approaches [Hub02, Sch03]. Due to potential subject motion, a respective reconstruction algorithm has to be able to acquire a reliable geometry with an accuracy below 1 mm within minimal time.

As a result, we came to the conclusion that the best suited approach would use a static random speckle pattern projection in conjunction with passive stereo cameras, allowing for single-shot acquisition of the geometry. The speckle pattern provides a well-correlating texture that allows for an unambiguous stereo reconstruction of the surface [Sch03]. Luckily, such a system already existed as a commercial product: purchasing a face scanning system from 3QTech Inc. freed us from the burden of re-implementing the 3D reconstruction within our acquisition dome. The 3QTech scanner comes in the form of two acquisition pods that could easily be integrated into the dome construction. The scanner provides a geometric accuracy of ± 1 mm and additionally acquires an RGB texture of the face. Although not incorporated in our reflectance modeling pipeline, the texture is used in our system to identify regions of interest (Section 7.2) and to register the calibration frame of our 16 cameras with the 3QTech scan (Appendix B.4).

5.1.2 Final Hardware

The actual construction and assembly of the acquisition device has been assigned to an external contractor that is closely located to MERL in Boston. The contractor received a complete specification of the construction.

The scanner has been built in the course of 2003. The construction took place in three phases, allowing us to evaluate the system components before the entire system had been built. Most of the evaluation has been done in the context of my stay at MERL in Boston, parts were performed remotely from ETH.

Phase I. This first phase consisted of a single camera and two light sources. Using this stage of development, we evaluated different lenses, camera settings, and the diffusers for beam smoothening.

Phase II. In a second expansion stage, 2 cameras and 12 light sources have been attached to two mobile bars, allowing us to generate small-scale reflectance fields and to evaluate the light source control. Figure 5.6 shows the respective camera and light source setup, as recovered from the geometric calibration data.

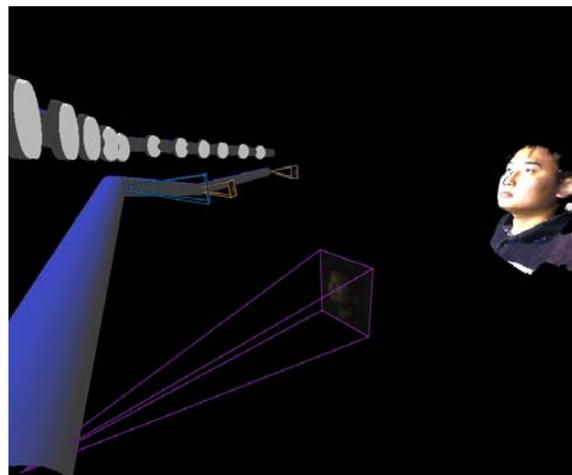


Figure 5.6: *Second expansion stage of the scanner. Light source and camera positions have been recovered from calibration data.*

Phase III. The final stage, depicted in Figure 5.7, comprises the full dome construction with 150 light sources, 16 cameras, and a small PC cluster to retrieve the stream of images from the cameras. Light sources and cameras are attached to locking ball swivel mounts to allow a careful alignment with the dome center. We attached switchable laser pointers to the construction, providing two laser beams

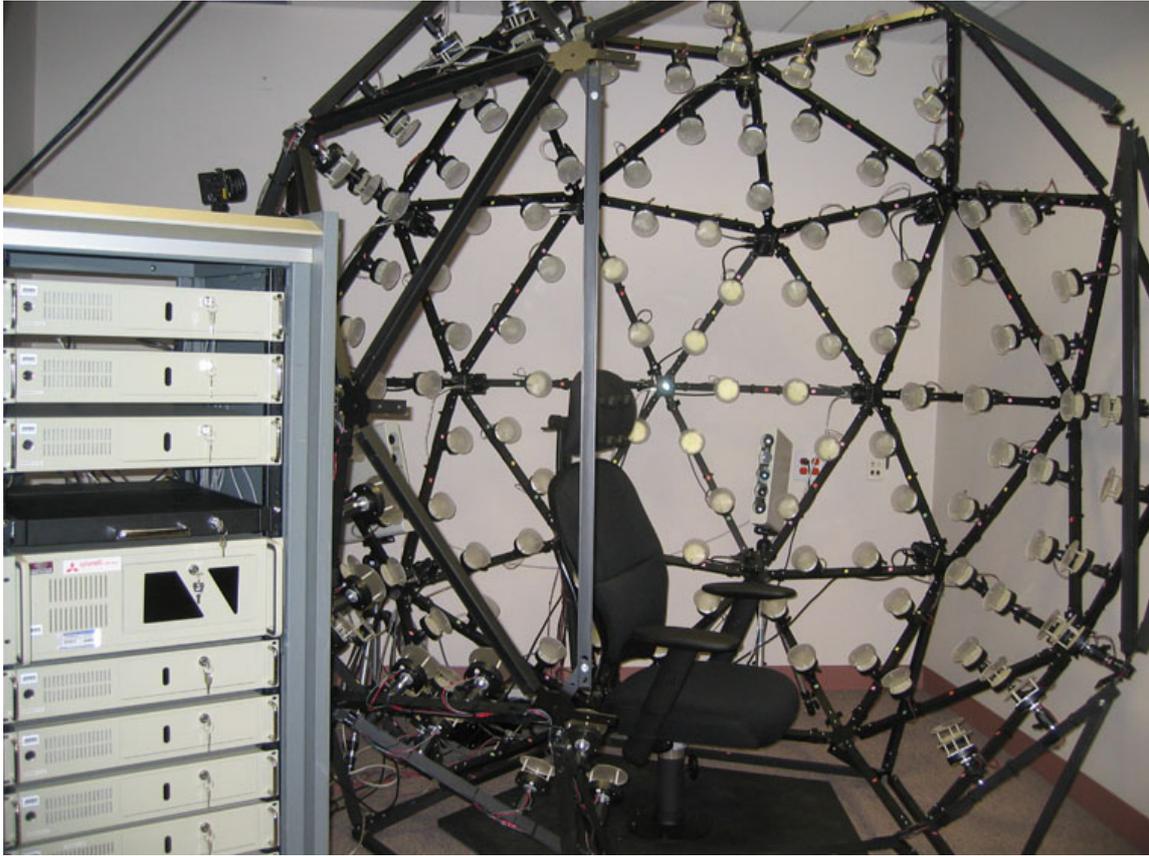


Figure 5.7: *The final dome construction. To the left, the rack with the client PC and the controlling server is visible.*

that intersect at the center of the dome, assisting the operator with the camera and light source alignment. In the center of the dome, a swivel chair accommodates the subject to be scanned, see Figure 5.8. A headrest minimizes head motion during the acquisition. The dome construction has been set up within a dedicated laboratory at Mitsubishi Electric Research Laboratories.

5.1.3 Acquisition Dome Software

The whole system is synchronously triggered. A server PC drives the light source controller, which is based on an FPGA that triggers each of the 150 light sources. Unfortunately, the contractor did not manage to provide all requested lighting sequences, which restricted some of the scanner's functionality (see Section 5.2.1). All camera triggers are hooked up with the light source control, allowing for synchronous image acquisition. The camera settings are controlled by a cluster of 8 client PCs connected to the server. Each client controls two cameras. During acquisition, it retrieves all image data from its cameras and stores them in local

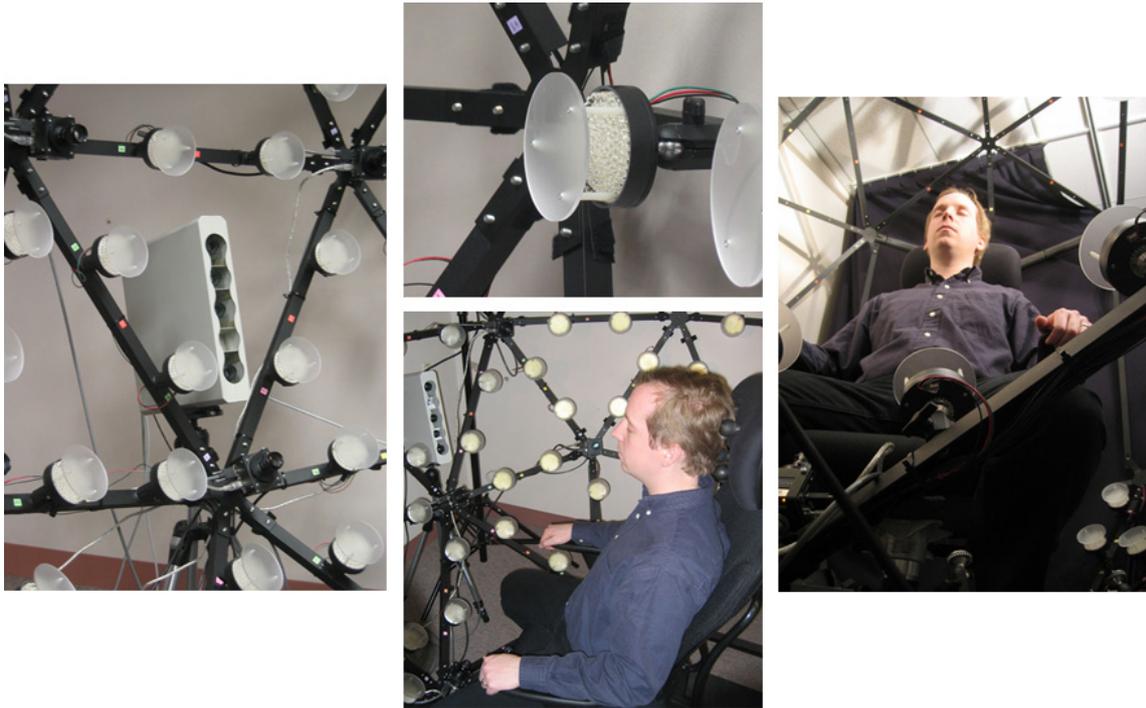


Figure 5.8: *Close-ups of the dome construction. Left: One of the two 3D reconstruction pods. Top: Light sources mounted in the dome. Bottom and Right: A subject is seated in the center of the dome.*

RAM. Upon scan completion, all clients store their data to a network drive. The respective capture software has been developed by the contractor.

Our own processing pipeline picks up the raw camera images as stored by the client PCs, converts them to color images (demosaicing), corrects for lens distortions (cf. Appendix B.2.1), and places the whole reflectance field data of approximately 4.5 GB in a dedicated directory structure for each scanned subject. Figure 5.9 shows a subset of such a reflectance field.

5.2 Motion Compensation

Unintentional motion by the subject tends to invalidate the reflectance data. In order to realize if a subject has moved, a rough motion detection is performed immediately after each scan. We use the fact, that due to the double-exposure acquisition, each lighting condition is acquired twice, at two points 13 seconds apart in time. A simple Fourier-based image correlation between two such images yields a translation vector in the image plane, which is a rough estimate of the subject's motion within that period of time as seen from the camera. If the subject did not move, both images should be identical up to scale, yielding a motion vector



Figure 5.9: Raw reflectance images of a subject acquired with all 16 cameras and 14 (of 150) lighting conditions. Each row shows images with different camera viewpoints and the same lighting condition. Each column displays images with different lighting conditions and the same viewpoint. The displayed images were all captured using the higher exposure setting. The images are not yet color corrected, revealing differences between camera characteristics and light source colors.

near zero. For a rapid motion detection after the reflectance field acquisition, we perform this test for four selected camera / light pairs.

Although this kind of motion detection does not provide quantitatively accurate information about the subject's motion, it is still sufficient to detect whether the motion exceeds a certain limit. As it is computed even before the reflectance field data is stored, it allows to promptly discard scans with too much motion and to repeat the procedure while the subject is still seated in the scanner.

In some cases, however, a certain head motion can not completely be avoided. While some subjects have no difficulties to sit perfectly still, some others could not be scanned without a motion of up to 5 pixels in image space. In order to improve these critical datasets, we developed a motion compensation pipeline that can optionally be used before further processing of the scan. In three stages, the pipeline estimates subject motion with respect to the camera planes, reconstructs a motion trajectory, and finally re-samples the reflectance images to compensate for the recovered motion. As it only corrects the input data, the motion compensation can be used independently from any further use of the data.

5.2.1 Motion Estimation

The motion estimation is based on optical flow, which estimates motion vectors for image regions in each camera's image plane. Naively applying optical flow to subsequent images in the acquisition sequence, however, is bound to fail, as the flow would be computed between images of different lighting. The resulting flow field would rather capture the flow in shading and shadowing, but not real motion of the underlying surface.

One way to exclude virtual motion due to illumination changes would be to insert constant illumination patterns into the light sequence. Such tracking frames have been successfully used for example by Wenger et al. [WGT⁺05], however, as our acquisition runs at 12 frames per second, tracking frames would have prolonged the acquisition time significantly. Hence the original plan has been to reorder the light sequence in a pseudo-random order. This allows to approximate a constantly even illumination by combining images of light conditions adjacent in time. The pseudo-random sequence is statically generated using an adapted algorithm to create ordered dither matrices. Figure 5.10 shows the light sequence within the dome layout. Unfortunately, it turned out that our contractor was not able to include this pattern into the dome driver. This in effect rendered a direct motion analysis between adjacent points in time impossible. Consequently, a motion estimation for the acquired data has to comply with the existing sequential light sequence.

As the only feasible way to rule out effects due to shading remained the use of flow fields between images of identical illumination only, similar to the rough

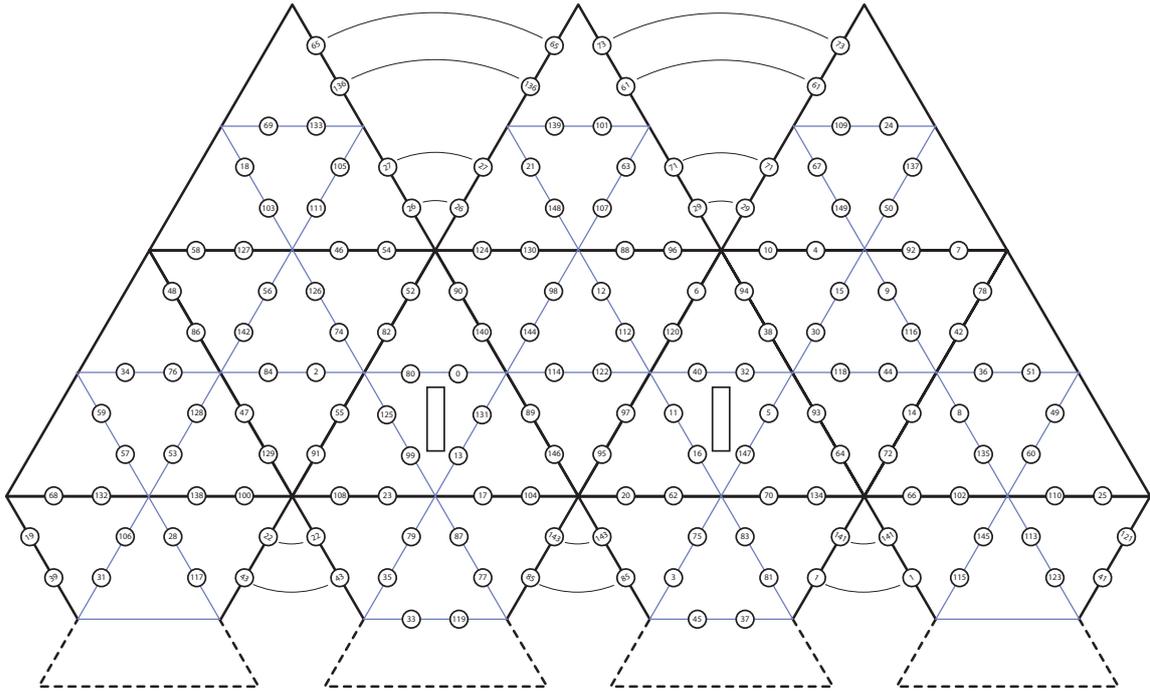


Figure 5.10: Pseudo random order of light sources for motion compensation. The original design targets a pseudo random order of the light sources, allowing to produce a near-diffuse illumination by overlaying lighting conditions adjacent in time. The numbers denote the frame index when a light source is triggered.

motion detection described above. For each camera, this results in pair-wise relations between frames at points in time t_i and t_{i+n} , with n the number of light sources. Note that the relations are disjunct in the sense that for each camera, no frame t_i is connected to more than one other frame $t_{i\pm n}$ by a flow field relation (see Figure 5.11). As only sufficiently lit images can be used for optical flow, just a fraction of frames have such a relation. Many other reflectance images show mainly shadowed conditions of the face and can not be used for motion estimation.

Using the pyramidal Lukas-Kanade algorithm [LK81, OCV], we determine optical flow fields from exposure-corrected pairs of images. We use a combination of heuristics to constrain the resulting flow field to regions where the optical flow is likely to correspond to actual subject motion: background regions are filtered out by masking the flow field by the projection of the acquired geometry. We also discard shadowed regions and highlights using intensity thresholding. Shadows produce unreliable data due to noise, and highlights tend to invalidate the motion field, as they move independently from the real surface motion. Figure 5.12 shows a flow field before and after masking out unreliable regions.

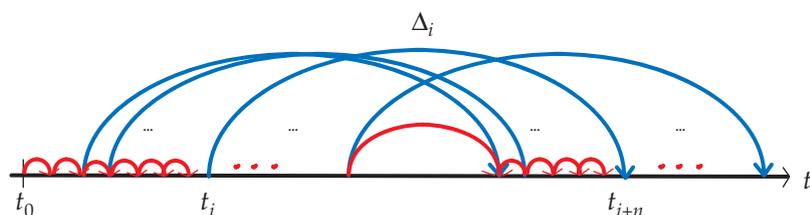
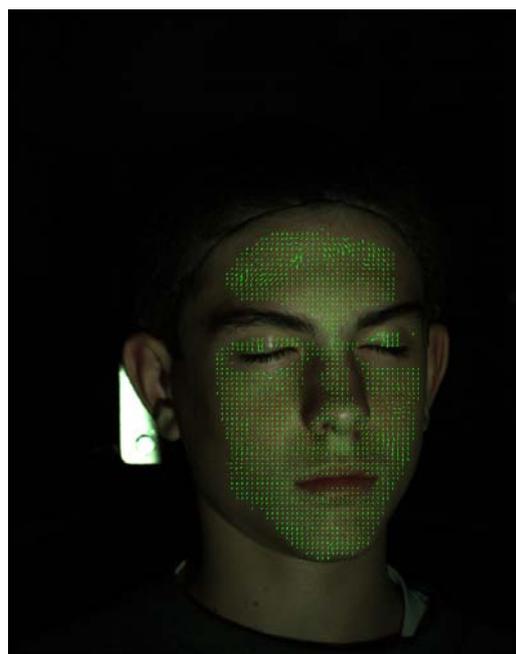


Figure 5.11: Flow field relations. Blue arrows denote the sparse flow relations between points t_i and t_{i+n} in time. Red arrows denote the relations that have to be found for a motion compensation.



(a)



(b)

Figure 5.12: Flow masking. Background, shadows, and highlights are filtered from the flow field in order to obtain a reliable flow estimation.

5.2.2 Trajectory Recovery

From this sparse set of inter-frame relations, the trajectory recovery has to reconstruct a consistent motion hypothesis. Two general approaches exist. An image-space motion recovery would reconstruct for each camera a two-dimensional motion trajectory that is compliant with the observed flow fields. Alternatively, one could fit a single three-dimensional head trajectory to the flow fields of all cameras. We decided for the former variant, as its realization is less complex, not requiring the incorporation of an explicit face geometry.

Prior to a trajectory recovery per camera, we derive a two-dimensional face motion hypothesis for each flow field separately. For stability reasons, we use a very simple motion model, assuming that the two-dimensional motion is rigid. We use the Procrustes method to obtain a rigid transformation Δ_i^c between frames t_i and t_{i+n} that is consistent with the flow field of camera c . For each camera, we then aim to find a trajectory \mathcal{P}^c that is consistent with those transformations.

Let a trajectory \mathcal{P} be defined by poses p_i for acquisition frames $i = 0 \dots 2n - 1$. Each pose is defined by a rigid transform relative to the first frame with $p_0 = id$. The estimated pose transformations Δ_i constrain the trajectory to

$$p_{i+n} = \Delta_i(p_i), \quad (5.1)$$

for all frames i with a flow relation. Despite these constraints, the trajectory is still under-determined, calling for the introduction of a prior. A reasonable prior is to assume a constant head pose over time. This makes sense, as the subject actively tries to sit still. Let $\|p_i\| = \|p_i - p_0\|$ be a norm that quantifies the degree of motion of p_i with respect to the initial frame. The constant head pose constraint can be expressed as a minimization task

$$\min_{\mathcal{P}} \sum_{i=0}^{2n-2} \|p_{i+1} - p_i\|^2, \quad (5.2)$$

constrained by Equation (5.1). We recover each motion trajectory \mathcal{P}^c by minimizing (5.2) using a non-linear least-squares optimization.

In order to increase stability of the recovery, the optimization is only performed with the 20% most reliable flow relations. Further stability is obtained by fitting a fourth-degree polynomial to each pose parameter. Figure 5.13 shows the translational part of an exemplary head trajectory. The output by the non-linear least-squares optimization is shown in points, the fitted polynomial motion is drawn as a curve. The plots show that the subject moved by up to eight pixels during the acquisition.

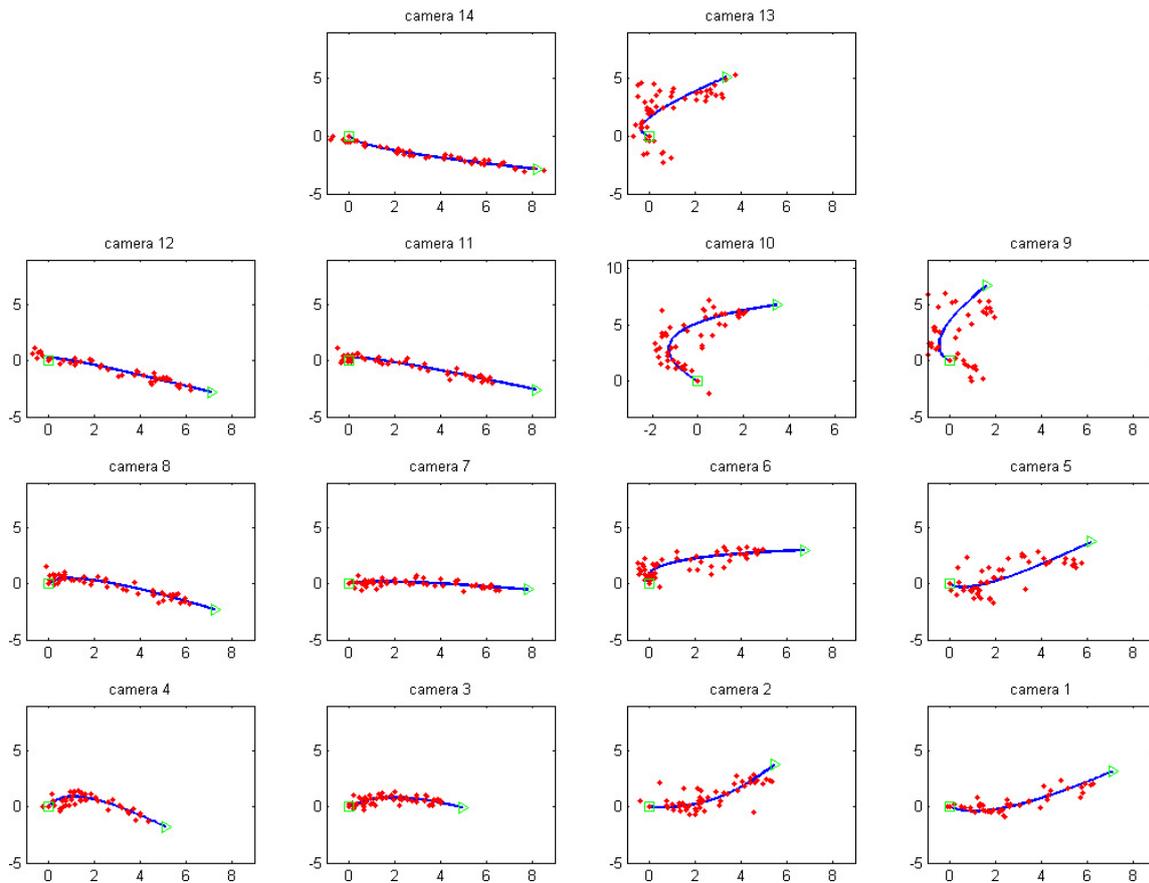


Figure 5.13: *Translational parts of an exemplary head trajectory. For each of 14 cameras, the output by the non-linear least-squares optimization is shown in points, the fitted polynomial motion is drawn as a curve. The motion trajectory always starts at $(0,0)$, and its end is denoted by \triangleright . Units are pixels in the respective camera plane.*

5.2.3 Image Warp

In a final step, all images in the reflectance field are unwarped to compensate for the recovered motion trajectory. Each image is translated and rotated according to the pose estimate for its camera and frame. We use bi-linear interpolation for the respective re-sampling. We refrained from using higher-order interpolation in order to avoid over-shooting of the resulting intensity values due to camera noise. The result is stored as a new reflectance field, serving as input for all further processing steps. Figure 5.14 shows two example reflectance image pairs before and after motion compensation.

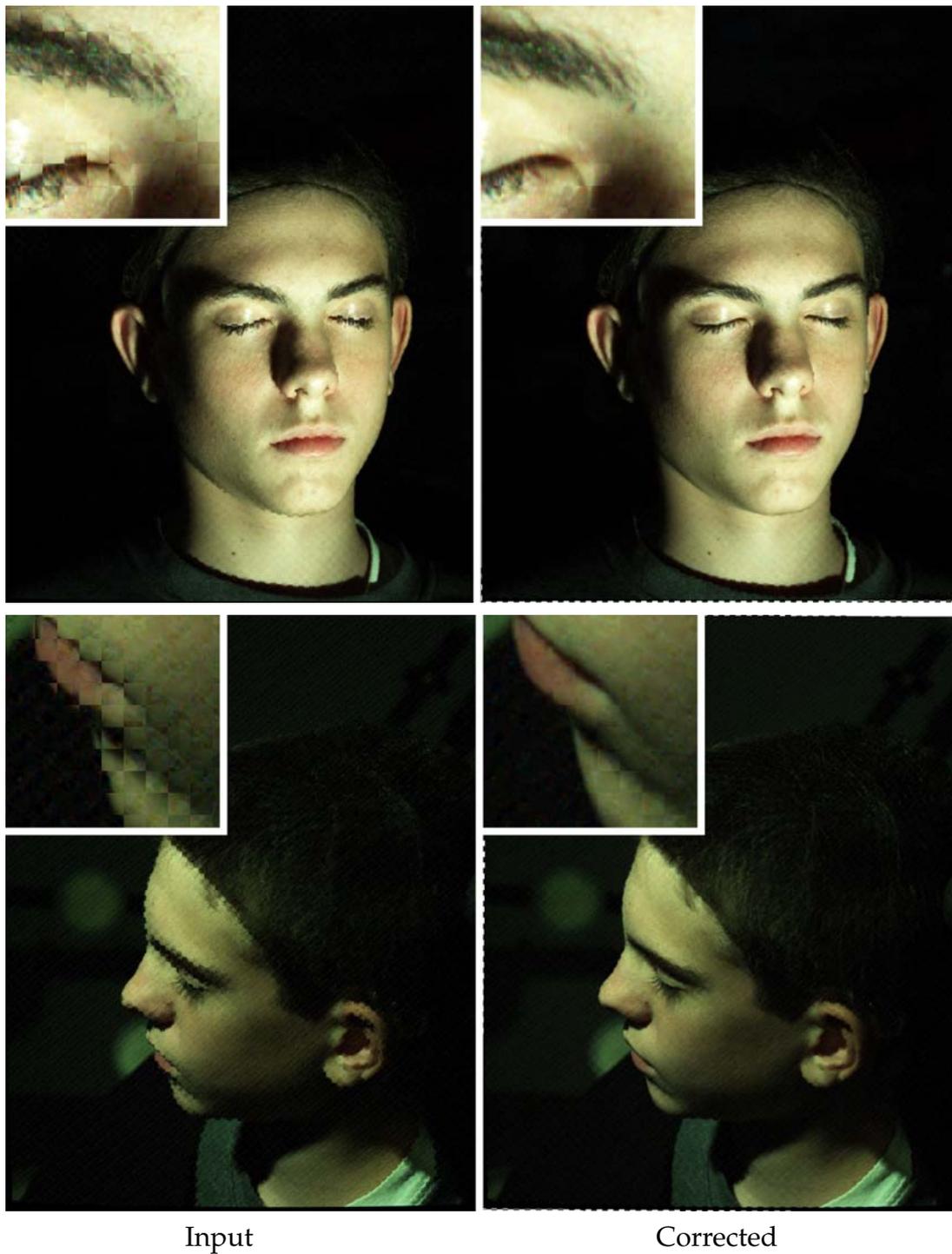


Figure 5.14: Motion compensation. Left: From the raw input images a chess board mosaic is computed by interleaving intensity normalized images of two different exposure times (150 frames apart in time). The apparent chess board structure in the mosaic shows a significant motion between the two frames. Right: Chess board mosaics of motion-compensated image pairs. Chess board artifacts in the face are mostly eliminated. Note that, however, the static background now appears to be jittered. Insets show close-ups.

5.3 Discussion

In the previous sections we discussed the design considerations that led to our custom device for facial surface reflectance field acquisition. Short acquisition time as a central design criterion to minimize motion artifacts required a static design with multiple light sources and cameras to obtain a dense reflectance field. The final scanner consists of 150 light sources, 16 cameras, and a commercial 3D range scanner, allowing to acquire the full reflectance field of a human face within 25 seconds. The reflectance field consists of 2 400 high dynamic range (HDR) images. We further developed a novel motion analysis to compensate for unconscious head motion by the scanned subject.

This system is unique of its kind. It allows to collect an unprecedented amount of reflectance data of a human face, and it is well-suited for our projected acquisition of the large face database presented in Chapter 10. Nevertheless, certain technical constraints introduce error sources that potentially affect the quality of the reflectance field.

Despite the motion compensation, motion artifacts still remain a potential error source. Subtle changes of the facial expression over time, e.g., cannot be compensated by our proposed method. The only way to reliably suppress all kinds of motion artifacts would be to reduce the acquisition time even further. With the current setup, acquisition time is mainly constrained by the 400 Mbit/s bandwidth limitation of the firewire interface between cameras and client PCs. However, even with a higher transmission bandwidth, the existing light source brightness and camera sensitivity would still limit the achievable acquisition frame rates. This is due to the fact that the large number of components required a cost-efficient choice of cameras and light sources. Effectively increasing the acquisition speed while retaining the current directional sampling density of the reflectance field would have drastically increased the costs of the acquisition device.

A second potential error source is the geometrical precision of the 3D scanner. The commercial scanner in our setup is designed for an absolute precision of ± 1 mm, which is a typical resolution as required, e.g., for surgical simulation. However, when re-parameterizing the reflectance field at the face geometry, see Section 7.4, this precision still leads to unacceptable registration errors when projecting the camera images onto the face geometry. Section 7.6 presents a geometry refinement technique that improves the scanned geometry to obtain a higher consistency with the camera images. An alternative, however, would have been to directly use the dome cameras for multiple-baseline stereo, obtaining a high-quality geometry from a static speckle pattern projection. With a back-projection error below ± 0.15 pixels, our camera calibration fulfills the respective precision requirements.

From our experience, however, the current geometric and radiometric quality of the acquired data is sufficient for our application. Chapter 7 describes the fur-

ther refinement of the data, before it is used in our skin reflectance model fit, as described in Chapter 8.

Subsurface Scattering Acquisition

With the decision for an analytic skin reflectance model, as motivated in Section 4.7, a deeper analysis of the constituents of skin reflectance becomes necessary. One of the most prominent effects is subcutaneous scattering of light. In fact, 93 to 96% of the light reflected from skin is due to subsurface scattering [INN05]. In our attempt to capture this effect, we decided to build a custom device.

6.1 Problem Statement

Internal light transport within skin can be described by its bi-directional surface scattering reflectance distribution function (BSSRDF), see Section 2.2.2. As the BSSRDF is a 9 dimensional function, measuring a general BSSRDF requires drastically more samples than measuring a BRDF of 6 to 7 (spatially varying) or 4 to 5 (uniform) degrees of freedom, respectively. Hence, in order to efficiently measure subcutaneous scattering properties, simplifying assumptions have to be made, reducing the number of samples to be acquired.

A common assumption is that the medium is homogeneous and of a high density, which means that exiting light is the result of a very high number of scattering events. This allows for a simplified scattering model, as described in Section 2.2.2. For such materials, light transport can be defined by an intensity fall-off function across the surface: the radiant exitance at a surface point x_0 due to a constant

incident illumination at point \mathbf{x}_i is proportional to a characteristic fall-off kernel function k in dependence from $\mathbf{x}_o - \mathbf{x}_i$:

$$M(\mathbf{x}_o) = k(\mathbf{x}_o - \mathbf{x}_i)E(\mathbf{x}_i) = k(\mathbf{x}_o - \mathbf{x}_i) \int_{\Omega_+} L(\mathbf{x}_i, \boldsymbol{\omega}) \mathbf{n} d\boldsymbol{\omega} \quad (6.1)$$

If the material is furthermore isotropic, the fall-off $k(r)$ is only dependent from the distance $r = \|\mathbf{x}_o - \mathbf{x}_i\|$ to the point of incident light.

6.1.1 Requirements

Any suited measurement device has to determine this fall-off function. That is, the skin's impulse response to incident light is to be measured. The near-exponential fall-off of k requires measurements with a high dynamic range. For instance, the characteristic impulse response of skin covers up to 5 orders of magnitude within the first few millimeters from the point of incident light.

An additional requirement occurring particularly for *in vivo* measurements of human skin is that the device would have to be robust against motion during the measurement procedure.

6.1.2 Previous Designs

Homogeneity and high density cause the directional distribution of outgoing radiance to be constant across the surface. This allows to measure the integral of outgoing radiance within a fixed solid angle Ω only, relating radiant exitance M to the measurement by a constant scalar K :

$$M(\mathbf{x}_o) = K \int_{\Omega} L(\mathbf{x}_o, \boldsymbol{\omega}) \mathbf{n} d\boldsymbol{\omega} . \quad (6.2)$$

All existing measurement devices rely on (6.2), which allows to determine k as

$$k(\|\mathbf{x}_o - \mathbf{x}_i\|) = \frac{M(\mathbf{x}_o)}{E(\mathbf{x}_i)} = \frac{K}{E(\mathbf{x}_i)} \int_{\Omega} L(\mathbf{x}_o, \boldsymbol{\omega}) \mathbf{n} d\boldsymbol{\omega} . \quad (6.3)$$

However, not all of them calibrate for K , which is dependent from Ω :

$$K = \frac{\int_{\Omega_+} L(\mathbf{x}_o, \boldsymbol{\omega}) \mathbf{n} d\boldsymbol{\omega}}{\int_{\Omega} L(\mathbf{x}_o, \boldsymbol{\omega}) \mathbf{n} d\boldsymbol{\omega}} = \text{const} . \quad (6.4)$$

A short overview over previous acquisition systems for subsurface scattering follows.

In Vivo Measurements

Schmitt et al. [SZW90] perform *in vivo* measurements of k in skin, using two optical fibers. One source fiber couples light into the skin, and one receiving fiber is subsequently placed at different distances by a calibrated positioner. Outgoing radiance is measured using a photoamplifier. Using neutral-density filters of known transmittance, the photoamplifier is calibrated relative to the incident light intensity. K is analytically determined based on the numerical aperture of the optical fibers. They compare their skin measurements to simulations obtained with a dual-layer diffusion model of skin.

[FPW92] measure skin using a probe containing ten $400\text{-}\mu\text{m}$ -diameter optical fibers, placed in contact with the skin. One of the optical fiber acts as a light source while the other fibers are used to detect diffusely reflected light at distances ranging from 1.0 to 10.0 mm. The light from the detection fibers is measured by photodiodes connected to a lock-in amplifier, allowing precise measurements within a high dynamic range. They are the first to fit the approximating dipole model introduced by [EVNT78] to their data. Using this simple model, they show that it is in principle possible to derive all model parameters just by the shape of the fall-off, that is, they implicitly receive K from the fit of their model to the data. See Section 2.2.2 for a detailed model description. Nickell et al. [NHE⁺00] use the same device to analyze anisotropy of light diffusion within skin. In addition, they built an extended measurement probe consisting of a 2D bundle of optical fibers, coupled into a CCD receiver. This design allows to image directional scattering variations in a single measurement. Scale issues are not discussed, why it has to be assumed that K was again retrieved solely from the model fit.

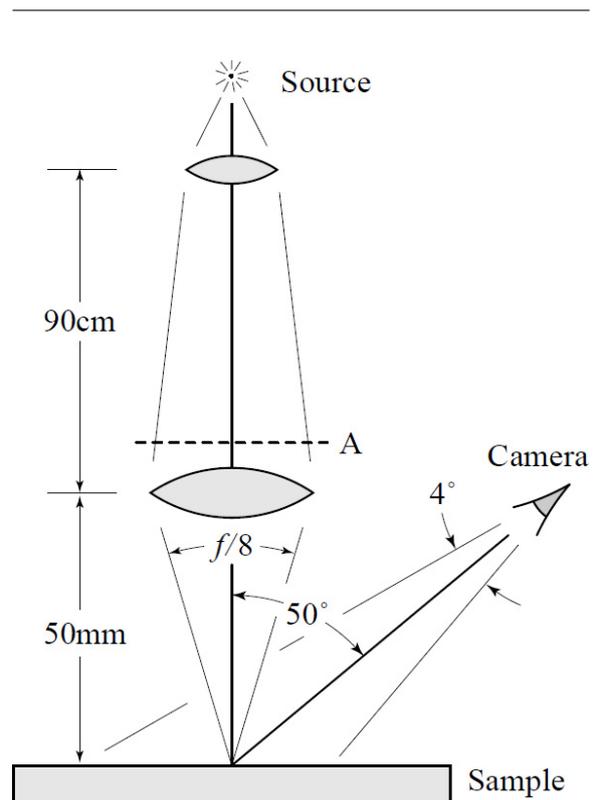


Figure 6.1: Contact-less measurement apparatus, as used by [JMLH01]. The camera observes exiting radiance around a spot of illumination. For evaluation, a cross-section that avoids single scattering is used. (Courtesy of H. W. Jensen.)

Jensen et al. [JMLH01] introduced the dipole model in a seminal paper to computer graphics. They also suggest a novel contact-less acquisition device, see Figure 6.1. The setup allows to acquire radiant exitance very close to the point of incidence. As a non-contact device, it allows measurements without interfering with the material. It is reported that despite of the excellent quality of the acquired data, a model fit with unknown K appears to be ill-conditioned, infer a large uncertainty especially in the total diffuse reflectance, which is of visual importance. They propose to explicitly calibrate for K using a calibrated reflectance target. Using a Spectralon™ sample that is calibrated to reflect 99.9% of the incident light, K is determined directly according to (6.4). Multiple translucent materials are analyzed, including skin at the forearm of two subjects. Recently, [JDJ06] extended the setup by [JMLH01] to include measurements where light source and camera are exchanged. This allows for direct measurements of the phase function. They also use a white reflectance target to calibrate for scale¹.

Static Objects

Scanning static objects, in contrast to *in vivo* measurements of human skin, allows for alternative, more time-consuming acquisition procedures. Gösele et al. [GLL⁺04] use a purely data-driven approach to capture translucent appearance. Instead of using an explicit reflectance model, they efficiently store radiance transport within an object in a hierarchical data structure. Using an RGB laser projector and a Silicon Vision Lars III monochrome HDR video camera that captures a dynamic range of over 1 : 1 000 000 at 15 frames per second, they acquire millions of images of an object under point-illumination by the laser. They use a hierarchical approach to store low-frequency light transport and model fine-scale detail by a modulation texture. It is clear that this approach is not suited for *in vivo* measurements, as object acquisitions takes up to 24 hours. Tong et al. [TWL⁺05] also use controlled laser illumination and a data-driven modeling approach. However, they use the dipole model to describe low-frequency radiance transport. Nothing is known on how K is determined. A large parameter mismatch of materials acquired from phantoms of known material properties suggests that K is derived from an instable model fit.

6.2 Acquisition Device

This section presents our own development of a subsurface scattering acquisition device, specifically built for skin measurements.

¹personal communication

6.2.1 Preparatory Experiments

Initially, experiments with a device similar to [JMLH01] have been performed. A critical issue with this design is motion during the measurement, as the observed surface point has to be kept precisely within focus of the illuminating beam, and the relative orientation of the skin surface must not change. Consequently, we aimed at a comparatively short measurement time. One way to achieve short measurements is to increase the brightness of the light source. An ideal light source provides a stable illumination within a wide spectral range and constrains the source of illumination to a small volume, which makes it easier to focus the light onto a fine spot. Gas arc discharge lamps, such as Xenon arc lamps, are a good choice, providing a good combination of these features. However, directly using an arc discharge bulb in our setup is prohibitive, as the bulb is under high pressure, and the operation temperature is high. The actual danger of explosion does not allow its use in proximity to a human face.

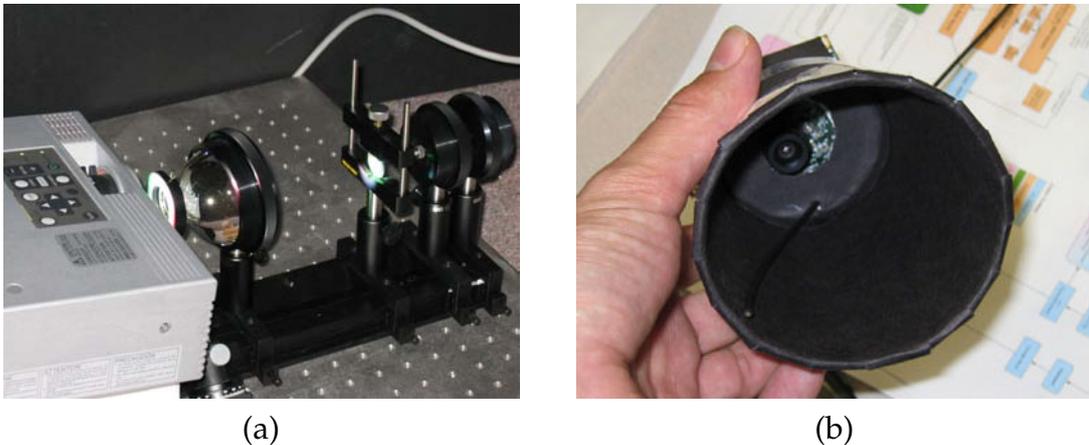


Figure 6.2: *Experimental subsurface scattering measurement devices. (a) Contact-free measurements using a lens system and a projector as a focused light source. Although short acquisition time became possible, the solution had to be discarded due to safety reasons. (b) Apparatus for light injection using an optical fiber with surface contact and remote sensing using a CCD camera. The device suffered from inter-reflections, falsifying the fall-off intensities.*

Instead, we decided to use an LCD projector as a light source, providing a safely mounted arc lamp within a cooled apparatus. Figure 6.2 (a) shows our experimental setup, using two convex lenses to focus a white projector image onto a small area of less than 1 mm diameter. The setup provides enough brightness to acquire a bracketed image sequence with very short exposure times, allowing for a full high-dynamic range acquisition under five seconds. However, it quickly became clear that the light spot on the skin gets too warm to be comfortable. Moreover, the setup can not be considered to be eye-safe when used for measurements within

human faces. As a result, increasing light intensity alone turned out to be not sufficient to address subject motion.

As a resort, we aimed at a mechanical alignment between measurement device and skin surface. Nickell et al. show that light can comfortably be led into the skin using optical fibers, fixing the point of illumination by surface contact [NHE⁺00]. After a short experiment with an illuminating fiber and an HDR camera imaging the surface (Figure 6.2 (b)), we turned toward a design that follows the linear probe by Nickell et al.

6.2.2 Design

Our subsurface measurement device is an image-based version of the linear fiber detector by [FPW92], see Figure 6.3. As a contact device, the sensor can reliably be placed at a fixed position, keeping this position even under moderate motion by the subject. We decided against a two-dimensional fiber layout, as the high curvature of the facial surface prohibits the planar placement of a wider sensor.

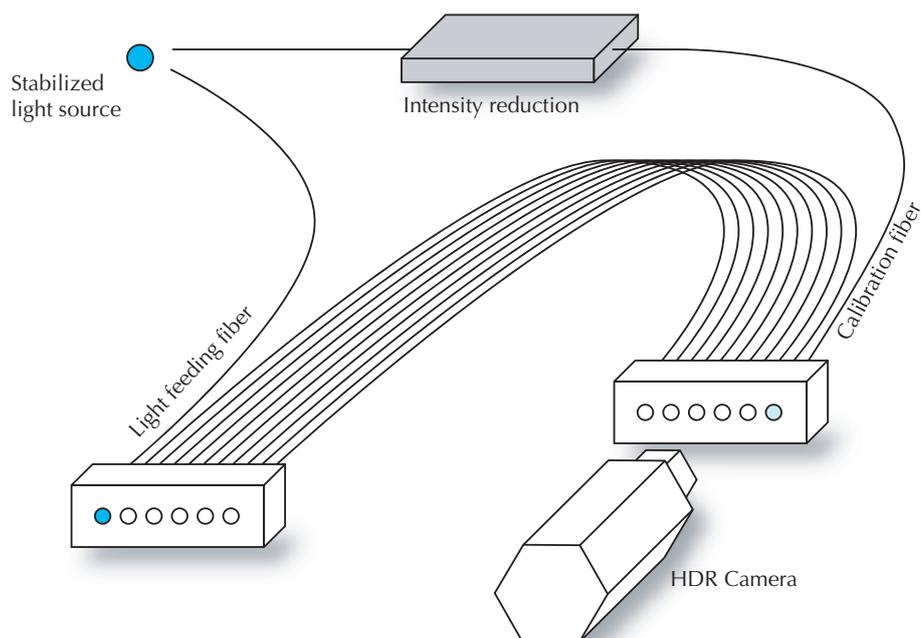


Figure 6.3: *Subsurface scattering measurement device. A schematic of the system.*

White light from an LED with stabilized power supply is coupled into a source fiber. A sensor head holds the source fiber and 28 detection fibers. Each fiber has a core diameter of $980\ \mu\text{m}$ and a numerical aperture of 0.51. In order to increase the radial resolution of the sensor, fibers are aligned in two interleaved rows, which

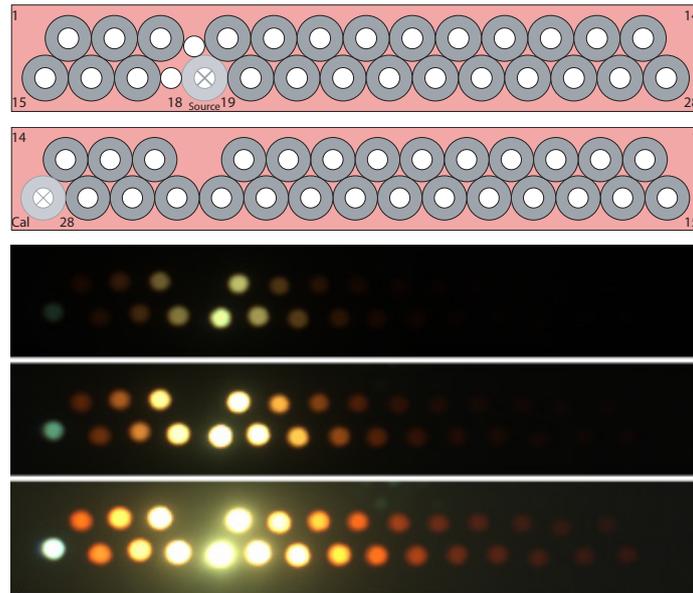


Figure 6.4: BSSRDF sensor fiber layout. Top: Fiber layout of sensor (top) and camera-side fiber bundle (bottom). Bottom: An HDR observation of the camera, displayed at three different exposures.

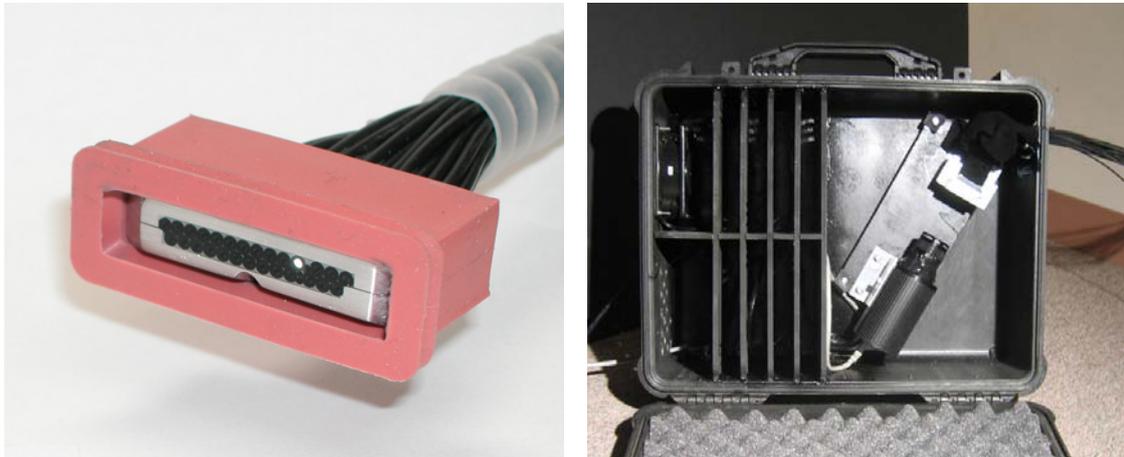


Figure 6.5: Left: A picture of the sensor head with linear fiber array. The source fiber is lit. Right: The fiber array leads to a camera in a light-proof box. The box is cooled to minimize imaging sensor noise.

increases the number of sampled distances. To further increase the resolution toward the point of illumination, a second, shorter array that is shifted by a quarter of a fiber diameter extends the sensor at the opposite side of the source fiber. The sensor layout is displayed in Figure 6.4. Figure 6.5 shows the sensor head and the housing of the acquisition camera.



Figure 6.6: Usage of the BSSRDF sensor.

A digital camera records the light collected by the detector fibers. We use a QImaging QICAM camera with a 1360×1036 10-bit per pixel CCD sensor. The camera and detector fibers are encased in a light-proof box with air cooling to minimize imager noise. We capture 23 images bracketed by $2/3$ f-stops to compute an HDR image of the detector fibers. The total acquisition time is about 88 seconds. Each fiber is imaged by approximately 80 camera pixels, further improving the signal-noise-ratio. Figure 6.4 shows an HDR image of the detector fibers with three different exposure values.

Figure 6.6 shows the sensor head placed on a face. We found that pressure variations on the skin caused by the mechanical movement of the sensor head influence the results. To maintain constant pressure between skin and sensor head we attached a silicon membrane connected to a suction pump. This greatly improves the repeatability of the measurements.

6.2.3 Calibration

Similar to Nickell et al. [NHE⁺00] we calibrate for differences in fiber transmittance by a flatfield measurement. We use a light table with an additional opal glass diffuser to generate a constant stimulus. An additional “black image” acquisition records light spill within the apparatus. In contrast to Nickell et al., however, who were mainly interested in anisotropy measurements, we have to calibrate for the proportionality constant K as well. Determining K by a reference image of an opaque reflectance standard, as done in [JMLH01], is not possible, as the sensor has to stay in contact with the sample during calibration. As soon as the sensor was lifted from the surface, the air gap would introduce further transitions of indices of refraction between sensor and sample surface, falsifying the measurement.

The solution has been to use a scattering medium of known parameters with a similar refraction index as skin. We decided to use skim milk, as it has been previously measured by Jensen et al., and because its composition is largely standardized across the U.S. We fit the dipole model to the measured skim milk data, fixing μ'_s and μ_a while solving for K . All calibration steps are depicted in Figure 6.7.

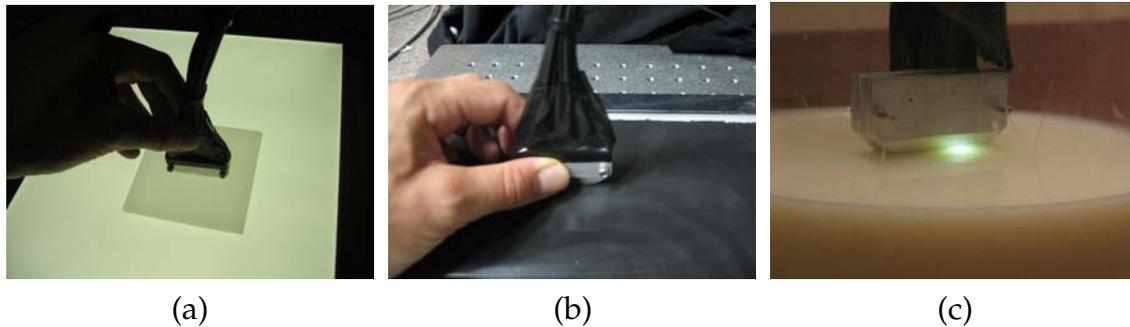


Figure 6.7: BSSRDF gun calibration procedure. (a) White calibration of relative fiber transmittance. (b) Black calibration for internal light spill and camera background noise. (c) Using skim milk as a secondary standard to calibrate for irradiance.

6.2.4 Measurement Procedure

During acquisition, it is crucial that the sensor stays in closed contact with the skin to prevent light spill between the sensor and the skin. Although the suction pump supports the contact, an operator is required to hold the sensor during the procedure. To keep the blood flow unaltered, the sensor should only gently touch the skin, no pressure should be applied. For obvious reasons, the sensor should be kept as still as possible during the acquisition.

The measuring device has been used to scan a large number of people. Great care has been taken to keep the sensor hygienically clean. Before and after each measurement, the sensor is cleaned using ethanol. In addition, we restrained from measuring lips.

6.3 Representation

The measurements deliver 28 radial samples of the skin's impulse response. Unfortunately, this data cannot directly be used for rendering. On one hand, it is still comparatively noisy. On the other hand, which is more severe, there is no data point at the source fiber itself. Hence, direct rendering of the data requires an extrapolation between $r = 0$ and the first sample point. For the near-exponential

fall-off function, this extrapolation is an ill-conditioned task. Consequently we decided to fit a subsurface scattering model to the data, providing smoothing and a realistic extrapolation toward $r = 0$ at the same time.

[SZW90] observe that steady-state diffuse reflectance is not very sensitive to the structure of layered tissues. [FPW92] and [JMLH01] confirm that a single-layered diffusion model and its dipole approximation does already fit well to experimental skin measurements. Hence, we believe that a single-layer dipole diffusion model is best suited for our purposes. Donner and Jensen [DJ05] show that visual accuracy can still be increased by simulating multiple skin layers, however, it is not clear how the additional model parameters, such as layer thicknesses, should be derived from the measured data.

6.3.1 Model Fit

We use a non-linear least-squares solver to fit the dipole diffusion model to our log transformed data points. Without a fixed K , we observed the same ill-conditioning as reported by [JMLH01]. Further attempts to improve the error norm did not succeed. However, convergence could greatly be improved by using previously measured skin parameters as starting conditions. Figure 6.8 shows an exemplary dataset and the corresponding model fit.

6.3.2 Re-Parameterization

Although μ'_s and μ_a are of immediate physical meaning, their effect on the visual material appearance is strongly entangled. A visually more intuitive parameterization of the diffusion model uses total diffuse reflectance $R_d(r)$, mainly responsible for the perceived skin albedo, and the effective transport coefficient μ_{tr} , that is closely related to the mean free path ℓ . It is

$$R_d(r) = \frac{\alpha' z_r (1 + \mu_{tr} d_r) e^{-\mu_{tr} d_r}}{4\pi d_r^3} - \frac{\alpha' z_v (1 + \mu_{tr} d_v) e^{-\mu_{tr} d_v}}{4\pi d_v^3}, \quad (6.5)$$

with $\mu'_t = \mu_a + \mu'_s$ the reduced extinction coefficient and $\alpha' = \mu'_s / \mu'_t$ the reduced albedo, and

$$\mu_{tr} = \sqrt{3\mu_a \mu'_t} \approx 1/\ell. \quad (6.6)$$

The mean free path ℓ is a measure of translucency, causing the soft appearance of shading and shadow boundaries upon human skin. Due to its visual importance, our subsequent analysis concentrates on μ_{tr} as the skin parameter most important for the perception of subsurface scattering.

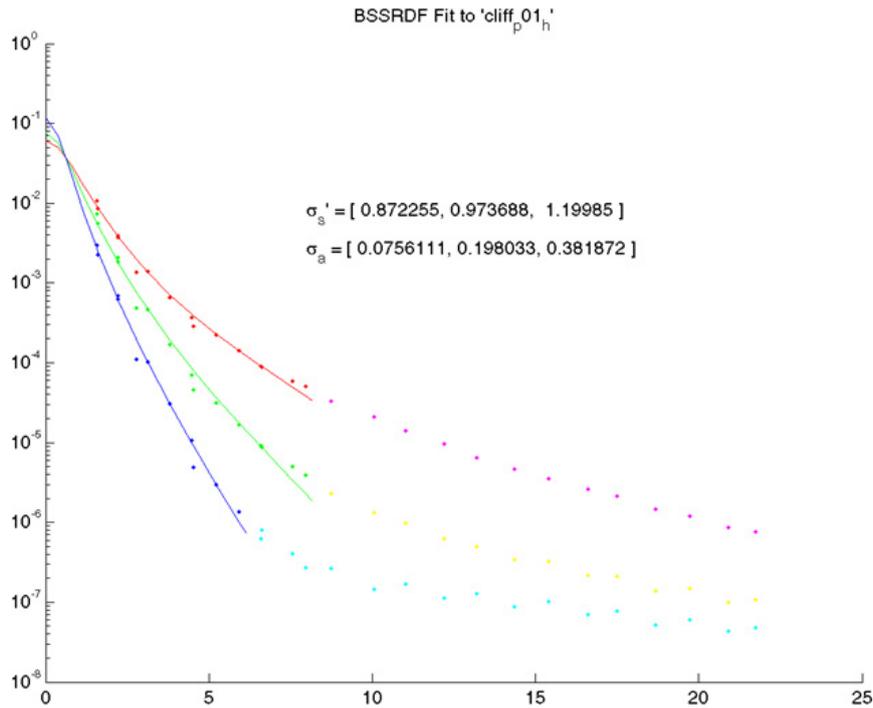


Figure 6.8: Subsurface scattering model fit in log space. The dipole diffusion model is fitted for each color channel separately. Data points that exceed a dynamic range of 3.5 orders of magnitude are not used for the fit.

6.4 Evaluation

We compared our measurements to values previously reported by [JMLH01] for Caucasian skin. Our fitted parameters μ'_s and μ_a were well in the same range, whereas the parameters slightly differed for Asian subjects. See Chapter 10 for a more detailed analysis of cross-subject variance of our measurements.

The first measurements have been performed without the silicon membrane. Although it has been possible to acquire stable measurements, great care had to be taken to minimize contact pressure. Attaching the suction silicon membrane to the sensor greatly improved precision and repeatability of the measurements.

6.4.1 Spatial Variance

Before measuring a larger population, suited measurement points had to be found. For that, a central question is how much variance in scattering properties there is across a single face. We measured two subjects (Asian and Caucasian males), positioning the probe at 52 different locations and orientations. See Figure 6.9 for a map of the respective measurement points. For many facial areas it

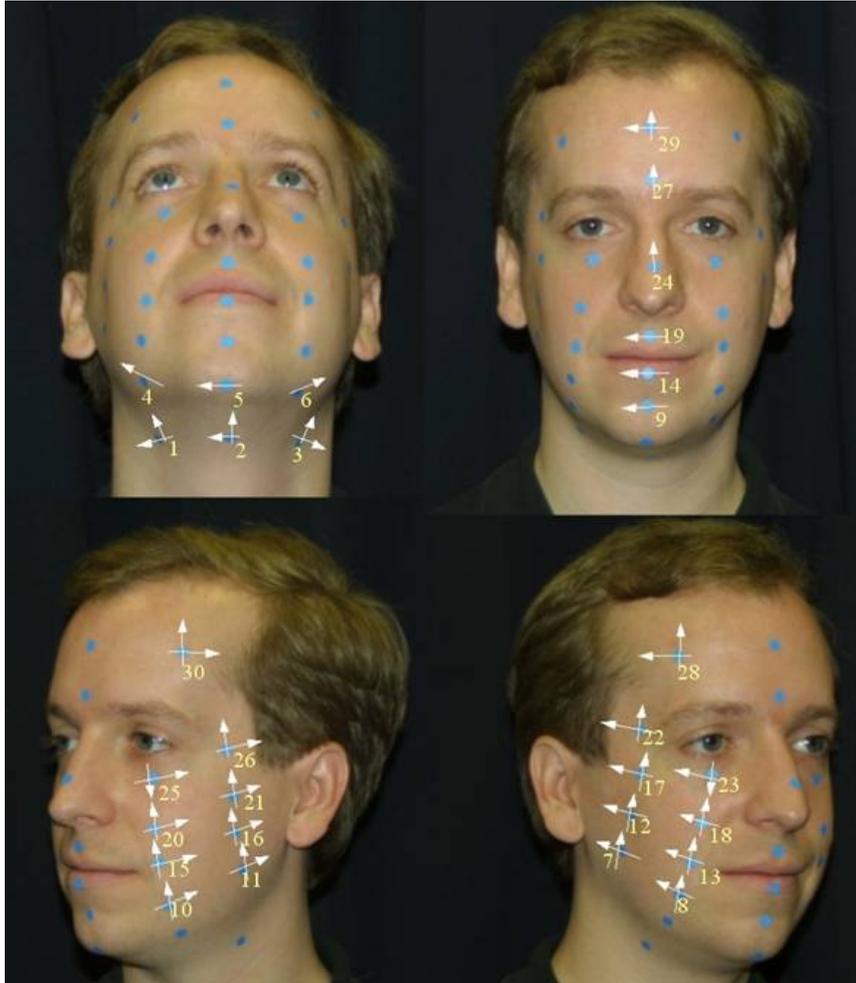


Figure 6.9: *Spatial translucency variance measurements. The sensor has subsequently been located at 30 locations and with up to two different orientations.*

proved to be difficult to place the sensor planar on the surface, affecting stability of the measurements. However, where stable measurements were possible, and comparing measurements on both sides of the symmetry plane, measured translucency variations appear to be within our measurement error. We conclude that for visual accuracy, it is sufficient to measure maximally three locations to obtain representative translucency values. Subsequent measurements were taken at the positions *cheek*, *neck*, and *forehead*, samples 16, 1, and 29 in Figure 6.9, respectively. All measurements have been taken with a horizontally orientated probe.

6.4.2 Anisotropy Measurements

As we use the dipole approximation to model our data, it is of particular interest whether the isotropic assumption of the underlying diffusion approximation

holds for human faces. Scattering anisotropy in facial skin, if existing, would not only be an important visual effect, but would also require subsurface scattering measurements along multiple directions. Nickell et al. report scattering anisotropy of varying degree for multiple body regions. From the regions that have been measured, abdomen is reported to show the highest anisotropy. So far, no respective measurements have been performed for facial skin.

In order to show that our device is in principle able to measure such anisotropy, we performed measurements on the abdomen of a subject. We subsequently place the sensor at 16 different orientations, covering 360° in steps of 22.5° . To obtain a regular placing of the sensor, we use a circular cardboard template that had been taped onto the skin. Figure 6.10 (a) visualizes the radially retrieved intensity values in dependence from the sensor's angular orientation. Note that raw sample data is displayed, that is, the data has not yet been smoothed by the model fit. The plot clearly reproduces the expected scattering anisotropy of abdominal skin.

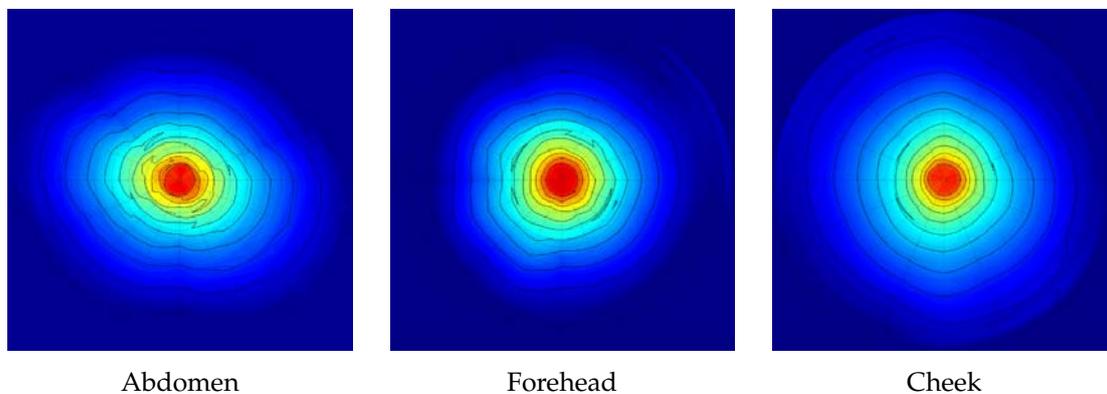


Figure 6.10: *Anisotropy measurements for subsurface scattering in human skin.*

Subsequent measurements at cheek and forehead, shown in Figure 6.10 (b) and (c), turned out to be close to isotropic. The remaining anisotropy is clearly too low to become visible in artificial renderings of skin. The complex facial geometry prohibits measurements of further facial regions, as the sensor has to be placed along multiple directions without colliding with curved surface regions. Future, more advanced, measuring techniques may provide a better coverage of facial anisotropy measurements, but as far as we can verify, facial skin anisotropy appears to be negligible.

6.5 Discussion

In this chapter we presented a custom-design of a subsurface scattering measurement device. We evaluated a number of existing designs, finally deciding for an

adaption of the design by Nickell et al. [NHE⁺00]. Our device is easy to use and allows for rapid measurements of a large number of subjects. An added suction mechanism improves repeatability and reliability of the measurements.

We conducted experiments on spatial variance and analyzed the scattering anisotropy of human skin. In contrast to other body regions, facial skin anisotropy turns out to be negligible for visual accuracy in computer graphics. Spatial variation of scattering parameters across the face is low. These insights greatly simplify our skin reflectance model, as described in Chapter 8. In particular, scattering isotropy allows to use the dipole diffusion model to describe facial subsurface scattering.

Processing Pipeline

Starting from the acquired geometry and reflectance data, synthetic face models have to be constructed. Before an actual skin reflectance model can be fitted, extensive pre-processing of the input data is required. This chapter describes the necessary steps of data refinement. As our project aims at the construction of a face database, a large number of faces has to be processed. Hence, the processing is largely automated in a data processing pipeline that allows for an unsupervised face model construction.

7.1 Overview

Figure 7.1 provides an overview over the processing pipeline. The raw geometry retrieved from the 3QTech 3D scanner is cleaned, parameterized, and up-sampled to obtain a highly resolved model of the facial geometry. Based on this geometry, the acquired reflectance field is re-parameterized into a *lumitexel* representation [LKG⁺01]. This requires consideration of photometric and geometric calibrations as well as the computation of camera and light source visibilities for every vertex of the geometry. Based on the *lumitexel* representation, detailed surface normals are estimated. The 3D scanner resolution itself does not suffice to capture fine-scale normal variations. As an optional path in the processing pipeline, it is possible to use the normal estimates to refine vertex positions in the source geometry, and to re-iterate the processing pipeline. As a result of this processing

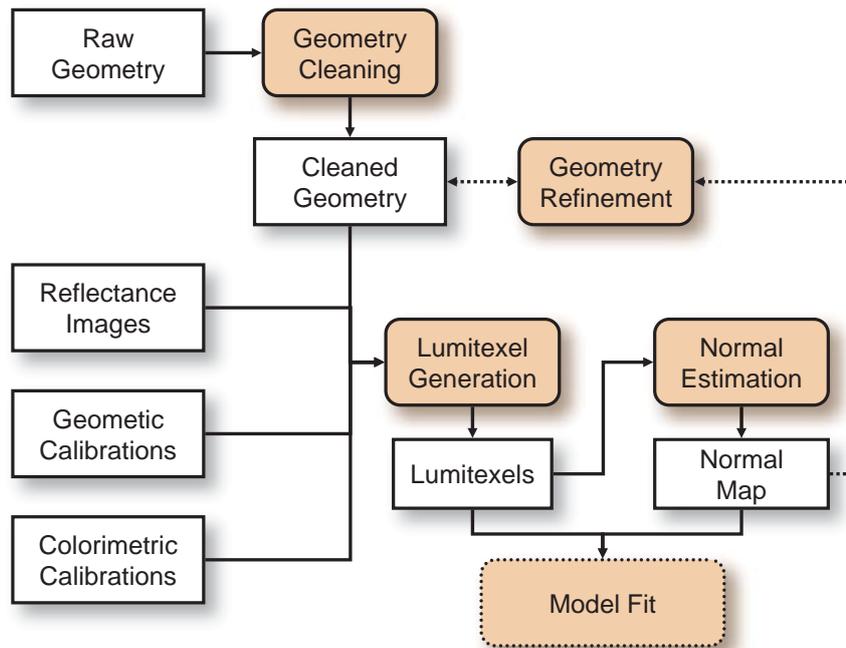


Figure 7.1: Overview over the processing pipeline. An optional feed-back loop refines the acquired geometry based on normal estimates.

pipeline, lumitexels and a refined normal map are used in a subsequent model fit. The remainder of this chapter describes the different pipeline stages.

7.2 Geometry Processing

The 3QTech 3D scanner provides a triangle mesh that typically consists of around 30 000 vertices, covering a frontal view of head and shoulders. In the geometry processing stage, the geometry has to be processed to end up with a clean manifold, single-boundary mesh of the facial geometry. During processing, geometry that does not belong to the face, e.g. hair, ears, or parts of the upper body, have to be removed. Mesh degeneracies and non-manifold situations have to be removed, and the mesh is up-sampled, so that every vertex corresponds to a surface sample for the spatially varying reflectance reconstruction. In addition, a continuous uv parameterization of the mesh is computed.

The face scanner processing pipeline generally uses two different geometry representation. For visibility computation, lumitexel generation, and model fit, we use an elliptical surfel geometry, as described in Section 4.2. In addition to this, a traditional mesh structure is maintained for rendering and surface parameter-

ization, as this allows us to use existing tools. Both models have a one-to-one correspondence between surfels and mesh vertices.

The parallel use of triangle meshes forced us to drop the purely point-based scan cleaning, as described in Section 3.1: while the conversion from meshes to point clouds is a reliable technique, the opposite direction of conversion is not yet stable enough to be used in our project where large numbers of models have to be processed. We experimented with the surface reconstruction algorithm COCONE, [ACDL00, DG01], but our data did not always meet the algorithm’s sampling criterion and low-noise requirements. Consequently, the geometry pre-processing is completely triangle-based, converting only the final mesh to a surfel geometry.

In a first version, all processing steps have been fully automated, only requiring the area of interest to be manually specified. The automated pipeline aims at a rapid processing of a large number of scans, as needed for unsupervised cross-subject analysis. However, as it turned out, automated cleaning was not sufficient to meet the requirements to produce highest quality datasets that allow for photo-realistic renderings of individual datasets, which agrees with Section 3.1. Hence, we alternatively use a manual geometry processing pipeline that allows careful tuning of the models.

While the automated processing is designed to run unsupervised on a large cluster of Linux workstations, the manual pipeline has been ported to Windows, as the cleaning makes use of an interactive Windows application. The windows port generally turned out to be a wise decision, as the actual Linux cluster that we had access to turned out to become unreliable toward the end of our project. Figure 7.2 provides an overview over both the automated and the manual geometry pipeline. The remainder of this section presents their constituents.

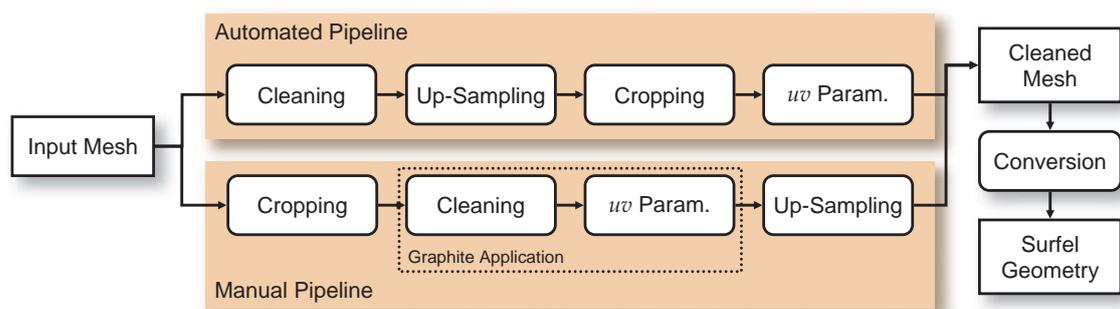


Figure 7.2: *Two alternative geometry pipelines. The automated pipeline is designed for rapid processing with minimum user interaction. The interactive pipeline provides more control to produce optimal results.*

7.2.1 Mesh Cleaning

The automated mesh cleaning software performs the following steps in the presented order:

1. Removal of all non-manifold triangles that share an edge of more than two adjacent triangles in total.
2. Cluster analysis of the mesh. Removal of all triangle clusters but the biggest one. This leaves the facial surface as a single cluster, also removing potential stray clusters produced by step 1.
3. Hole filling. The software determines all closed boundaries of the mesh and keeps the largest one as the facial boundary, considering the remaining smaller boundaries to define holes in the surface. These hole boundaries are filled using triangle fans.

In general, this leads to a single, hole-free, mesh surface, which is manifold and has one boundary. The software does currently not cope with non-orientable manifolds where front and back side of the mesh coincide, a situation as found in a Moebius band. Unfortunately some of the meshes output by the 3QTech scanner contain such conditions, even though mostly outside of the facial region. Inconsistent triangle orientations, however, are fixed as long as the surface is orientable.

In the manual variant of the cleaning procedure, the mesh is interactively checked for degeneracies, orientability, and non-manifold situations using Graphite, a GPL open source application courtesy by Bruno Lévy [Gra03]. The manual editing provides finer control over the cleaning process.

7.2.2 Up-Sampling

Sample generation for normal estimation and model fit requires a vertex per surface sample. Hence, the number of vertices in the mesh has to be increased, until the sampling density lies beyond the sampling by the camera pixels. For efficiency reasons, an even sample distribution is desired. This could theoretically be achieved using a particle simulation as presented in Section 3.3.3, however, the 3Q Tech scans already show an even triangle distribution, with well-shaped triangles of approximately the same size. This allows us to use simple triangle subdivision schemes to generate up-sample the geometry.

We implemented Loop and modified butterfly subdivision [Loo87, ZSS96], both primal subdivision schemes that typically lead to 700 000 to 1 000 000 vertices after

three subdivision steps, corresponding to an average sample spacing of under $200\mu\text{m}$, which meets the camera resolution very well.

The modified butterfly algorithm as an interpolating, i.e. non-shrinking, subdivision scheme would be expected to be better suited for up-sampling. However, the noisy input data also requires a smoothing step, which is implicitly performed by the Loop subdivision. Hence, Loop subdivision became our preferred choice. As the shrinkage due to the subdivision scheme was below the geometric error of the scanner, we decided not to further investigate into alternative techniques. Possible alternatives would have been non-subdividing low-pass filters that minimize shrinkage, such as Taubin filtering, implicit fairing, or Laplacian smoothing [Tau95, DMSB99, JLW05], followed by an interpolating subdivision.

7.2.3 Cropping

We decided to specify the facial region of interest using a 2D interface: prior to the geometry processing, the user has to paint a mask on top of the acquired 3QTech texture, selecting all parts of the surface that should be kept. This boils down to painting a selection within Photoshop, which is a much simpler interface in comparison with a 3D editing tool. Cropping a dataset using these *skin masks* simply requires to project each vertex into the texture domain and to look up the mask value, discarding vertices outside the selection. Figure 7.3 shows an exemplary skin mask.

In the automated pipeline, cropping is deferred until after the up-sampling step to prevent jaggy crop boundaries. This requires to maintain the 3QTech uv parameterization throughout the pipeline, which complicated the implementation of the subdivision schemes in the presence of discontinuities in the parameterization. When porting the geometry pipeline to Windows we decided to use a subdivision implementation from the OpenMesh library [BSBK02] instead, which does not provide subdivision with maintenance of uv -coordinates. Hence, skin mask cropping had to be performed *before* editing and up-sampling of the the mesh. Within the manual pipeline, however, the resulting jaggy crop boundaries can still be smoothed manually. Figure 7.4 shows a source geometry and the manually cleaned and cropped model before up-sampling.

7.2.4 uv Parameterization

Our implementations of the geometry refinement, rendering, and some operations during the model fit benefit from an existing uv parameterization of the face model. The parameterization allows to interpolate surface attributes across the

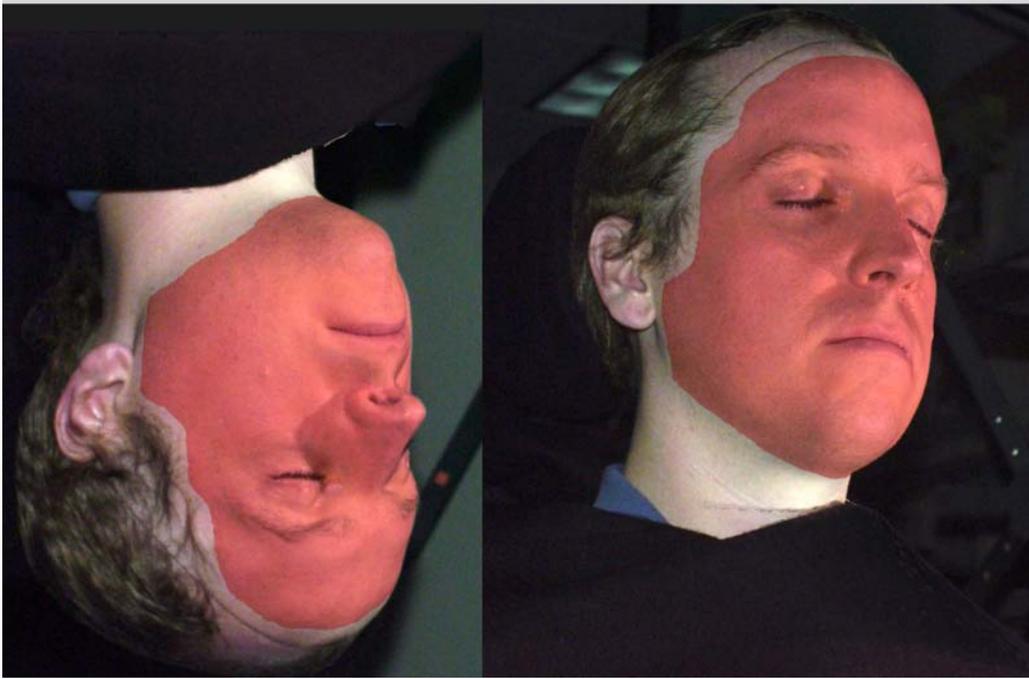


Figure 7.3: A skin mask superimposed over the acquired 3QTech texture.

surface, and allows to transform different datasets into a common parameterization frame. During rendering, the uv parameterization is required to encode model parameters in surface textures.

The original 3QTech parameterization is not well suited, as it contains discontinuities and considerable stretch. A continuous, conformal, area-preserving parameterization would be desired. In general, not all of these properties can be fulfilled at the time, but various techniques exist to compute a parameterization that concurrently minimize angular distortion and stretch. Our automated processing pipeline contains an implementation of the intrinsic Desbrun parameterization [DMA02] that allows to control the trade-off between conformality and area

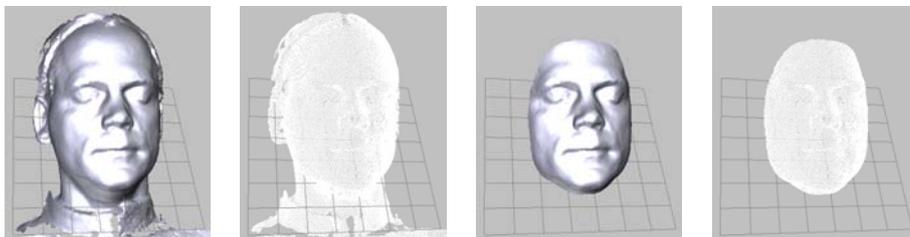


Figure 7.4: Left: Source geometry acquired with the 3QTech scanner. Right: Manually cleaned geometry, cropped to the face region using a skin mask.

preservation. The algorithm requires a pre-defined mapping of the boundary into the parameterization domain. To produce a rotationally invariant parameterization, we map the mesh boundary onto a circular region in the parameterization domain, preserving the relative lengths of the boundary edges. For our meshes with hundreds of thousands of vertices, computing the parameterization takes in the order of 30 minutes. Determining the natural intrinsic parameterization after [DMA02], however, which does not require a predefined boundary mapping would have been too computationally expensive for our mesh sizes. Figure 7.5 shows iso-lines of a Desbrun parameterization on one of our face scans.

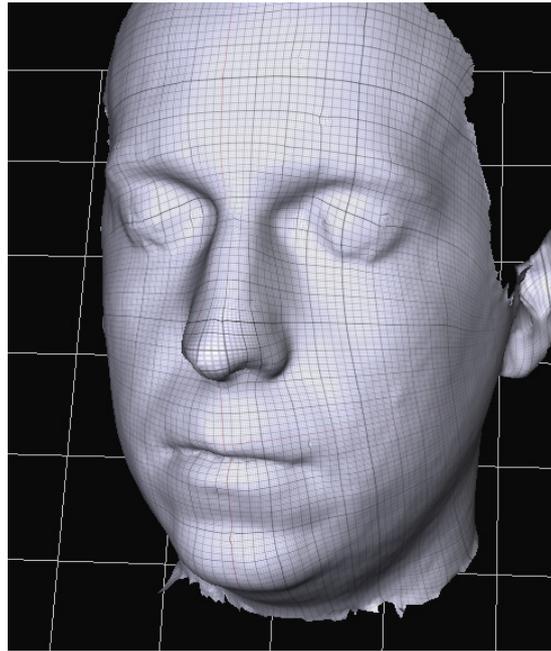


Figure 7.5: *Desbrun parameterization of a face model.*

In the manual processing pipeline we decided to use a faster parameterization, as it recently became available as an extension within the Graphite application: ABF++ introduced by Sheffer et al. generates conformal parameterizations with low stretch and is fast enough to deal with very large meshes [SLMB05]. To further increase the processing speed, we decided to compute the parameterization for the low-resolution mesh before up-sampling, reducing the processing time to the order of seconds. As the computed parameterization is continuous, a parameterization of the up-sampled mesh can be obtained by treating the uv coordinates analogously to the vertex positions within the subdivision scheme used for up-sampling. Although the resulting parameterization is not strictly conformal anymore, it is of sufficient quality for our purposes. See Figure 7.6 for a comparison of the two parameterizations.



Figure 7.6: Comparison of face model parameterizations. Left: Circularly constrained intrinsic Desbrun parameterization. Right: ABF++ parameterization. (Albedo maps of two different subjects are shown; cf. Section 8.2.1.)

7.3 Visibility Classification

Reflectance sample generation from the reflectance field data requires to take occlusion into account. In contrast to image-based methods, such as the previously presented Surface Reflectance Fields [MPN⁺02] where lighting is implicitly encoded in the dataset, it is not sufficient to determine only camera visibility for each surface point. In addition, for each light source it has to be determined whether it directly illuminates the surface point, as the model fit expects direct observations. It is crucial that the visibility is conservative with respect to both cameras and light sources to prevent erroneous samples from influencing the model.

Previous work implicitly determines camera visibility during geometry construction [MBR⁺00, MPN⁺02], which is not applicable to light source visibility, or uses traditional shadow mapping [Wil78, Len04], suffering from potential aliasing effects. Hence, we decided to use an alternative technique that is still related to the shadow map algorithm but circumvents its sampling restrictions. Similar to [Len04], occlusion is computed based on the acquired geometry. Our approach renders the face geometry from the cameras, or light sources, respectively. As a key feature, rendering is performed using a modified EWA surface splatting software implementation after [ZPBG01], directly rendering the surfel geometry that defines the reflectance sample locations. Due to the EWA screen-space filter, each sample point is guaranteed to be projected onto at least one pixel. This makes

depth buffer resolution less critical, as no sample can fall “between” the pixels, and effectively removes mis-classifications due to aliasing.

The renderer implementation allows to determine for each surfel the number of n_s screen samples it projects to. Furthermore, we attach to each rendered pixel the list of surfels that contribute to this pixel. From these lists, each surfel’s number n_v of visible samples is computed. A surfel is considered visible, if a certain percentage ϑ_v of its samples contributes to a visible pixel, that is, if and only if

$$\vartheta_v \leq \frac{n_v}{n_s}. \quad (7.1)$$

We usually choose $\vartheta_v = 0.6$. Increasing ϑ_v makes the visibility test more conservative.

The acquired datasets are never entirely free of motion artifacts, and the precision of the 3D scanner lies in the order of a millimeter. Both error sources can lead to mis-classifications of surface observations in the vicinity of depth discontinuities: a sample on the cheek might be considered visible, although it is just occluded by the tip of the nose. If samples were drawn from this configuration, samples from the nose would be assigned to this cheek sample, which invalidates this sample and may lead to visible artifacts. In order to keep the visibility computation conservative, samples around depth discontinuities are masked out. In particular, samples at a pixel p are discarded if its circular neighborhood $\mathcal{N}_r(p)$ with radius r covers a depth-range larger than ϑ_d , i.e., if

$$\vartheta_d < \max_{q,r \in \mathcal{N}_r(p)} \|\text{depth}(q) - \text{depth}(r)\|. \quad (7.2)$$

Figure 7.7 shows three depth range images as they are used to mask visibility continuities. The proposed heuristic also masks out areas where the depth image contains a large slope, see Figure 7.7 (c). This is a desired side-effect, as these areas contain extremely grazing observations that would be too unreliable anyway. A similar masking of depth-discontinuities can be found in [LKG⁺01], however, they use a Gaussian filter to propagate depth-discontinuities across the image-plane, introducing a dependency from the shape and size of a depth-discontinuity that is hard to predict.

The depth discontinuity masking works very robustly, and it was possible to fix ϑ_v , r , and ϑ_d for all datasets, only choosing different values for cameras and light sources. Nevertheless, when fitting normals and BRDFs to the data, we still observed artifacts along visibility discontinuities with respect to the light sources. These artifacts could not be removed by changing depth discontinuity thresholds. They are due to the spatial extent of the light sources, occurring in penumbra regions where the light sources are only partially visible, while the visibility computation, treating light sources as point lights, considers the light source fully visible. This leads to under-estimated reflectance at these samples. Although

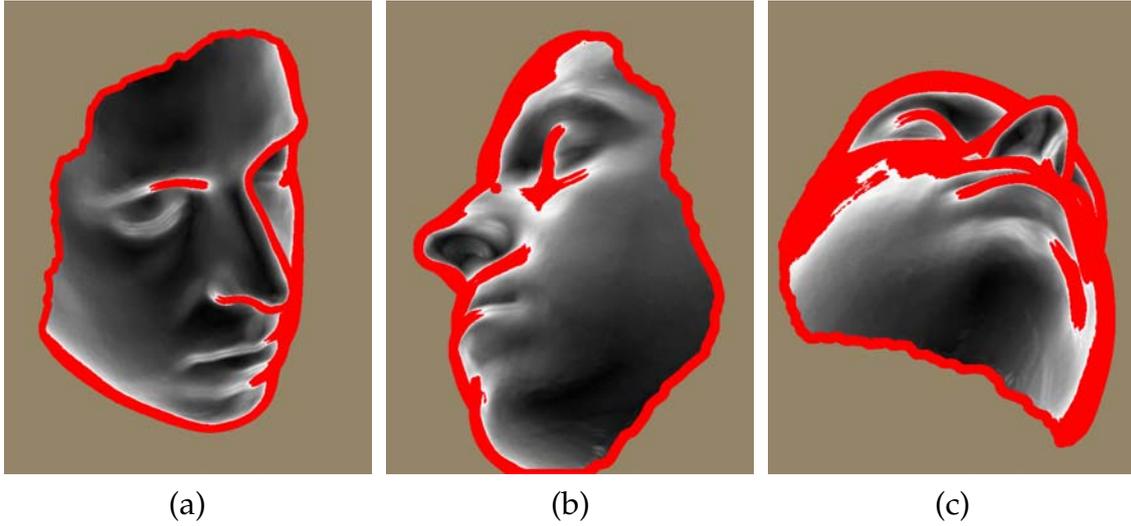


Figure 7.7: Depth discontinuity images. Brighter values at a pixel p denote a larger depth-range within $\mathcal{N}_r(p)$. Pixels where the depth-range exceeds ϑ_d are marked in red. Samples at these locations will be ignored. (a) Shows depth discontinuities as seen from one of the cameras. (b) and (c) show light source perspectives. Note that oblique observations are also masked out.

explicit treatment of the light sources' extents would have been an option, we decided to discard these samples by eroding the set of visible light sources: a light is only assumed fully visible if it is completely surrounded by other light sources in the dome that are reported as visible by the depth buffer algorithm.

The visibility computation described here is also used by the deformable surface reflectance field evaluation presented in Section 4.2.

7.4 Sample Generation

Any further processing of the reflectance field data is based on a *lumitexel* representation [LKG⁺01]. For each point x_o , a lumitexel stores reflectance observations of the form $R_i = (r^i, \omega_1^i, \omega_o^i) \in \mathcal{R}(x_o)$, relative to a local surface frame $(\mathbf{n}, \mathbf{t}_1, \mathbf{t}_2)$. r^i stores the cosine-weighted surface reflectance under incident angle ω_1^i and outgoing angle ω_o^i :

$$\begin{aligned} r^i &= \langle \mathbf{n}, \omega_1^i \rangle f_r(x_o^i, \omega_o^i, \omega_1^i) \\ &= \langle \mathbf{n}, \omega_1^i \rangle f_r^i. \end{aligned} \quad (7.3)$$

Deferring the cosine normalization until the model fit allows for readjusting the normal directions at a later point in time. f_r^i can be determined from the respec-

tive reflectance field image: \mathbf{x}_o projects to a camera pixel with an intensity proportional to

$$L(\mathbf{x}_o^i, \boldsymbol{\omega}_o^i) = \int_{\Omega_+} f_r(\mathbf{x}_o^i, \boldsymbol{\omega}_o^i, \boldsymbol{\omega}_i) L(\mathbf{x}_o^i, \boldsymbol{\omega}_i) n d\boldsymbol{\omega}_i . \quad (7.4)$$

The respective proportionality constant is part of our calibration procedure, as described in Appendix B.3.2. Considering the small angular footprint Ω of the light source, f_r can be approximated to be constant for $\boldsymbol{\omega}_i \in \Omega$, leading to

$$\begin{aligned} L(\mathbf{x}_o^i, \boldsymbol{\omega}_o^i) &\approx f_r(\mathbf{x}_o^i, \boldsymbol{\omega}_o^i, \boldsymbol{\omega}_i) \int_{\Omega_+} L(\mathbf{x}_o^i, \boldsymbol{\omega}_i) n d\boldsymbol{\omega}_i \\ &= f_r^i E(\mathbf{x}_o^i) . \end{aligned} \quad (7.5)$$

$E(\mathbf{x})$ is known from our calibration, as described in B.3.2. Hence, f_r^i is determined from (7.5) as $f_r^i = L(\mathbf{x}_o^i, \boldsymbol{\omega}_o^i) / E(\mathbf{x}_o^i)$.

On average, each lumitexel contains about 900 BRDF samples per color channel, with many lumitexels having up to 1 200 samples. The numbers vary depending on the lumitexel visibility. Color correction, incorporation of the light source calibration, and HDR evaluation are performed on the fly during the lumitexel generation. Figure 7.8 shows a visualization of lumitexels from two different points in the same face. The images give an impression of the hemispherical sampling for each camera viewpoint.

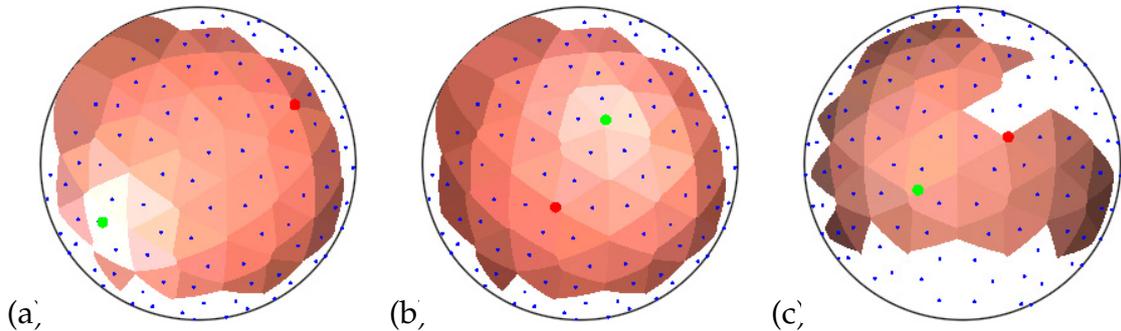


Figure 7.8: Visualization of two skin lumitexels for different viewpoints. Only the samples for a single camera are shown at the time. The red and green dots are the intersection of the view vector and reflection vector with the hemisphere, respectively. Reflectance values for each light position (blue dots) are used to shade the hemispherical triangulation. (a),(b) Two different camera views of the same lumitexel. Note that the reflection direction corresponds well with the specular highlight. (c) This lumitexel is more diffuse, and some light positions are occluded by the nose.

All lumitexels are sequentially stored in a lumitexel stream that occupies up to 7 GB for a face. The stream is indexed by surfel indices, which means an exhaustive

reordering of the original dataset where the data is organized in raw camera images, indexed by camera ID and light ID. As the overall size of the data requires out-of-core processing, special care had to be taken with respect to the data access pattern during lumitexel generation. In a first step, our implementation generates 30 independent lumitexel streams, each of them taking its samples only from five out of all 150 light sources. As the partial reflectance field of five light sources fits into main memory, the required data shuffle operations can be performed efficiently without generating additional file system traffic. The separate lumitexel streams are then merged in a second step. Merging requires simultaneous read from all partial streams and allows sequential writing of the final lumitexel stream. This procedure speeds up the lumitexel generation significantly, however, the process still takes in the order of 40 minutes.

7.5 Normal Estimation

Normal estimation is a critical point in the processing pipeline. Accurate normals are a necessary pre-condition for the BRDF fit, and a faithful reproduction of high-frequency normal variations is crucial for high-fidelity face renderings. We evaluated four different approaches to derive an individually corrected normal \hat{n} at each lumitexel [Bic06].

7.5.1 Normals From Specular Direction

For pure specular reflectance, the law of reflection states that the reflection vector r and the light vector l are co-planar with the normal vector \hat{n} , and the angle of incidence is equal to the angle of reflection. We exploit this fact in a normal estimator based on specular highlight direction. Conceptually, we determine for each camera viewpoint the light source that contributes the brightest sample, using the direction to this light source as l . As the brightest observation, this lighting direction is assumed to reflect light in direction v of the camera. To avoid quantization due to the discrete light source spacing, we use an interpolator known from unstructured lumigraph rendering [BBM⁺01] to interpolate l between the directions to the three brightest light source observations (cf. Section 2.5).

With this approximation to l and r , the normal vector can easily be estimated as

$$n_i = \frac{l + r}{\|l + r\|}, \quad (7.6)$$

delivering a normal estimate for each camera i . For the micro-normals (see Section 8.1.3) in our model, we directly use these normals.

If a single, view-independent normal is required, we determine $\hat{\mathbf{n}}$ by applying a voting scheme. The normal is interpolated from the three normal estimates that agree most. This is implicitly performed by determining a per-camera weight

$$w_i = \sum_j \langle \mathbf{n}_i, \mathbf{n}_j \rangle. \quad (7.7)$$

Using the same interpolator as before, $\hat{\mathbf{n}}$ is interpolated between the three \mathbf{n}_i with highest w_i . This effectively discards outliers, for instance normal estimates from cameras where no specular reflection is observed. This method, as a colloquial term internally referred to as “Weyrich’s method”, has been designed to preserve high frequency details in the presence of imprecise geometry and motion artifacts. It has to be noted, however, that the underlying reflectance assumption does not necessarily hold for skin, where off-specular highlights have been reported [MWLT00].

7.5.2 Helmholtz Stereopsis

This normal estimator is based on the method by Zickler et al. [ZBK02]. The method does not use any prior, such as specular observations or Lambertian reflectance, making it an approach of maximum generality. It exploits Helmholtz reciprocity, which states that any light transport, and surface reflectance in particular, is symmetrical and can be inverted. Exchanging light source and camera in a given setup does not change the ratio of irradiance and radiant exitance. Figure 7.9 shows such a setup, acquiring a reciprocal pair of images. This fact allows

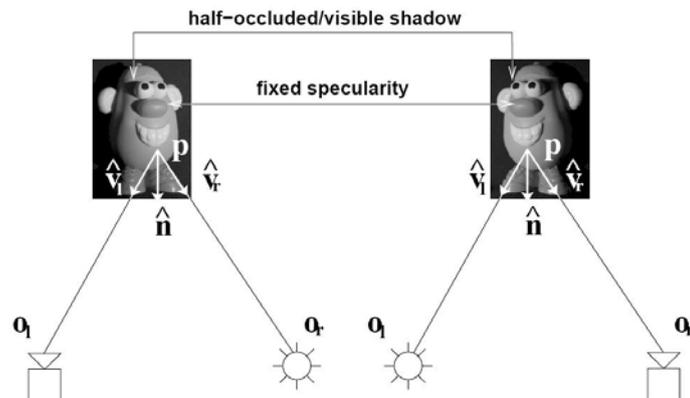


Figure 7.9: Setup for Helmholtz reciprocity [ZBK02]. Two images are acquired, whereby the only difference between them is that the positions of camera and point light source are swapped. (Potato man by courtesy of Todd Zickler.) (Will be changed to an example with faces .)

us to determine normals without making any assumptions about the BRDF of the surface.

Given a reciprocal pair of images, the image irradiance values i_l and i_r observed at the same surface point \mathbf{p} satisfy the constraint

$$\left\langle i_l \frac{\hat{\mathbf{v}}_l}{|\mathbf{o}_l - \mathbf{p}|^2} - i_r \frac{\hat{\mathbf{v}}_r}{|\mathbf{o}_r - \mathbf{p}|^2}, \hat{\mathbf{n}} \right\rangle = 0, \quad (7.8)$$

where \mathbf{o}_l and \mathbf{o}_r are the left and right camera centers, and $\hat{\mathbf{v}}_l$ and $\hat{\mathbf{v}}_r$ are the normalized view vectors. If \mathbf{p} and the camera centers are known, the normal $\hat{\mathbf{n}}$ is constrained by i_l and i_r . Two to three reciprocal image pairs suffice to determine the normal [ZBK02].

The acquisition dome configuration, however, does not allow to exchange camera and light position. Instead of using a real point light source, we approximate a point light source at a camera's position by superimposing the illumination of a ring of dome light sources adjacent to the camera. Figure 7.10 shows an example.

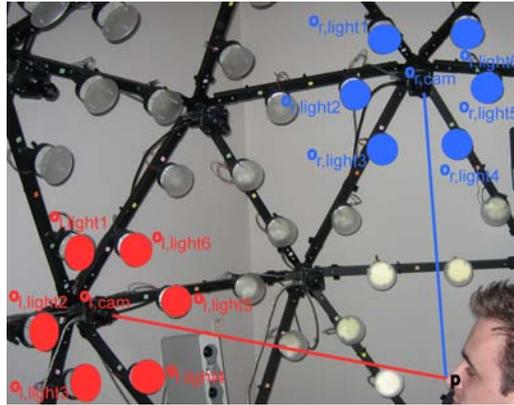


Figure 7.10: Approximate Helmholtz reciprocity with our dome setup. The point light source required by equation (7.8) is approximated by taking an average of light sources surrounding a camera.

The light source rings contain 5 or 6 light sources, respectively, depending on the location of the camera node. The joint illumination by the light sources is computed from each lumitexel by averaging the respective light source observations.

This allows us to construct several reciprocal pairs with the dome setup for each surface element. The exact number of possible reciprocal pairs depends on light source and camera visibility and can vary significantly between lumitexels. A light and camera pair is accepted only if all lights surrounding the camera are visible from the surface element \mathbf{p} . All observations build a linear system according to equation (7.8). The system is generally over-determined and is solved in the least-squares sense using singular value decomposition (SVD).

7.5.3 Standard Photometric Stereo

A very common approach for normal estimation is photometric stereo. The method is based on the assumption that the surface is Lambertian. Reflectance by the Lambertian law can be written as

$$\mathbf{i} = \frac{1}{\pi} \langle \mathbf{l}, \hat{\mathbf{n}} \rangle \mathbf{c}, \quad (7.9)$$

where $\mathbf{i} = (i_r, i_g, i_b)^\top$ is the observed intensity for a particular illumination direction \mathbf{l} of a surface point with albedo $\mathbf{c} = (c_r, c_g, c_b)^\top$. Multiple observations with different illumination directions $\mathbf{l}_1, \mathbf{l}_2, \dots$ yield an intensity matrix $I = (\mathbf{i}_1 | \mathbf{i}_2 | \dots)^\top$, which again leads to an over-determined linear system of equations that can be solved in the least-squares sense.

Color channels are theoretically independent, that is, the method should yield the same normal when applied to each color channel separately. Due to camera noise and Bayer pattern interpolation, however, this is not the case, potentially making photometric stereo unstable. To improve the conditioning of the problem, Barsky and Petrou [BP01] introduce an improved approach, reducing the problem to a monochromatic case using principal component analysis (PCA). The idea is that all observed colors should be located on a line through the origin in color space, that is, they correspond to the same color direction. This is a natural consequence from equation (7.9). Due to noise, the observations appear to deviate from this line. Barsky and Petrou suggest to recover the color direction using PCA. Projecting all observations onto the axis of main variance yields scalar intensities that are then used to determine the normal according to (7.9).

A significant drawback of the photometric stereo method is the assumption that the surface is diffuse Lambertian. This is obviously not true for skin. However, as we have hundreds of intensity measurements for each surface position, we only choose values in a range of the actual skin reflectance that comes close to the Lambertian assumption. We do this by sorting \mathbf{i}_k according to their values of the green color channel, which has a high signal-to-noise ratio, and select the 30 to 50 percentile. The upper 50 percent are discarded because they contain with high probability a specular contribution which definitely cannot be modeled by equation (7.9), and the lower 20 percent could contain invalid measurements caused by shadows of very small objects like single hair strands.

7.5.4 Photometric Stereo With Highlight Removal

Mallick et al. [MZKB05] present a photometric stereo method that explicitly addresses the presence of specular surface reflectance. The idea is that the specular reflectance of an object can be separated from the diffuse, Lambertian, reflection by a data-dependent rotation of the color space. Images in the transformed color

space, which are approximately free of specularities, are then used for normal reconstruction with traditional photometric stereo.

The method relies on the dichromatic reflectance model. It assumes that reflectance can be separated into the sum of a diffuse part of the material color and a specular part that has the color of the light source. In our case, the setup is colorimetrically calibrated so that each light source color is $\mathbf{s} = (1, 1, 1)$. Projecting all color observations into the plane orthogonal to \mathbf{s} produces color samples that are free from any specular color. The projected samples are then used to compute a normal using a photometric stereo implementation after [BP01].



Figure 7.11: Results of the normal estimation step. From left to right, top row: normals from Specular Direction, Helmholtz Stereopsis. Bottom row: photometric stereo with highlight removal, standard photometric stereo.

7.5.5 Comparison

Figure 7.11 compares the result of the four normal estimation techniques we described. The per-vertex normals are encoded in a normal map. The data is densely interpolated using push-pull interpolation into maps of 2048×2048 resolution. Apparently, the approximate Helmholtz stereopsis method fails in the chin region. For this facial area not enough light/camera pairs are visible and a reliable normal estimation is not possible. Therefore we conclude that for our requirements the approximate Helmholtz stereopsis method is not competitive with the other two methods, given the layout constraints of our acquisition dome. Normal estimation from specular highlight direction preserves high frequency details very well, but it shows some discontinuities and increased normal noise. The photometric stereo method with a transformation in color space suffers from over-smoothing. Fine facial details, e.g. wrinkles in the eye or forehead region, are missing. The standard photometric stereo approach reproduces these facial details very well. It is more stable compared to the approximate Helmholtz stereopsis, as it works even at surface locations where only one camera and a few light sources are visible. Consequently, the two methods that are used in this project are normals from highlight direction and standard photometric stereo.

7.6 Geometry Refinement

The output of the normal estimation stage should ideally be a noise and bias free normal map without any discontinuities. Unfortunately, due to small calibration errors normal estimations based on different cameras tend to be biased, which can lead to an inconsistent normal field. In addition, the geometry from the 3D scanner suffers from imprecisions, namely from a lack of high-frequency details and from high-frequency noise. Nehab et al. [NRDR05] analyze this general problem and propose a method to combine the reliable low-frequency information of a 3D range scan with the high-frequency content of photometric normals, thereby removing the low-frequency bias from the normal orientations.

The method by Nehab et al. is a two-step procedure. In a first stage, high-frequency details of the normal map are combined with the bias-free low-frequency normals of the smoothed geometry. Afterward, geometric detail is added to the smoothed geometry that is consistent with the new normals. We use this method to improve both normal maps and face geometry, adapting this method to our datasets [Bic06]. Figure 7.12 shows a schematic overview of the geometry refinement.

In our setting, the facial normal map has been estimated using 16 different cameras, each having a different systematic bias. Hence, we perform the bias removal on separate normal maps for each camera. We generate low-frequency versions

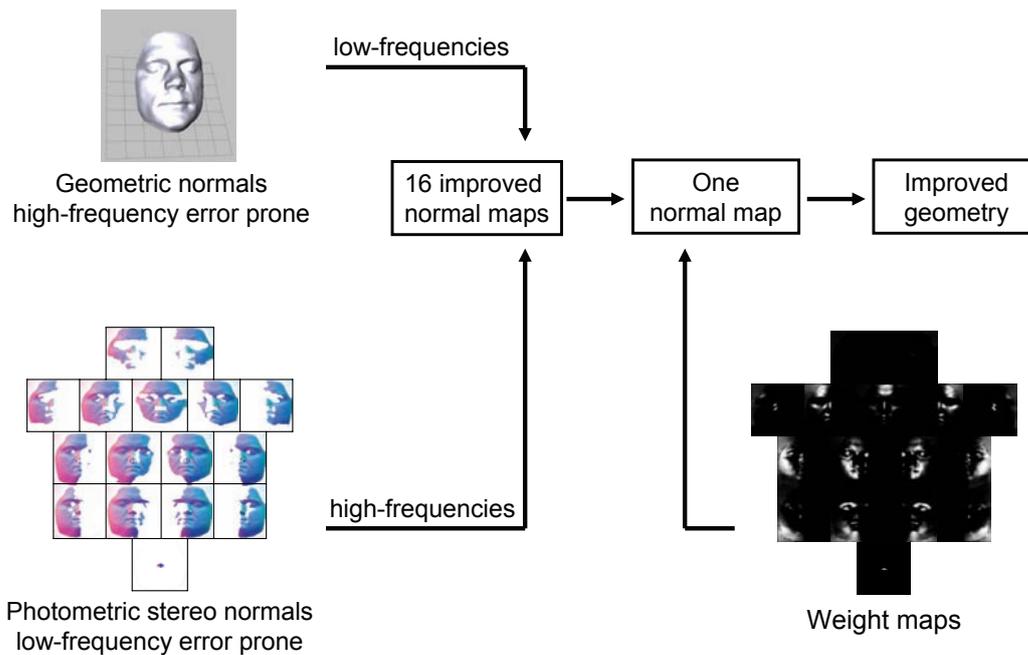


Figure 7.12: *Geometry refinement work flow. Using only low error prone frequency bands of estimated normals and 3D geometry allows us to compute an improved high-resolution geometry. We estimate for each camera (16 in total) a normal map, remove their bias to create 16 improved normal-maps and combine them to a single normal map using an appropriate weighting. Based on the improved normal map an improved geometry is computed.*

of both the geometry and the normal maps by applying Gaussian filters of a carefully chosen size. The high-frequency content of the normal map is determined by comparing the normals of the original normal map with the smoothed version. This yields a rotation field that transforms smoothed normals to original normals. This rotation field represents the unbiased high-frequency content of the original normal map. Applying the respective rotations to the normals of the smoothed mesh produces new normals that contain the low-frequency information from the mesh and details from the photometric normal map. These normals are stored in bias-corrected per-camera normal maps.

Before refining the geometry based on the bias-corrected normals, the new per-camera normal maps have to be blended into a single normal map. The respective fusion of normal maps is performed by weighted averaging. We generate weight maps for each camera, encoding confidence in the respective normal observation. Near-orthogonal views are considered more reliable than grazing observations. Pixel locations with no observation are assigned a zero weight. Due to sudden changes in surface visibility, the resulting weight maps may contain discontinu-

ities where weights abruptly vanish. We smooth out these discontinuities, using the following algorithm.

Algorithm 1: *Smoothing of weight maps*

```

1: zero_pixel_indices  $\leftarrow$  Indices of all pixels where (weight_map == 0)
2: for k = 1 to 200 do
3:   weight_map  $\leftarrow$  Apply gaussian filter to weight_map with kernel size 1
4:   weight_map[zero_pixel_indices]  $\leftarrow$  0
5: end for

```

When blending the normal maps according to the weight maps, we use unstructured lumigraph interpolation (Section 2.5), ensuring that no more than three camera's normal estimations contribute to a final normal. This way, we avoid over-blurring of the blended normals, which can occur due to back-projection errors due to misalignment of the geometric reconstruction.

As a final step, the vertex positions of the smoothed geometry are adjusted to be consistent with the blended normal map. This is done by a linear optimization with two error terms, penalizing deviations from the original vertex locations and normal deviations between the mesh and the normal map, respectively [NRDR05].

Figure 7.13 shows the result of the geometry refinement. Even small facial details such as fine wrinkles around the eye are captured in the geometry, which is a very significant improvement compared to the original geometry. This processing step combines two separate sources of information, the 3QTech 3D geometry scan and the reflectance images, to obtain an improved geometric representation. The high-fidelity geometry in turn leads to more accurate renderings, as presented in Section 9.2. As the lumitexel generation (see Section 7.4) depends on the 3D geometry, we repeat the lumitexel generation with the improved geometry to get more accurate lumitexels for further processing.

7.7 Reflectance Model Fit

The final normal map and the lumitexel stream deal as input for any subsequent skin model construction. Details of our skin model and of the respective model fit are presented in the next chapter. In our processing pipeline, the model fit is embedded into the data flow, yielding final model descriptions that can be rendered within a standard ray tracer, as described in Chapter 9.

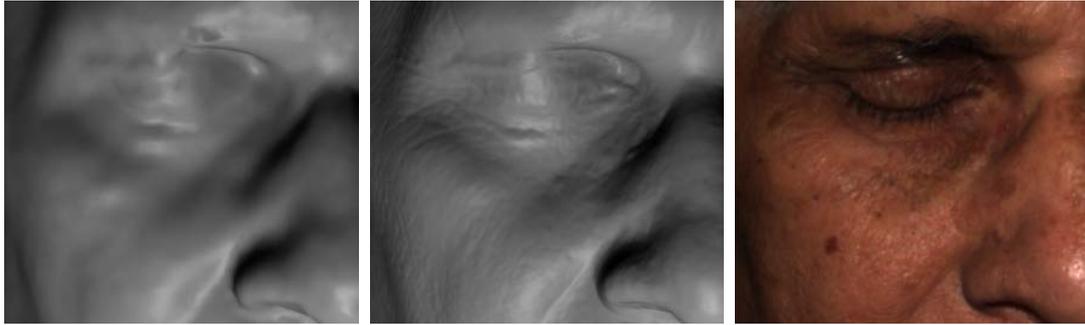


Figure 7.13: Left: 3QTech geometry. Center: Optimized geometry. No normal map was used for the rendering; the surface detail is due to the actual geometry of the model. Right: Original photo.

7.8 Discussion

We presented a data processing pipeline that refines and re-parameterizes the raw input data from the face scanning dome into a form that can directly be used in a model fit. This includes visibility classification, sample generation, and geometry refinement. The pipeline is largely automated to allow for an efficient processing of a large number of scans. The possibility to manually clean and edit the acquired 3D geometry, however, proved to be necessary if a certain model quality has to be reached. This is consistent with our analysis in Chapter 3. Unfortunately, the point-based editing tools presented in that chapter could not be used as-is in the final processing pipeline, as some of the tools employed in this project require a triangle-based representation.

An important stage of the scan processing pipeline is the geometry refinement step. It uses a multimodal approach to combine normal observations from the reflectance field with the acquired 3D geometry to obtain an improved geometric representation. It becomes possible to combine the reliable low-frequency component of the acquired geometry with the high-frequency details of the observed normal field. In Chapter 9 it will be shown that geometry refinement is vital for the fidelity of a face reconstruction.

Skin Reflectance Model

In this chapter we develop a reflectance model that is capable of describing skin reflectance based on the acquired data. Additionally, the representation should contain intuitive parameters that can be manually changed, and it should roughly reflect the physiology of skin. These goals leave a wide range of models that are more or less strictly aligned with physical reality [INN05]. A key decision in our project, however, is that the model should not provide more degrees of freedom than can be defined by the acquired data. A model that contains too many parameters would require ad-hoc definitions of quantities that have not been measured, which falsifies the goal of a measurement-based skin model. On the other hand, physical quantities within the model should be well-exposed to further an elementary understanding of skin reflectance.

8.1 Constituents

We use a simple two-layer model of skin which reflects the coarse physiological classification of skin into epidermis and dermis. On top of the epidermis, we assume a thin oil layer, see Figure 8.1. This classification is simple enough to be accessible to modeling from measured data.

The reflectance model separates reflectance into specular and diffuse reflectance, that is, outgoing radiance is the sum of two independent radiance terms:

$$L(\mathbf{x}_o, \boldsymbol{\omega}_o) = L_{\text{specular}}(\mathbf{x}_o, \boldsymbol{\omega}_o) + L_{\text{diffuse}}(\mathbf{x}_o, \boldsymbol{\omega}_o), \quad (8.1)$$

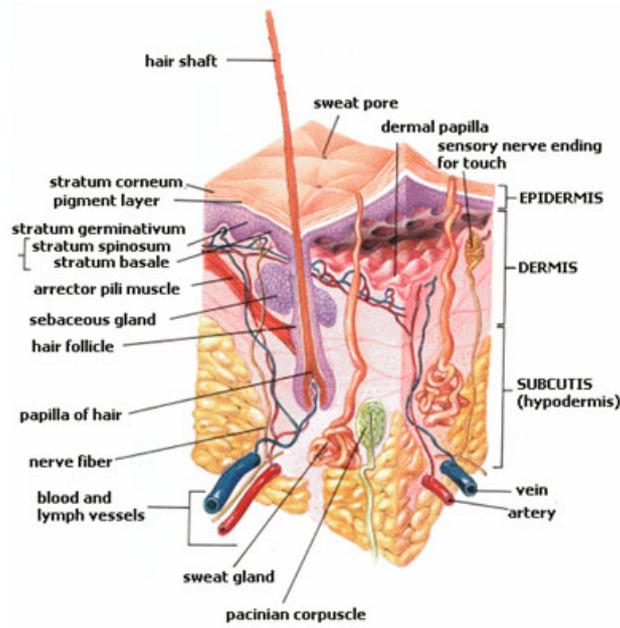


Figure 8.1: *Physiology of skin. On a micro-scale, human skin is a very heterogeneous tissue. However, at scales relevant for rendering, it is sufficient to consider the two prominent layers epidermis and dermis. The visual impact of scattering within the hypodermis is negligible. (Image from Wikipedia.)*

with

$$L_{\text{specular}}(\mathbf{x}_o, \boldsymbol{\omega}_o) = \int_{\Omega} f_s(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i) L(\mathbf{x}_o, \boldsymbol{\omega}_o) \mathbf{n} d\boldsymbol{\omega}_i, \quad (8.2)$$

$$L_{\text{diffuse}}(\mathbf{x}_o, \boldsymbol{\omega}_o) = \int_A \int_{\Omega} S_d(\mathbf{x}_o, \boldsymbol{\omega}_o; \mathbf{x}_i, \boldsymbol{\omega}_i) L(\mathbf{x}_o, \boldsymbol{\omega}_o) \mathbf{n} d\boldsymbol{\omega}_i d\mathbf{x}_i, \quad (8.3)$$

for the ease of formulation, wavelength is omitted in the formulas. They correspond to spatially varying surface reflectance $f_s(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i)$ at the skin-air interface, and to diffuse reflectance $S_d(\mathbf{x}_o, \boldsymbol{\omega}_o; \mathbf{x}_i, \boldsymbol{\omega}_i)$ due to scattering and absorption effects within the skin, see Figure 8.2. Note that subsurface scattering is modeled to take place in the dermis only. Epidermal scattering is neglected, reflecting the mainly absorbing effect of epidermis to the overall appearance. We evaluated different modeling approaches both for surface reflectance and for subsurface scattering.

8.1.1 Surface Reflectance

Surface reflectance is modeled either using the data-driven BRDF approach introduced by [MPBM03], or using common analytical BRDF models, such as the Blinn-Phong [Bli77], the Torrance-Sparrow [TS67], and the Lafortune [LFTG97]

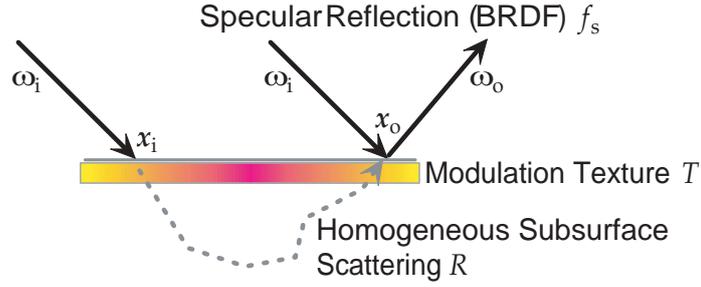


Figure 8.2: Model constituents. Spatially-varying skin reflectance can be explained by a specular BRDF f_s at the oil-layer and a diffuse reflectance component due to subsurface scattering. We model the diffuse reflectance by a subsurface scattering component R_d with an epidermal modulation texture T . Variations in skin albedo are generated either by spatially varying R_d or T .

model. While the data-driven approach allows for an easier model fit and for the application of linear machine learning techniques for dimensionality reduction (see Section 10.1.1), the analytical models are more widely used in Computer Graphics and provide a more compact representation compared to a data-driven approach. Furthermore, the meaning of the model parameters is well-defined for the analytical models, which simplifies the analysis of general skin reflectance. A comparison of the different surface reflectance models is provided in Section 8.2.3.

8.1.2 Diffuse Reflectance

A general diffuse BSSRDF S_d is very difficult to acquire directly. Consequently, the diffuse BSSRDF is often decomposed into a product of lower-dimensional functions. A commonly used decomposition for S_d is:

$$S_d(\mathbf{x}_o, \boldsymbol{\omega}_o; \mathbf{x}_i, \boldsymbol{\omega}_i) = \frac{1}{\pi} f_o(\eta, \boldsymbol{\omega}_o) R_d(\mathbf{x}_o, \mathbf{x}_i) f_i(\eta, \boldsymbol{\omega}_i) \quad (8.4)$$

Here $R_d(\mathbf{x}_o, \mathbf{x}_i)$ describes the diffuse subsurface reflectance of light entering at point \mathbf{x}_i and exiting at point \mathbf{x}_o . The diffuse surface reflectance/transmittance functions $f_i(\mathbf{x}_i, \boldsymbol{\omega}_i)$ and $f_o(\mathbf{x}_o, \boldsymbol{\omega}_o)$ further attenuate the light at the entrance and exit points based on the incoming and outgoing light directions, respectively.

Like previous work [HK93, JMLH01], we assume that skin is mostly smooth due to the oil layer and use transmissive Fresnel functions $F_t(\eta, \boldsymbol{\omega}_i)$ and $F_t(\eta, \boldsymbol{\omega}_o)$ to modulate incoming and outgoing radiance. We set $\eta \approx 1.38$ as the relative index of refraction between facial skin and air for the Fresnel terms [Tuc00]. In addition, the diffuse reflectance is modulated by the epidermal modulation layer T' :

$$S_d(\mathbf{x}_o, \boldsymbol{\omega}_o; \mathbf{x}_i, \boldsymbol{\omega}_i) = \frac{1}{\pi} F_t(\eta, \boldsymbol{\omega}_o) T'(\mathbf{x}_o) R_d(\mathbf{x}_o, \mathbf{x}_i) T'(\mathbf{x}_i) F_t(\eta, \boldsymbol{\omega}_i) \quad (8.5)$$

We assume that modulation of incident light $T'(\mathbf{x}_i)$ can be approximated by a constant average modulation T'_{avg} :

$$\begin{aligned} S_d(\mathbf{x}_o, \boldsymbol{\omega}_o; \mathbf{x}_i, \boldsymbol{\omega}_i) &\approx \frac{1}{\pi} F_t(\eta, \boldsymbol{\omega}_o) T'(\mathbf{x}_o) R_d(\mathbf{x}_o, \mathbf{x}_i) \left(\frac{1}{A} \int_A T'(\mathbf{x}) d\mathbf{x} \right) F_t(\eta, \boldsymbol{\omega}_i) \\ &= \frac{1}{\pi} F_t(\eta, \boldsymbol{\omega}_o) T'(\mathbf{x}_o) R_d(\mathbf{x}_o, \mathbf{x}_i) T'_{\text{avg}} F_t(\eta, \boldsymbol{\omega}_i). \end{aligned} \quad (8.6)$$

This approximation is based on the consideration that the frequency of the diffuse skin texture, and hence of the modulation layer T' , is higher than the low-pass characteristic of the diffusion kernel $R_d(r)$. To further simplify the representation, we model the modulation as a joint *modulation texture* $T(\mathbf{x}_o) = T'(\mathbf{x}_o) T'_{\text{avg}}$:

$$S_d(\mathbf{x}_o, \boldsymbol{\omega}_o; \mathbf{x}_i, \boldsymbol{\omega}_i) = \frac{1}{\pi} F_t(\eta, \boldsymbol{\omega}_o) T(\mathbf{x}_o) R_d(\mathbf{x}_o, \mathbf{x}_i) F_t(\eta, \boldsymbol{\omega}_i). \quad (8.7)$$

A similar modulation texture is used by Fuchs et al. [FGCS05]. Intuitively, the modulation texture represents a spatially-varying absorptive film of zero thickness on top of homogeneously scattering material, affecting only outgoing radiance. To model homogeneous subsurface scattering we use the dipole diffusion approximation by Jensen et al. [JMLH01]. It simplifies $R_d(\mathbf{x}_o, \mathbf{x}_i)$ to a rotationally-symmetric function $R_d(\|\mathbf{x}_o - \mathbf{x}_i\|)$.

In the course of this project, we explored two special cases of (8.7). They differ in the way they model spatial albedo variation. In our first approach, we set $T(\mathbf{x}_o) \equiv 1$, modeling diffuse reflection solely by a spatially varying $R_d(\mathbf{x}_o, \|\mathbf{x}_o - \mathbf{x}_i\|)$. In a second approach, we fix R_d and vary $T(\mathbf{x}_o)$ across the surface. The two approaches have different advantages. A detailed discussion follows in Section 8.2.2.

8.1.3 Micro-Normals

The skin surface shows very detailed normal variations due to pores and follicles. Although the cameras of the acquisition dome are able to capture fine detail of skin albedo, resolution and calibration precision are not sufficient to reconstruct consistent normals at micro-geometry level. Nevertheless, due to the specular oil layer, normal variations at the micro-geometry are still visible in the camera images, being responsible for the “crisp” appearance of pore structures. In a sense, they produce a certain degree of aliasing that makes skin images appear “realistic”. In an attempt to make these effects accessible through our model, we use *micro-normals* [DHT⁺00]. We reconstruct a normal map from each camera view and use view-dependent angular interpolation to obtain a normal map for any virtual viewpoint, see Figure 8.3. This is a good approximation to the aliasing of highlights in real-world photographs, enabling one to render high-fidelity images of the acquired datasets.

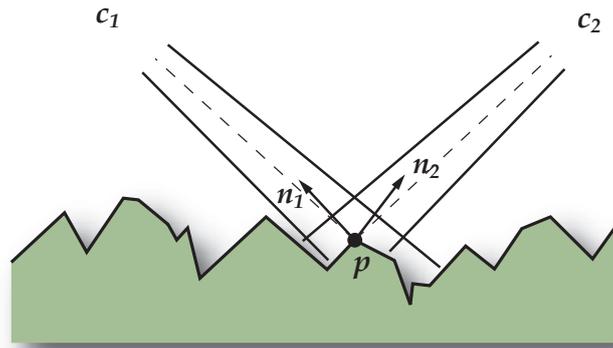


Figure 8.3: *Micro-normals.* The figure shows two cameras c_1 and c_2 that are observing the same surface point p . The two respective camera pixels integrate over different micro-normal distributions, leading to two different normals n_1 and n_2 that are perceived by each camera. Using per-view micro-normal maps which capture a normal map for each camera viewpoint, view-dependent normal variations due to micro-geometry can be approximated.

8.2 Model Fit

All skin model parameters are estimated from measured data, acquired using the surface reflectance field acquisition dome and the subsurface scattering device. Figure 8.4 shows a block diagram of the data flow during parameter estimation. The model is independently fitted to each surface point of the scan. In contrast to previous work [MWL⁺99, LKG⁺01, FLS05], we collect enough samples for a reliable BRDF fit at almost all lumitexels without clustering. Input reflectance samples \mathcal{R} have been transformed into local surface frames and take occlusions into account, see Section 7.4. Subsurface scattering input parameters stem from the fit of the dipole diffusion approximation described in Section 6.3.1. The acquired 3D geometry is not directly part of the parameter estimation.

According to our model, the skin reflectance can be described as a joint effect of specular reflectance f_s and a diffuse BSSRDF S_d , as formalized in (8.1). The dome's reflectance samples \mathcal{R} , however, were derived under the assumption that skin reflectance is due to a pure BRDF f_r . That is, they do not span the full parameter dimensionality of a BSSRDF. To be able to relate samples in \mathcal{R} to S_d , we will approximate S_d by a local BRDF f_d at x_o :

$$\int_A S_d(x_o, \omega_o; x_i, \omega_i) dx_i \approx f_d(x_o, \omega_o, \omega_i). \quad (8.8)$$

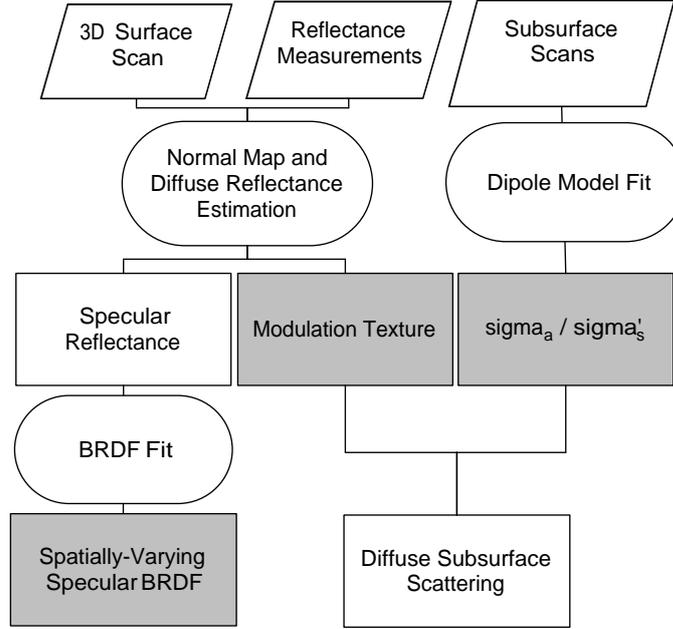


Figure 8.4: Block diagram of the model parameter estimation process. Blocks in gray are the parameters of our model.

Section 8.2.1 describes this approximation in more detail. Using (8.8), total skin reflectance can be approximated as

$$\begin{aligned}
 L(\mathbf{x}_o, \boldsymbol{\omega}_o) &= \int_{\Omega} f_s(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i) L(\mathbf{x}_o, \boldsymbol{\omega}_o) \mathbf{n} d\boldsymbol{\omega}_i \\
 &\quad + \int_A \int_{\Omega} S_d(\mathbf{x}_o, \boldsymbol{\omega}_o; \mathbf{x}_i, \boldsymbol{\omega}_i) L(\mathbf{x}_o, \boldsymbol{\omega}_o) \mathbf{n} d\boldsymbol{\omega}_i d\mathbf{x}_i \\
 &\approx \int_{\Omega} (f_s(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i) + f_d(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i)) L(\mathbf{x}_o, \boldsymbol{\omega}_o) \mathbf{n} d\boldsymbol{\omega}_i \\
 &= \int_{\Omega} f_r(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i) L(\mathbf{x}_o, \boldsymbol{\omega}_o) \mathbf{n} d\boldsymbol{\omega}_i, \tag{8.9}
 \end{aligned}$$

attributing skin reflectance to a *skin BRDF* f_r :

$$f_r(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i) = f_s(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i) + f_d(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i). \tag{8.10}$$

The following sections describe how we employ (8.8) and (8.10) to fit our model to the dome's BRDF measurements while simultaneously constraining S_d to measurements obtained with the subsurface scattering device.

The subsurface scattering measurements are given as pairs (μ'_s, μ_a) at sparse facial locations, while the dome reflectance observations at a dense set of surface points \mathbf{x}_o are given as $R_i = (r^i, \boldsymbol{\omega}_i^i, \boldsymbol{\omega}_o^i) \in \mathcal{R}(\mathbf{x}_o)$, with respect to a surface frame $(\mathbf{n}, \mathbf{t}_1, \mathbf{t}_2)$ and with r^i equal to the cosine-weighted surface reflectance $f_r^i \langle \mathbf{n}, \boldsymbol{\omega}_i^i \rangle$.

8.2.1 Diffuse Albedo Separation

We model diffuse reflectance as a modulated diffusion approximation

$$S_d(\mathbf{x}_o, \boldsymbol{\omega}_o; \mathbf{x}_i, \boldsymbol{\omega}_i) = \frac{1}{\pi} F_t(\eta, \boldsymbol{\omega}_o) T(\mathbf{x}_o) R_d(\|\mathbf{x}_o - \mathbf{x}_i\|) F_t(\eta, \boldsymbol{\omega}_i) . \quad (8.11)$$

In a first step, the model fit estimates parameters for S_d . The BSSRDF measurements of the subsurface scattering acquisition device are not dense enough to fit a spatially varying S_d . Hence, for each surface point, we estimate S_d from the reflectance data acquired with the face scanner dome, merely constraining its parameters to the translucency measurements obtained with the BSSRDF measurement device. The estimation of S_d can be seen as a separation of diffuse albedo from the reflectance field data.

As soon as S_d is known, surface reflectance f_s can be obtained by removing the contribution of S_d from the measured surface BRDF. To this end, we use (8.10) to determine f_s using a BRDF approximation f_d of S_d :

$$f_s(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i) = f_r(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i) - f_d(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i) . \quad (8.12)$$

Analogously to [JMLH01], we derive f_d from the dipole diffusion model as

$$f_d(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i) = \frac{1}{\pi} F_t(\eta, \boldsymbol{\omega}_o) T(\mathbf{x}_o) R_d(\mathbf{x}_o) F_t(\eta, \boldsymbol{\omega}_i) , \quad (8.13)$$

with total diffuse reflectance

$$R_d(\mathbf{x}_o) = 2\pi \int_0^\infty R_d(r) r dr = \frac{\alpha'}{2} \left(1 + e^{-\frac{4}{3}A\sqrt{3(1-\alpha')}} \right) e^{-\sqrt{3(1-\alpha')}} . \quad (8.14)$$

Note that f_d is solely dependent on the three scalars $T(\mathbf{x}_o)$, $R_d(\mathbf{x}_o)$ and η , while η has been fixed. Introducing $\alpha(\mathbf{x}_o) = T(\mathbf{x}_o)R_d(\mathbf{x}_o)$, simplifies (8.13) to:

$$f_d(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i) = \frac{1}{\pi} F_t(\eta, \boldsymbol{\omega}_o) \alpha(\mathbf{x}_o) F_t(\eta, \boldsymbol{\omega}_i) . \quad (8.15)$$

Our estimation of diffuse reflectance f_d assumes that specular reflectance f_s vanishes for at least q percent of the observations f_r^i :

$$\begin{aligned} f_d(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i) &\leq f_d(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i) + \underbrace{f_s(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i)}_{= 0 \text{ for at least } q\% \text{ observations}} = f_r(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i) \end{aligned} \quad (8.16)$$

Given the large directional coverage of the acquisition dome, this is a reasonable assumption for glossy and specular reflectance. Observations where f_s vanishes directly show f_d . Consequently, the estimator aims at maximizing f_d while retaining

$$f_d(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i) \leq f_r(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i) , \quad (8.17)$$

that is,

$$\frac{1}{\pi} F_t(\eta, \omega_o^i) \alpha(x_o) F_t(\eta, \omega_i^i) \leq f_r^i \quad \Leftrightarrow \quad (8.18)$$

$$\alpha(x_o) \leq \underbrace{\left(\frac{\pi f_r^i}{F_t(\eta, \omega_o^i) F_t(\eta, \omega_i^i)} \right)}_{=: \alpha^i}. \quad (8.19)$$

α can be maximized by setting $\alpha(x_o) = \min\{\alpha^i\}$. However, due to measurement noise and calibration errors, we determine a stable minimum of $\{\alpha^i\}$ by choosing

$$\alpha(x_o) = q\text{-percentile of } \{\alpha^i\}. \quad (8.20)$$

This also prevents that shadows from objects that are not captured by the geometry, e.g. hair, eyebrows, eyelashes, or protruding parts of clothes, are misinterpreted as valid measurements. To further increase stability, the percentile is not computed for each color channel separately. Instead, it is based on a sequence of α^i that has been sorted by luminance, that is, by the diffuse reflectance averaged across color channels.

The α values for each sample are projected into a texture map. The projected values are densely interpolated using push-pull interpolation and stored as a *diffuse albedo map*. Inserting α in equation (8.15) delivers diffuse reflectance samples f_d^i . Subtracting f_d^i from the dome observations f_r^i according to (8.12) yields surface reflectance samples f_s^i . They are passed on to the surface reflectance model fit.

Note that for $F_t \equiv 1$, i.e., under a Lambertian reflectance assumption, the proposed method becomes the standard albedo separation commonly used in computer vision.

8.2.2 Diffuse Reflectance

In the previous section the total diffuse reflectance $\alpha(x_o)$ has been determined for each surface point. Fitting the diffuse reflectance model to α means determining $T(x_o)$ and $R_d(x_o)$, while constraining R_d 's parameters by the subsurface scattering measurements. We followed two modeling approaches that fix either T or R_d to a constant value, respectively.

Spatially Varying Material Properties

In the first approach, we assume facial albedo variation entirely to be caused by spatially varying scattering and absorption properties μ'_s and μ_a . We set the modulation texture constantly to $T \equiv 1$, only deriving parameters for R_d .

The approach exploits a property of the actual implementation of the dipole approximation shader that we use in our rendering environment. We use a hierarchical BSSRDF implementation after [JB02]. In contrast to implementations after [JMLH01], it re-uses irradiance samples for a per-surface point evaluation of the dipole model. As a consequence, the shader is capable of displaying materials of spatially varying scattering and absorption parameters, and each surface point x_o is shaded as if it were part of a homogeneous material with the subsurface scattering parameters at x_o . That is, for homogeneous illumination the rendered color is faithfully described by the BRDF approximation $R_d(x_o)$ of the diffuse BSSRDF at x_o .

This allows to determine parameters for R_d that exactly reproduce the input data when rendered with this shader. Recalling that an $\alpha(x_o)$ has been reconstructed that suffices

$$\alpha(x_o) = T(x_o)R_d(x_o), \quad (8.21)$$

we know that $R_d(x_o) = \alpha(x_o)/T(x_o) = \alpha(x_o)$. From spatially BSSRDF measurements across a subject's face, see Section 6.4.1, we learned that variation of skin translucency is very low. Hence, we use the effective transport coefficient μ_{tr} derived from measurements of the subject to fix a constant translucency of $1/\mu_{tr}$ for the whole dataset. Jensen and Buehler [JB02] show that from R_d and μ_{tr} , reduced scattering and absorption coefficients, μ'_s and μ_a , can be derived. Following their line of computation, we solve (8.14) for the reduced albedo α' using a numerical secant root finder. With $\mu'_t = \mu_{tr}/\sqrt{3(1-\alpha')}$, we obtain

$$\begin{aligned} \mu'_s &= \alpha' \mu'_t & \text{and} \\ \mu_a &= \mu'_t - \mu'_s. \end{aligned} \quad (8.22)$$

In order to obtain dense coefficient textures for μ'_s and μ_a , we project the values $1/\mu_{tr}$ and R_d for each x_o into high-resolution textures and apply push-pull interpolation. For these quantities, linear interpolation kernels are still meaningful. The non-linear transformation to parameters μ'_s and μ_a is then performed on a per-texel level at the final output resolution.

In conjunction with the hierarchical BSSRDF implementation, this method reproduces the input data exactly, maintaining maximum realism even under spatially varying illumination, for instance at shadow boundaries. As a general method, however, it is less suited if a BSSRDF shader with different properties is used. Shader implementations after [HK93] or [JMLH01], e.g., would over-blur the albedo texture in renderings.

Epidermal Absorption Layer

In order to obtain a model with less dependence from the rendering architecture, we alternatively vary T , while keeping R_d fixed. This model can also easily be

justified by the observation that dermal diffusion is indeed largely homogeneous, while most albedo variations are due to absorption at the interface between dermis and epidermis. Furthermore, scattering effects within the epidermis are negligible compared to dermal scattering. T accordingly stands for an epidermal absorption layer on top of a dermal diffusion term.

As $R_d(r)$ is assumed to be homogeneous across the surface, it can again be approximated by its BRDF approximation $R_d(\mathbf{x}_o)$ during the fit. As in the previous approach, we constrain μ_{tr} to a value measured with the subsurface scattering acquisition device. As a spatially constant value $R_d(\mathbf{x}_o)$, we choose

$$R_d(\mathbf{x}_o) \equiv \alpha_{\text{avg}} = \frac{1}{A} \int_A \alpha(\mathbf{x}) \, d\mathbf{x} , \quad (8.23)$$

from which we derive subsurface scattering parameters according to (8.22). $T(\mathbf{x}_o)$ is then set to values that explain the $\alpha(\mathbf{x}_o)$ derived from the dome observations. From

$$\alpha(\mathbf{x}_o) = T(\mathbf{x}_o)R_d(\mathbf{x}_o) = T(\mathbf{x}_o)\alpha_{\text{avg}} , \quad (8.24)$$

it follows that

$$T(\mathbf{x}_o) = \frac{\alpha(\mathbf{x}_o)}{\alpha_{\text{avg}}} . \quad (8.25)$$

Accordingly, a densely interpolated $T(\mathbf{x}_o)$ is computed directly from the albedo map.

The approach has a clear advantage. Under constant illumination, it reproduces the input data exactly, independent from implementation details of the dipole approximation shader. On the downside, an error is introduced under spatially varying illumination, e.g. at shadow boundaries. In these areas, modulation of incident light is not negligible anymore. However, in our experience the visible effect is rather low. A rather theoretical disadvantage is the fact that the modulation layer contains values greater than one, i.e., it is not purely absorptive but also amplifies radiant exitance. Although strictly speaking, this leaves the bounds of physical reality, the combined model (8.7) will always stay within the physical bounds of a maximum total reflectance of one.

8.2.3 Surface Reflectance

The specular reflection is computed at each lumitexel by subtracting the BRDF approximation of the total diffuse reflectance:

$$f_s(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i) = f_r(\mathbf{x}_o, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i) - \frac{1}{\pi} F_t(\mathbf{x}_o, \boldsymbol{\omega}_o) \alpha(\mathbf{x}_o) F_t(\mathbf{x}_i, \boldsymbol{\omega}_i) . \quad (8.26)$$

Highlights on dielectric materials like skin are of the same color as the light source (white, in our case). Consequently, we convert the specular reflectance to gray

scale to increase the stability of the BRDF fit. Unlike previous work [MWL⁺99, LKG⁺01, FLS05] we do not need to cluster reflectance data because our acquisition system collects enough samples for a spatially-varying BRDF fit at each vertex. The data for lumitexels with a badly conditioned fit is interpolated during creation of the BRDF coefficient texture. Measurements close to extreme grazing angles are generally unreliable due to imprecisions in the registration between 3D geometry and camera images. Hence, reflectance samples with more than 80° deviation from the normal are not considered for the fit.

Data-Driven BRDF

In the beginning, we were using the data-driven BRDF model [MPBM03] to describe interface reflectance. The model is build from tabulated real-world BRDFs that are superimposed to produce individual BRDFs. We are using 59 basis BRDFs from the database by Matusik et al. [MPBM03]. These basis BRDFs are not specific to human skin—they are capable of representing reflectance of a much wider range of materials. We converted the original basis to gray-scale and removed a small number of overly specular BRDFs from the database, as they would have ill-conditioned the fit. In advance to the model fit, we pre-convolve each basis BRDF by the angular footprint of the light sources, thereby compensating for the approximation in Equation (7.5).

When fitting the data-driven model, a solution of non-negative BRDF weights is desired. Allowing for negative weights potentially leads to over-fitting and visible artifacts, such as “negative highlights” [MPBM03]. We use a non-negative least squares optimization based on quadratic programming. Our error metric is defined in image space, that is, on reflectance samples that are not yet cosine-normalized. At each surface point \mathbf{p}_j we obtain a surface BRDF

$$f_s^{\text{DD}}(\mathbf{p}_j, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i) = \sum_{i=1}^n w_{ij} f_i(\boldsymbol{\omega}_o, \boldsymbol{\omega}_i) \approx f_s(\mathbf{p}_j, \boldsymbol{\omega}_o, \boldsymbol{\omega}_i), \quad (8.27)$$

with all weights $w_{ij} \geq 0$. The spatially varying weights can be written as a per-surface-point weight vector $\mathbf{w}_j = (w_1 \dots w_n)^\top$. We use dimensionality reduction to discover a smaller linear basis that is specific to an acquired dataset. We construct a weight matrix $W = (\mathbf{w}_{k_1} | \dots | \mathbf{w}_{k_m})$ and apply non-negative matrix factorization (NMF) to it. W contains only weight vectors for a random subset $\mathbf{p}_{\{k_i\}}$ of surface points, as NMF is generally not capable of dealing with very large matrices.

Unlike principal component analysis (PCA), the NMF basis is not orthogonal. However, the original data can be expressed as *positive* linear combinations of the NMF basis vectors. That is, when fitting reflectance data to the reduced NMF basis, non-negative least squares optimization still prohibits negative coefficients

in the original basis. Using PCA we would have to allow for negative coefficients, leading again to artifacts.

To determine the smallest NMF basis that is necessary to model the data well, we plot the reconstruction error as a function of the number of NMF basis BRDFs (see Figure 10.4). There is a considerable fall-off in the sequential values seen in this plot. The plateau for a single person is reached around four basis BRDFs when the relative reconstruction error does not change much anymore. Figure 8.5 shows synthesized images of a face using different numbers of NMF basis BRDFs.

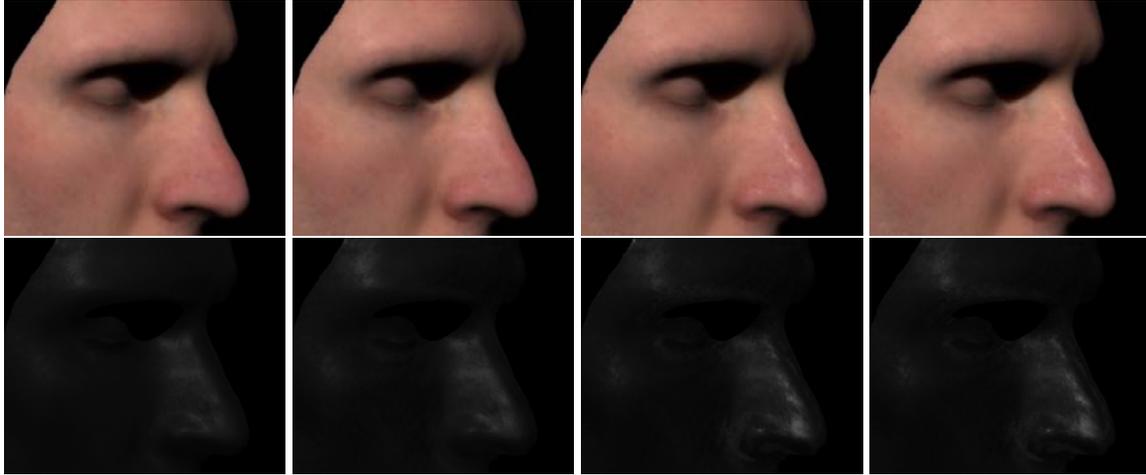


Figure 8.5: Synthetic images of a face using, from left to right, 1, 2, 4, and 24 NMF basis BRDFs. Top: Rendered images. Bottom: Surface reflection only. Four NMF basis BRDFs are sufficient to represent the surface reflectance of a single person.

As expected, the results do not improve much beyond four basis BRDFs.

The coefficients for each surface point are stored in four BRDF coefficient textures. In addition, we store the respective NMF basis BRDFs per face. We analyze a more general BRDF basis for a large population of people in Section 10.1.1.

Analytic Models

We also fit the Blinn-Phong [Bli77] and Torrance-Sparrow [TS67] isotropic BRDF models to the data. The Blinn-Phong model is the most widely used analytic reflectance model due to its computational simplicity:

$$f_s^{\text{BP}} = \rho_s \frac{n+2}{2\pi} \cos^n \delta. \quad (8.28)$$

Here δ is the vector between the normal \mathbf{n} and the half-way vector \mathbf{h} , ρ_s is the scaling coefficient, and n is the specular exponent. The factor $(n+2)/2\pi$ is for

energy normalization so that the cosine lobe always integrates to one. This assures that 100% of the incident energy is reflected for $\rho_s = 1$.

We also use the physically-based Torrance-Sparrow model:

$$f_s^{\text{TS}} = \rho_s \frac{1}{\pi} F_r(\langle \boldsymbol{\omega}_o, \mathbf{h} \rangle) \frac{DG}{\langle \mathbf{n}, \boldsymbol{\omega}_i \rangle \langle \mathbf{n}, \boldsymbol{\omega}_o \rangle}, \quad (8.29)$$

with:

$$G = \min\left\{1, \frac{2\langle \mathbf{n}, \mathbf{h} \rangle \langle \mathbf{n}, \boldsymbol{\omega}_o \rangle}{\langle \boldsymbol{\omega}_o, \mathbf{h} \rangle}, \frac{2\langle \mathbf{n}, \mathbf{h} \rangle \langle \mathbf{n}, \boldsymbol{\omega}_i \rangle}{\langle \boldsymbol{\omega}_o, \mathbf{h} \rangle}\right\},$$

$$D = \frac{1}{m^2 \cos^4 \delta} e^{-[(\tan \delta)/m]^2}. \quad (8.30)$$

G is the geometry term, D is the Beckmann micro-facet distribution, and F_r is the reflective Fresnel term. The free variables are the scaling coefficient ρ_s and the roughness parameter m .

For the fit we use the fitting procedure by Ngan et al. [NDM05]. The objective function of the optimization is the mean squared error between the measured BRDF f_s and the target model f_s^{M} (either f_s^{BP} or f_s^{TS}) with coefficient vector \mathbf{c} :

$$E(\mathbf{c}) = \sqrt{\frac{\sum w (f_s(\boldsymbol{\omega}_i, \boldsymbol{\omega}_o) \cos \vartheta_i - f_s^{\text{M}}(\boldsymbol{\omega}_i, \boldsymbol{\omega}_o; \mathbf{p}) \cos \vartheta_i)^2}{\sum w}}. \quad (8.31)$$

The sum is over the non-empty samples of the lumitexel, and ϑ_i is the elevation angle of incident direction. The weight w is a correction term that allows us to ignore data with $\boldsymbol{\omega}_i$ or $\boldsymbol{\omega}_o$ greater than 80° . We apply constrained non-linear optimization based on sequential quadratic programming over the specular lobe parameters to minimize the error metric.

To assess the quality of our fit, we compare renderings from camera viewpoints to the actual photographs. The camera and light source calibration data were used to produce identical conditions in the renderings. Figure 8.6 shows a comparison of the original and synthesized images with the same illumination and different BRDF models. The spatially-varying Blinn-Phong model (c) overestimates the specular component and does not match the photograph as closely as the spatially-varying Torrance-Sparrow model (b). Nevertheless, it produces more accurate results compared to the spatially uniform Torrance-Sparrow BRDF (d).

Another popular choice for BRDF fitting is the Lafortune model [LFTG97] because it models off-specular reflections while still being computationally simple. Because the non-linear Lafortune model has more parameters than the Torrance-Sparrow and Blinn-Phong models, fitting of the model becomes more demanding, even for a single-lobe approximation. A realistic employment of the Lafortune model would use at least three lobes. This fit has been reported by [NDM05] to be

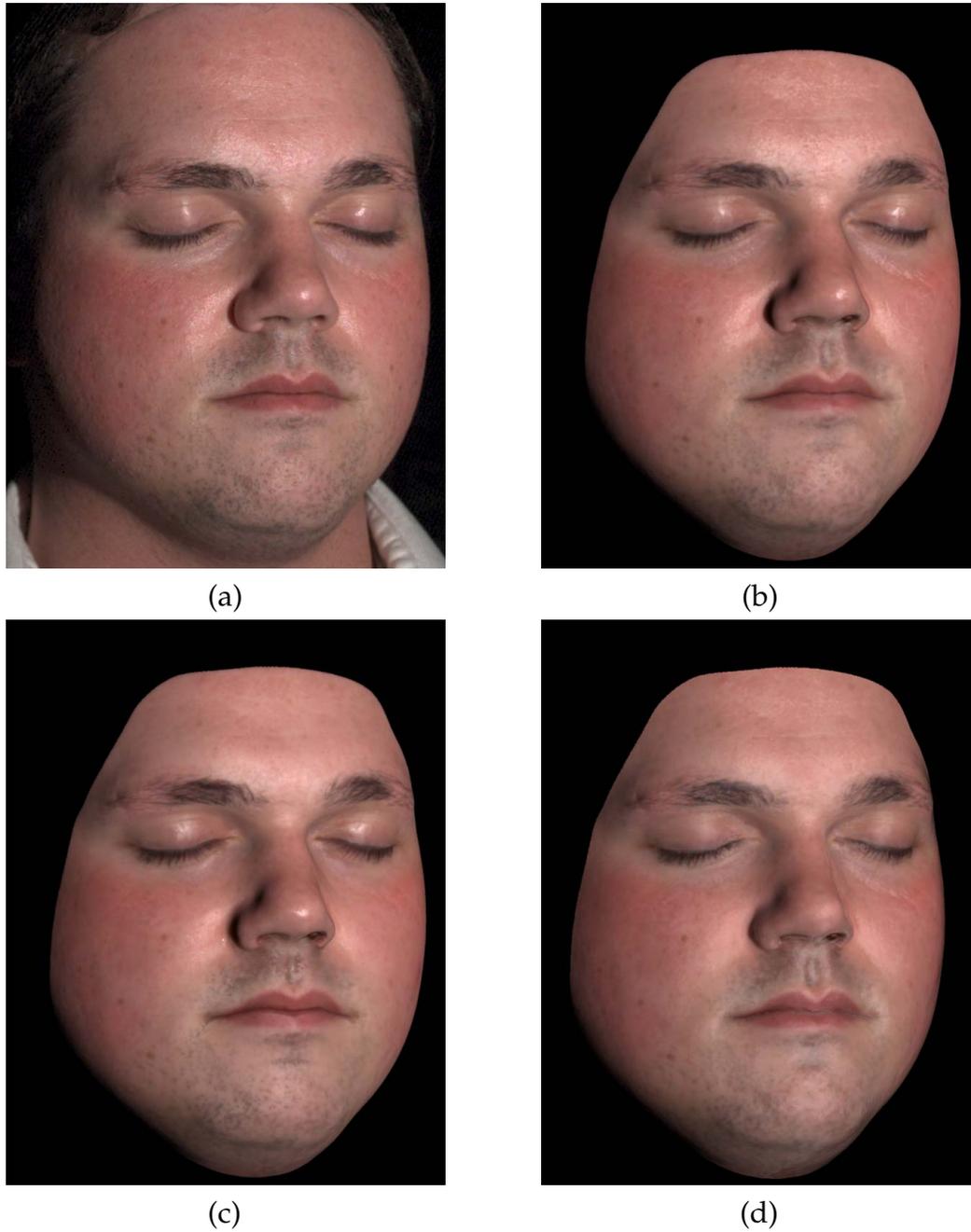


Figure 8.6: *BRDF Models. (a) Input photograph, compared to our face model with (b) spatially-varying Torrance-Sparrow; (c) spatially-varying Blinn-Phong; and (d) uniform Torrance-Sparrow BRDF models.*

very unstable compared to other reflectance models. In addition, [LRR04] explain that “Often, a nonlinear optimizer has difficulty fitting more than two lobes of a Lafortune model without careful user intervention.” Given the high quality that we could achieve with the Torrance-Sparrow model, we decided not to further investigate the Lafortune model.

8.2.4 Micro-Normals

Micro-normals are being determined for each camera using the half-way vector method described in 7.5.1. The method has been chosen due to its capability of capturing fine details. As the micro-normals are only used for visualization, the non-integrability of the resulting normal field is not critical.

8.3 Discussion

In this chapter we proposed a practical modeling approach to describe skin reflectance. The reflectance is generally separated into two components: surface reflectance and diffuse reflectance due to subsurface scattering effects. We propose alternative reflectance models for each component. The final choice of model will be application dependent and is also influenced by the rendering architecture employed. The data-driven surface BRDF model allows for linear analysis and editing operations on BRDF coefficients, while the Torrance-Sparrow model as an industry standard has a wider support on existing rendering platforms. Spatially varying subsurface scattering coefficients may be closer to physical reality, while a constant diffusive material with a modulation texture on top provides more independence from the BSSDF shader implementation.

We further presented a method to fit our model to measured data. It relies on the separation of diffuse and specular effects as a core mechanism. The fitting procedure runs unsupervised, which is important as the model has to be independently fitted to up to 800 000 lumitexels for a single face. We observe that spatial variance of surface reflectance is a necessary prerequisite for faithful appearance reconstruction. As shown in Figure 8.6, it would not be sufficient to use a constant surface reflectance function for the whole face.

In the remainder of this thesis, Chapter 9 demonstrates the performance of two instances of our model when fitted to measured data. Chapter 10 performs an analysis of model parameter variance between multiple subjects.

Rendering

The skin reflectance model was implemented using the off-line ray tracer Dali by Henrik Wann Jensen. This high-quality Monte Carlo renderer is extensible using a C++ interface for custom shaders. Although rendering speed and quality profited from the choice of renderer, the face model is principally not limited to a particular rendering architecture. This chapter shortly summarizes the implementation of the model within Dali, presents rendered images of reconstructed face models that were obtained from the acquired data, and discusses the quality of the results.

9.1 Implementation

Face models are defined as uv -parameterized triangle meshes of 500 000 to 800 000 vertices. A texture stack defines model parameters, such as normal directions, scattering coefficients, albedo values, etc., with a texture resolution between 1024×1024 and 2048×2048 texels. Custom shaders have been written to support the different surface reflectance models, and to implement the modulation texture $T(x_o)$. A hierarchical implementation of the dipole diffusion model is already part of Dali. As a Monte Carlo ray tracer, the renderer benefits from the definition of importance functions to support an improved quality-speed trade-off by importance sampling of surface BRDFs. By averaging across multiple data-driven skin BRDFs, we obtain a tabulated average skin BRDF that is used as importance function for all employed surface reflectance models, including the analytic mod-

els. This is a natural choice, as efficient importance sampling of general BRDFs would be implemented using BRDF tabulation anyway. Note that this choice of importance function does not alter the rendered BRDFs. It has only an effect on the signal-noise-ratio in the resulting images.

An additional set of custom shaders is used to simulate lighting and camera calibrations of the acquisition dome, including the beam-spread of the light source. This allows to faithfully reproduce input reflectance images within the renderer. The exact reproduction of the acquisition conditions allows for side-by-side comparison between real photographs and artificial renderings. Such comparisons are presented in the next section to evaluate the different modeling approaches.

9.2 Results

In the course of the project, multiple variants of the skin model have been explored. We provide exemplary results achieved using the two main variants of our model. Confer Section 8.2 for more detail on the model variants.

9.2.1 Data-Driven Surface BRDF

Data-driven BRDFs and the Torrance-Sparrow model have been evaluated concurrently as a surface reflectance model. In this section, we describe our experiences with the data-driven BRDF model. Compared to analytic reflectance models that typically have a non-linear parameter relationship, the linearity of the data-driven BRDF model makes it better suited for data fitting and trait analysis (see Section 10.1). In order to base the model on physical material properties, we further decided to model albedo variations as a spatially varying diffusion process R_d .

Figure 9.1 shows results that were achieved with this model. It compares single input reflectance images with synthetic images for different faces and different viewpoints. The camera and light source calibrations of the dome were used to reproduce identical conditions. Figure 9.2 shows another example, including renderings of the different components in our model. We observe that the reconstructions come very close to the input photographs. Remaining differences are mainly due to geometric factors: the presented results do not yet use the geometry refinement after [NRDR05], which leads to a significant texture blur due to back-projection errors, and they use a single normal map instead of view-dependent micro-normals. The lack of micro-normals leads to mis-matches in the high-light structure, as can be seen, e.g., at the nose of dataset in Figure 9.2.



Figure 9.1: Comparison of real photographs of four subjects (top) with fits of our model (bottom).

9.2.2 Torrance-Sparrow & Improved Geometry

A BRDF analysis across a large number of subjects revealed that the subject-dependent variation of skin surface BRDF is rather negligible (see Section 10.1.1). With this insight, there remains less reason for a linear reflectance model such as the data-driven BRDF. As a consequence, we subsequently concentrated our investigation on the analytic Torrance-Sparrow model that proved earlier to be very well suited to model the typical gloss of skin. By carefully improving the non-linear model fit according to [NDM05], the fit became reliable enough to be used in an unsupervised framework [Bic06]. Another advantage that motivated this choice is that the Torrance-Sparrow model can be accounted as an industry standard, making the results of our analysis applicable within a large number of existing renderers. To further the independence of our model from the rendering architecture, cf. Section 8.2.2, we turned to the simpler model of a constant diffusion process with a modulation texture $T(x_o)$. Last but not least, geometry refinement and micro-normal maps have been included into the processing pipeline. These geometrical improvements in particular improved the appearance of the reconstructed datasets dramatically.

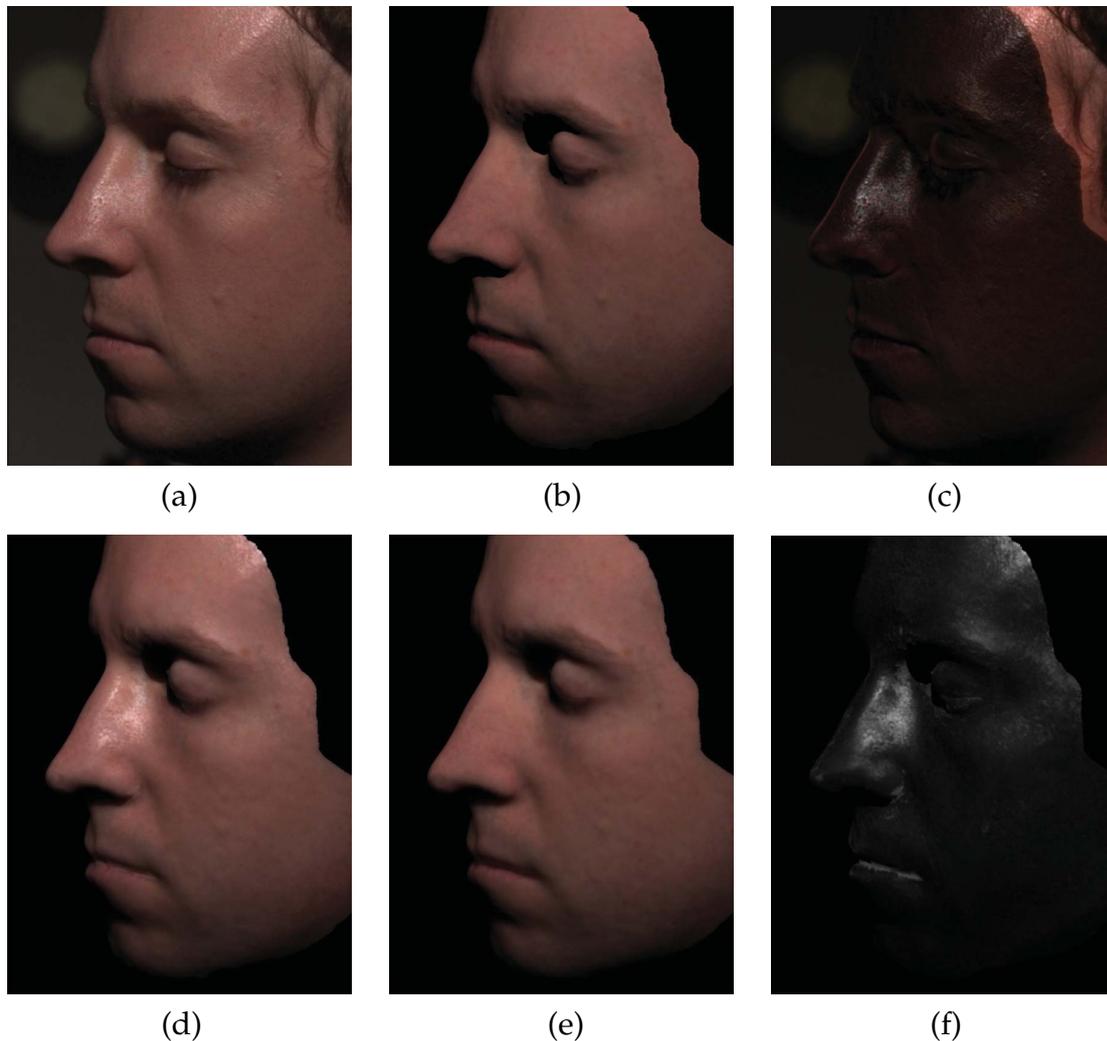


Figure 9.2: Components of our model using a data-driven BRDF. (a) One of the input images with single light source illumination. (b) 3D surface scan shaded with the BRDF approximation of the diffuse subsurface term. The model is lit using the camera and light source calibration derived for view (a). (c) Subtracting the estimated diffuse reflection term (b) from (a) reveals surface reflection as used for the BRDF fit. (d) Reconstructed model. (e) Underlying dipole diffusion. Note the smoother shading and soft shadows in comparison to (b), while preserving texture detail. (f) Reconstructed surface reflection.

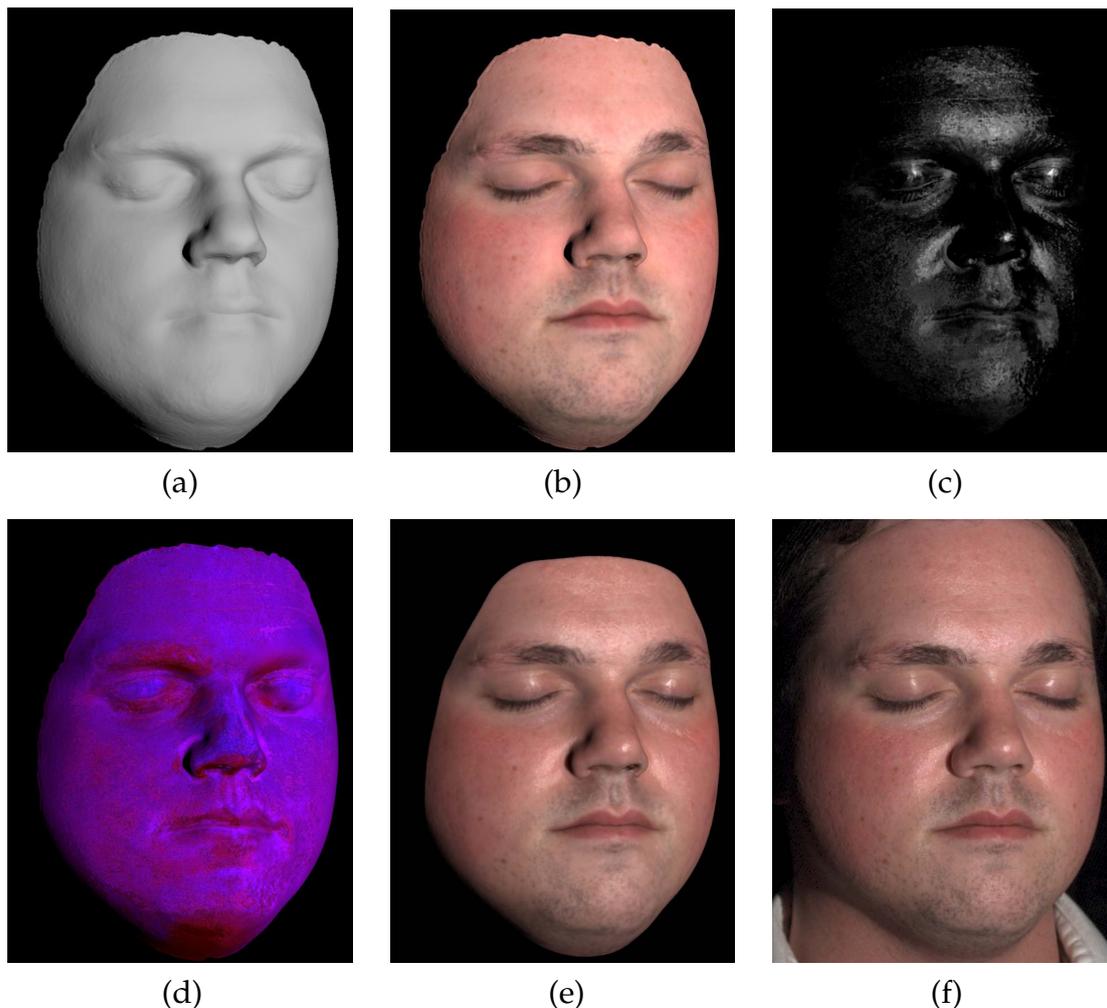


Figure 9.3: Components of the model using the modulation texture and a Torrance-Sparrow BRDF. (a) 3D geometry. (b) 3D geometry and albedo texture; no diffusion process is shown. (c) Specular reflectance. (d) Color encoded Torrance-Sparrow coefficients m and ρ_s . (e) Rendering based on the complete model. (f) Original input image.

Figure 9.3 shows a respective model fit and illustrates the constituents of the model. The reconstruction preserves significantly more detail of the original image than the previous models without geometry refinement. Only highest-frequency details such as the beard still look slightly blurred. This effect is already present in the diffuse albedo. Stubbles are not explicitly modeled, as they are beyond the geometrical resolution of our setup, which naturally leads to blur due to back-projection errors. Image (d) in Figure 9.3 shows the Torrance-Sparrow BRDF parameters color encoded, revealing the spatially-varying nature. The values appear to be clustered into regions of similar BRDF. For instance, obvious differences between the nose and cheek or chin region can be observed. See Section 10.1.1 for a more extensive analysis of facial surface reflectance.

In order to compare the two different modeling approaches for diffuse reflectance, Figure 9.4 shows the same face model rendered using fixed R_d with spatially varying T , and varying R_d with $T \equiv 1$, respectively. The renderings are nearly iden-

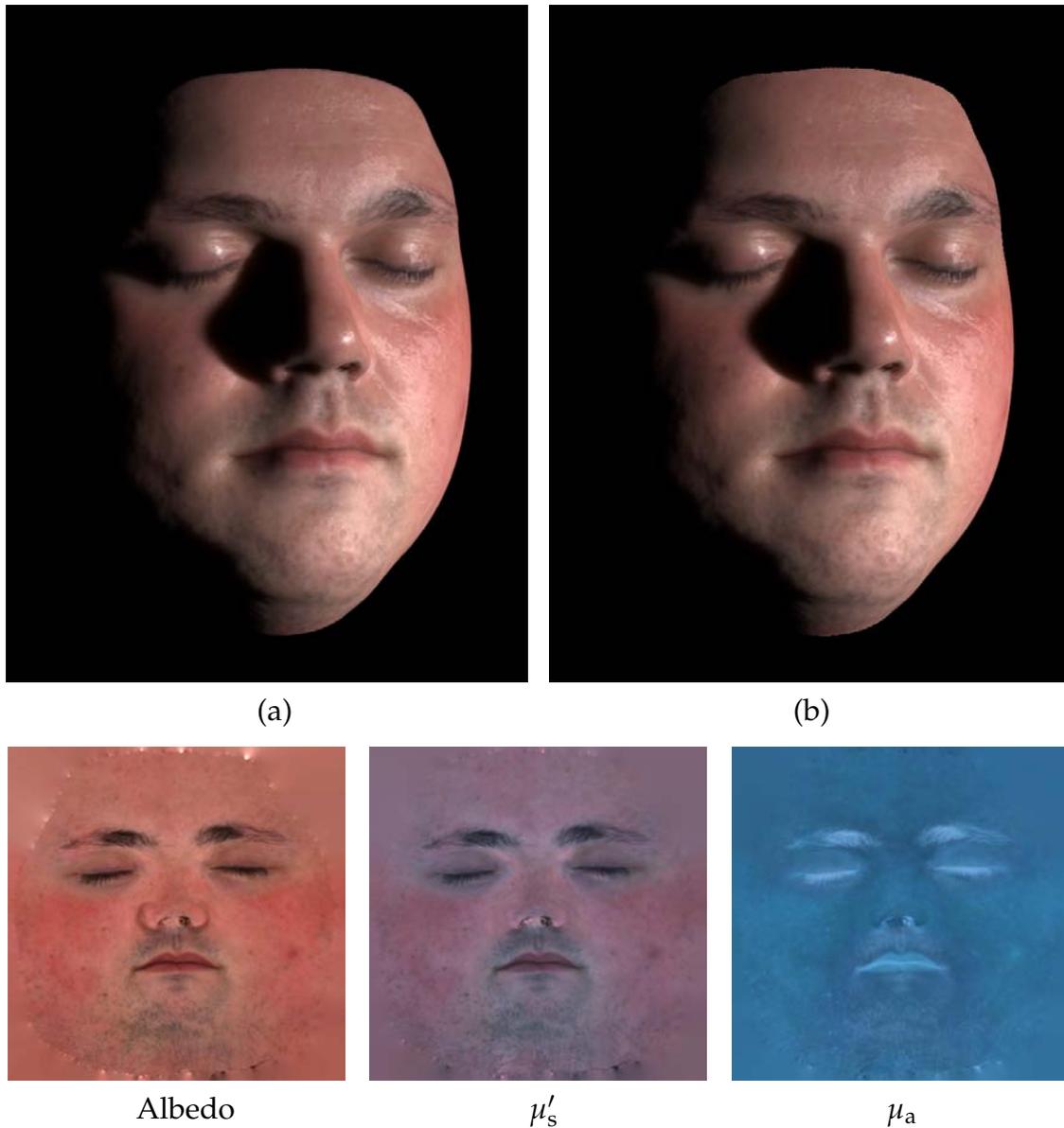


Figure 9.4: Two approaches to model diffuse reflectance. (a) A constant average diffusion term is modulated by an albedo texture. (b) Spatially varying material properties μ'_s and μ_a control the diffusion process. The bottom row shows the respective coefficient textures.

tical. The texture variation corresponds exactly. Only at shadow boundaries, the spatially varying R_d shows a slightly smoother transition. We conclude that the increase in fidelity from the modeling approach in Figure 9.2 to the one in Figure 9.3 is solely due to surface reflectance, which in particular reflects the added detail of the improved geometry and the micro-normals.

Figure 9.5 shows further comparisons between real and synthetic images for different faces and different viewpoints using Torrance-Sparrow reflectance and improved geometry. Each image took approximately five minutes to render on a single Intel Xeon 3.06 GHz workstation.

9.3 Discussion

The synthetic and real images look very similar, but not absolutely identical. The main limitation is a lack of sharpness, both in the texture and in geometry, mainly because accuracy in geometric calibration and alignment remain issues. The intensity and shape of the specular highlights in the synthetic images is sometimes underestimated. The shape of the specular highlights—especially on the forehead—is greatly affected by fine wrinkles and small pores. Several sources of error (measured 3D geometry, motion artifacts, calibration errors, camera noise, Bayer interpolation, errors in the photometric stereo estimations, etc.) prevent us from capturing the micro-normal geometry with 100% accuracy. Much higher-resolution and higher-quality face scans can be obtained by creating a plaster mold of the face and scanning it with a high-precision laser system. For example, the XYZRGB 3D scanning process originally developed by the Canadian National Research Center contributed significantly to the realism of the Disney project [Wil05] and the *Matrix* sequels. However, it would be prohibitive to scan more than one hundred subjects this way. It would also be difficult to correspond the 3D geometry with the image data from the reflectance field dome due to registration errors, subtle changes in expressions, or motion artifacts. In addition, the molding compound may lead to sagging of facial features [Wil05]. We believe that our system strikes a good tradeoff between speed, convenience, and high image quality.

Facial hair is currently not explicitly represented in our model. Although the overall reflectance and geometry of the bearded person in Figure 9.5 has been captured, the beard is lacking some details that are visible in the photograph. We avoided scanning people with full beards or mustaches. Velvety hair is not modeled either, which affects the appearance of the facial silhouette, in particular under back-lighting. However, we believe that in principle both beard and velvety hair can be added to our model, as they contribute complementary effects.

We model subsurface scattering using a single layer. However, skin consists of several layers that have different scattering and absorption properties [Tuc00]. The multi-layer dipole model of Donner et al. [DJ05] would provide a better approximation, but as previously discussed, it is not clear how to measure its parameters *in vivo* for a large number of people. Nevertheless, the presented results show that a single scattering layer suffices to reproduce a high visual fidelity.



Figure 9.5: Comparison of real photographs (first and third row) to our model (second and last row). All photographs were cropped according to the 3D model to remove distracting features. See Color Plates on pp. 184–188 for larger versions of these images.

Generally, it is worth to further investigate in subsurface scattering models. The dipole diffusion approximation overestimates the total fluence in the superficial areas of skin, making the skin appear slightly more opaque than actual skin [INN05]. Another limitation of our model is that our measurements are confined to RGB data. Krishnaswamy and Baranoski [KB04b] developed a multi-spectral skin model and emphasize the importance of spectral rendering in the context of skin.

Appearance Analysis

The ultimate goal of this project has been an editable skin reflectance model that contains meaningful parameters for face editing. The set of parameters should allow to intuitively change the skin appearance of any acquired face. To this end, we measured face reflectance for 149 subjects, followed by an analysis of translucency, diffuse albedo maps, and spatially-varying specular BRDF parameters of the respective surface BRDF models. The statistical results of our analysis can then be used to define meaningful appearance editing operations.

10.1 Analysis

Each subject had to fill in a questionnaire, allowing to classify all acquired data by the following criteria.

Skin type. The skin type is classified according to the Fitzpatrick system [Fit88]. Since we only have two individuals of skin type I in our database we group types I and II together for further analysis. Table 10.1 explains the Fitzpatrick system and shows the distribution of our measurements.

Age group. The distribution by age is listed in Table 10.2.

Skin Type	Skin Color	Sun Exposure Reaction	Subjects (M/F)
I	Very white	Always burn	–
II	White	Usually burn	8 / 6
III	White to olive	Sometimes burn	49 / 18
IV	Brown	Rarely burn	40 / 8
V	Dark brown	Very rarely burn	13 / 2
VI	Black	Never burn	4 / 1

Table 10.1: The Fitzpatrick skin type system and the number of subjects per skin type.

Age Range	Subjects (M/F)
< 30	61 / 18
30–39	37 / 10
40–49	10 / 2
≥ 50	6 / 5

Table 10.2: Age distribution.

Gender. Our database is significantly skewed toward a male majority. In addition, approximately half of the female datasets is affected by make-up, which has to be taken into account when interpreting skin appearance.

Facial region. We distinguish 10 different face regions (mustache, forehead, eyebrows, eyelids, eyelashes, nose, cheek, lips, chin and beard region) as shown in Figure 10.1. This is especially interesting for investigating the specular reflectance, which is directly related to the BRDF parameters. To classify the regions, a mask in uv space is manually drawn for each subject, encoding the regions with certain color codes.

Additional questions in the questionnaire were related to the subject’s smoking habits and to frequent sun or weather exposure. This information might be useful for future analysis, but has not been evaluated yet. Each capture session takes about 30 minutes and includes multiple acquisitions with the reflectance field dome and six subsurface scattering measurements.

Figure 10.2 shows a few representative subjects in the face database that we acquired. The study has been IRB approved by the MIT Committee on the Use of Humans as Experimental Subjects (COUHES), and written informed consent was obtained from all subjects prior to the measurement. All of the study personnel had to take and pass a MIT COUHES training course on human subjects research.



Figure 10.1: 10 face regions: (1) mustache, (2) forehead, (3) eyebrows, (4) eyelids, (5) eyelashes, (6) nose, (7) cheek, (8) lips, (9) chin, and (10) beard region.



Figure 10.2: *Representative subjects in dataset B.*

10.1.1 Interface Reflectance

The first question of interest has been whether there are subject-dependent variations in pure surface reflectance and how this reflectance varies under the influence of external conditions.

Data-Driven BRDF

In order to allow for a wider range of analysis techniques, we performed our first analysis using the linear reflectance model of the data-driven BRDF. For inter-subject variations, we expected a comparatively low variance. Hence, the first goal was to explore extreme variations, as they are to be expected for external skin conditions, such as sweat or make-up application. For this experiment, a single subject was subsequently exposed to the following conditions:

Cold. The subject is scanned immediately after being outside for 20 min under conditions of approximately -20°C (-4°F).

Hot. The face is held over a pot of hot water for 5 minutes. Afterward, sweat and condensed water are gently removed using a terry cloth.

Sweat. The face is again exposed to hot steam. For this measurement, however, the sweat is not to be removed.

Lotion. A fat lotion (Nivea Creme) is applied to to face.

Make-up. Application of make-up to the face.

Powder. The entire face is powdered using white cosmetic powder.

Neutral. As a reference measurement, the face is also scanned under neutral condition, that is, after two hours inside with no physical activity.

To analyze the surface BRDFs under these conditions, we randomly choose 5 000 points on the face and fit the complete non-parametric BRDF basis as discussed in Section 8.2.3. We then computed the average of the coefficients of all 5 000 points. Figure 10.3 shows the average BRDF for each external condition applied to a sphere and lit with point light sources from two different directions. As ex-

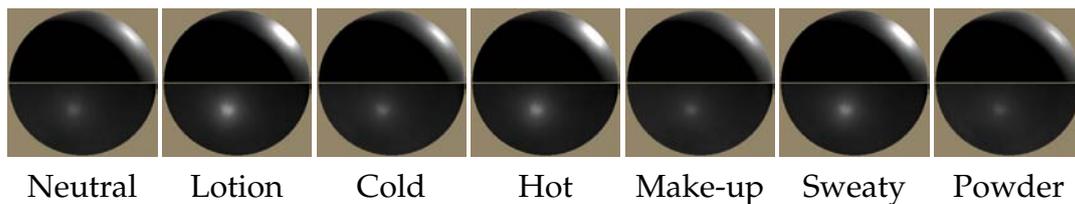


Figure 10.3: Visualization of the average surface BRDF for dataset A. Top hemispheres: Back lighting from grazing angle. Bottom hemispheres: Front lighting from lower left.

pected, there are noticeable differences between these BRDFs, especially between lotion / hot and cold / powder. An interesting observation is that the make-up BRDF closely resembles the neutral reference scan. This reflects the fact that make-up is designed to look as skin-like as possible.

To analyze the space of surface BRDFs of a larger population, we fit our non-parametric BRDF model to 2 000 randomly chosen points for 87 subjects from our database. As described in Section 8.2.3, we apply NMF dimensionality reduction to obtain a low-dimensional manifold that characterizes the BRDF space for each dataset. We perform the same reduction for the single subject under multiple external conditions, and, as a reference, for a single subject under neutral condition. For each analysis, Figure 10.4 shows relative reconstruction error as a function of the number of basis BRDFs. To show the three curves in the same plot we normalized the relative errors to a common scale. As expected, the plot suggests that four NMF basis BRDFs are sufficient for one person, whereas a single subject under varying external conditions requires six, and a common basis for all 87 subjects requires at least eight NMF basis BRDFs. Matusik et al. [MPBM03] concluded that a 45-dimensional linear (PCA) basis is required to span the space of isotropic BRDFs. Our results suggest that BRDFs representing skin are a small subset of all isotropic BRDFs.

Searching for inter-subject BRDF variations that reflect traits exposed by our pre-classification, we compared mean-differences between clusters of datasets, separated by age, skin type, and gender. This simple analysis did not show any

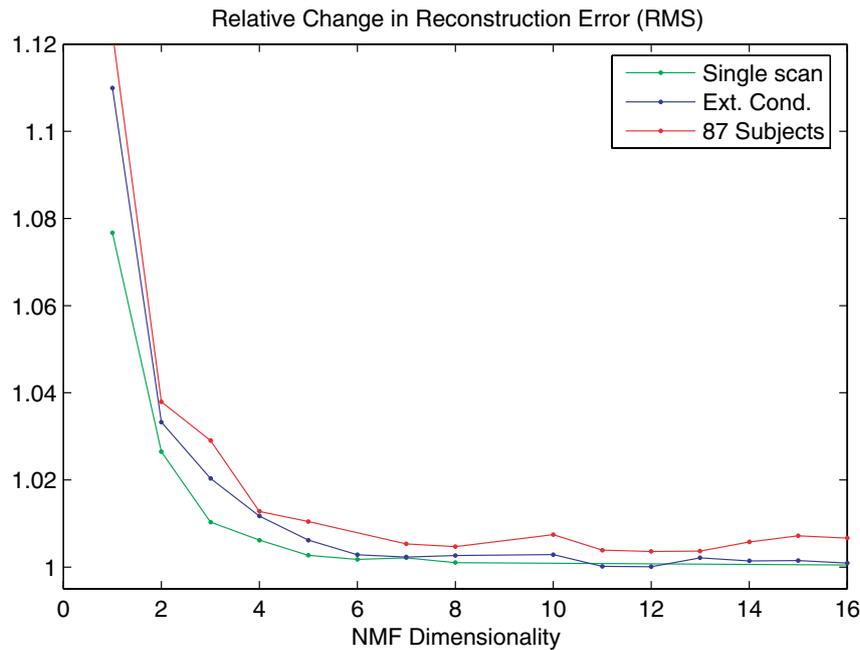


Figure 10.4: Reconstruction error as a function of the number of NMF basis BRDFs. The analysis is performed on the spatially varying BRDF of a single subject, of a single subject under 7 external conditions, and across surface BRDFs of 87 subjects from our database, respectively.

relevant difference depending on age or skin type. A very subtle, but statistically significant, difference exists between male and female surface reflectance. This difference may be due to the increased sebum production on male facial skin, however, it should not be over-interpreted, as variances can also be due to stubbles or make-up residuals that contribute a gender-specific bias to the data.

We come to the conclusion that subject-dependent variations in surface reflectance are not well-expressed. Gloss variations due to external conditions, however, are significant and can be described within our model.

Torrance-Sparrow

After turning to an analytic surface BRDF description, we continued our investigation in specular reflectance by analyzing the parameters of the Torrance-Sparrow BRDF model. During the model fit, the Fresnel coefficient has been fixed to a value from literature, leaving the free coefficients m and ρ_s for analysis. For future research, however, it would be interesting to include the Fresnel coefficient into the fit to investigate into differences in refractive index between different skin types and facial regions.

Figure 10.1 shows the 10 different facial regions we distinguished during the analysis. For each subject and each facial region we select a random subset of several thousand surface points. Figure 10.5 shows the parameter distribution for each face region. The ellipses are centered at the respective mean parameter values, with axes denoting the principal directions of variance. Principal Component Analysis (PCA) has been used to derive the variance ellipses.

A high variation can be observed especially in the (1) mustache, (3) eyebrows, (8) lips, (9) chin, and (10) beard region. All of these regions, except of the lips, are potentially covered with hair, potentially leading to erroneous normal estimates and over-fitting. Hence, the high variance does not necessarily correspond to a variance in skin reflectance. In contrast to these areas, the variation in the (7) cheek region, which is usually not covered by facial hair, is much lower. The increased reflectance variation of the lips may also be an artifact of lip-stick and lip-gloss application. Although all subjects have been instructed not to wear cosmetic products, many female subjects often still wore lip-stick. It is interesting that the mean value of the nose has the smallest roughness parameter m and the highest scaling factor ρ_s , corresponding to a very narrow specular lobe and a significant peak at the mirror reflection direction.

Despite the potential error sources, it can clearly be concluded that facial surface reflectance differs significantly between different face regions, which makes a spatially-varying model an absolute necessity for an accurate reproduction of human faces.

With the Torrance-Sparrow model, we also repeat the variational analysis in dependence of skin type, age, and gender. Equipped with the additional regional classification of surface points, we pre-processed each subject's dataset, determining the mean of m and ρ_s for each region. Experiments with linear (PCA) and non-linear dimensionality reduction techniques lead to a similar result as with the data-driven analysis: no significant correlation between skin properties and the pre-classifications can be found. The data is tightly clustered in a high-dimensional ball that cannot easily be reduced to a low dimensional space. As an alternative, we use Canonical Correlation Analysis (CCA). Canonical correlation analysis finds projections of two multivariate data sets that are maximally correlated. In our case, the data is projected (orthogonally) onto the two dimensions that are most strongly correlated with the experimental variables. Using CCA it turns out that the first dimension is best predicted by skin type; the second by gender. Figure 10.6 illustrates the projection of our data into this subspace. The skin type is encoded by color (light = I/II, black = VI), gender by marker type (\diamond = F, \circ = M), and age by marker size. The principal correlations are 0.768 and 0.585 and the direct correlations are 0.766 for skin type, 0.573 for gender, and 0.537 for age. This gives rise to the conclusion that in contrast to age and gender, skin type is significantly correlated with skin surface reflectance.

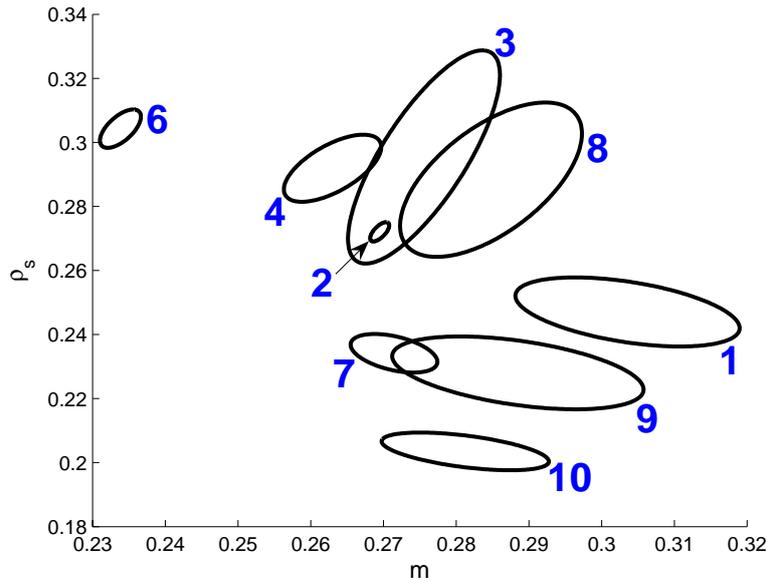


Figure 10.5: Variation of Torrance-Sparrow coefficients per face region, averaged over all subjects. The center indicates the mean and the axes show the result of a PCA. The direction and size of the axes correspond to the eigenvectors and eigenvalues of the covariance matrix.

10.1.2 Translucency Variations

The translucency component accounts for non-local subsurface scattering in the epidermis and dermis. It is a slowly varying effect that is responsible for much of the red color and soft shadows we see in human faces. It is important to note that translucency cannot be estimated directly from the reflectance images, which is why we use the additional subsurface scattering measurements.

Table 10.3 shows the mean and variance of μ_{tr} within our face database. The mea-

μ_{tr} [mm ⁻¹]	Cheek		Forehead		Neck	
	Mean	Var.	Mean	Var.	Mean	Var.
red	0.5572	0.1727	0.5443	0.0756	0.6911	0.2351
green	0.9751	0.2089	0.9831	0.1696	1.2488	0.3686
blue	1.5494	0.1881	1.5499	0.2607	1.9159	0.4230

Table 10.3: Mean and variance of μ_{tr} for dataset B.

surement points on cheek and forehead are quite similar in translucency. The measurement point on the neck underneath the chin shows a rather different mean, but also higher variance. This is probably due to measurement noise, as

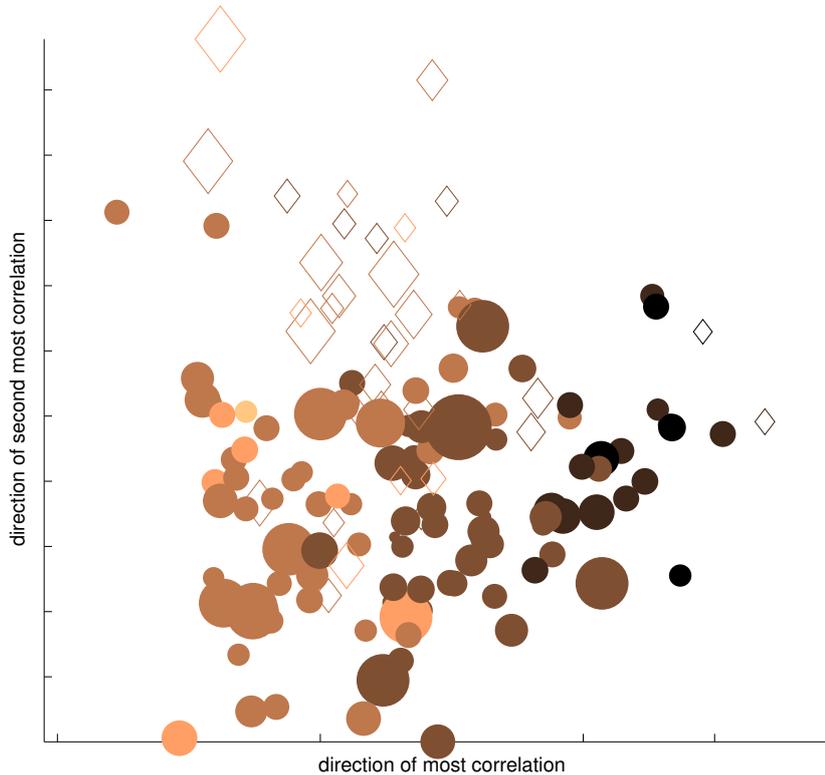


Figure 10.6: Canonical Correlation Analysis of the Torrance-Sparrow BRDF parameters. Color encodes skin type (light = I/II, black = VI); shape encodes gender (\diamond = F, \circ = M); marker size encodes age.

the sensor head is hard to place there. Overall, translucency values do not vary much between measurement points and between individuals. In practice, realistic result could be achieved approximating translucency by a single value for the whole face.

Figure 10.7 shows closeups of the subjects with minimum (0.3558, 0.7932, 1.5173) and maximum (0.9171, 1.5682, 1.6569) values for μ_{tr} in the face database. (Note that we define translucency as $1/\mu_{tr}$.) There are subtle differences visible at shadow boundaries. Figure 10.8 shows closeups computed with our model using the same minimum and maximum translucency values. Note that the model is capable of reproducing the subtle differences of Figure 10.7.

As described in Section 8.2.2, we determine subsurface scattering parameters μ'_s and μ_a from the albedo observations $\alpha(x_o)$ by fixing μ_{tr} . Figure 10.9 shows parameter estimates that were obtained from average diffuse albedo α_{avg} and average translucency $1/\mu_{tr}^{avg}$ for each subject. Subject classification is again denoted as in Figure 10.6.

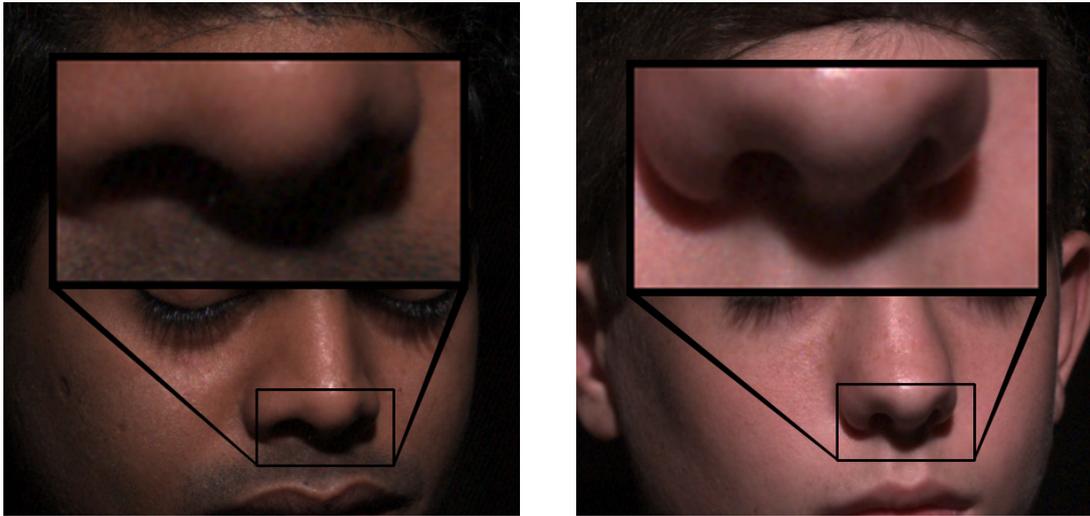


Figure 10.7: Photographs of subjects with minimum (left) and maximum (right) translucency values in dataset B. The differences at shadow boundaries are subtle.



Figure 10.8: Synthetic images with minimum (left) and maximum (right) translucency values.

As expected, subjects with a darker skin have higher absorption coefficients than subjects with lighter skin. Apart from this, there seems to be no significant correlation with gender or age. μ_a is significantly higher for the green and blue channel. This effect can be observed in real images as a redish color bleeding at shadow boundaries, and it is mainly due to hemoglobin.

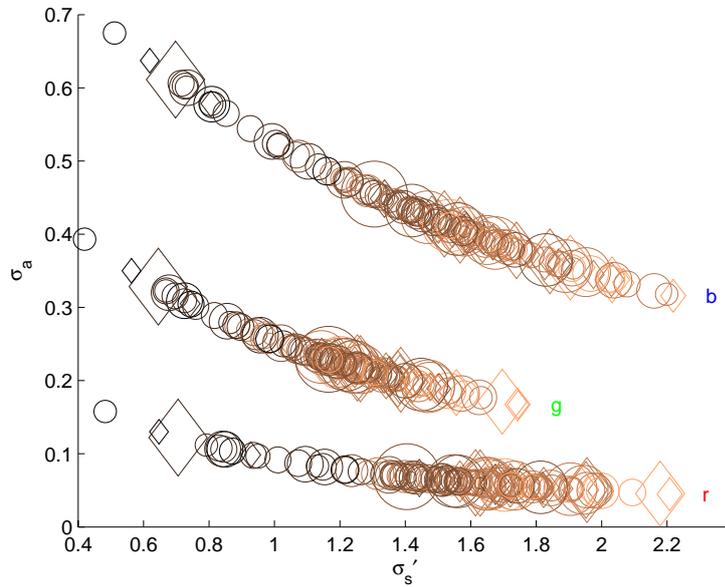


Figure 10.9: Plot of μ_a [mm^{-1}] and μ'_s [mm^{-1}] for all subjects. Color encodes skin type (light = 1, black = 6); shape encodes gender (\diamond = F, \circ = M); marker size encodes age. The values have been derived from α_{avg} for a constant μ_{tr}^{avg} .

10.1.3 Albedo Texture

The diffuse albedo maps capture the most individual-related information together with the 3D geometry. Skin color, freckles, moles, scars, facial hair, and a whole variety of other information is stored in the albedo map which makes it highly unique for each subject. For the analysis part we mainly extracted histograms of different skin types and regions and forgo a correlation analysis, e.g. with skin types, because the results would be obvious. Figure 10.10 shows albedo histogram for each skin type. We computed them using a random subset of albedo values from each subject without distinguishing different face regions. As expected, for darker skin types the curves cover a lower range on the abscissa. The data of these plots is useful for example to transfer a certain skin type to an individual subject by matching the histogram of the subject with the histogram of the skin type, as will be demonstrated in Section 10.2.

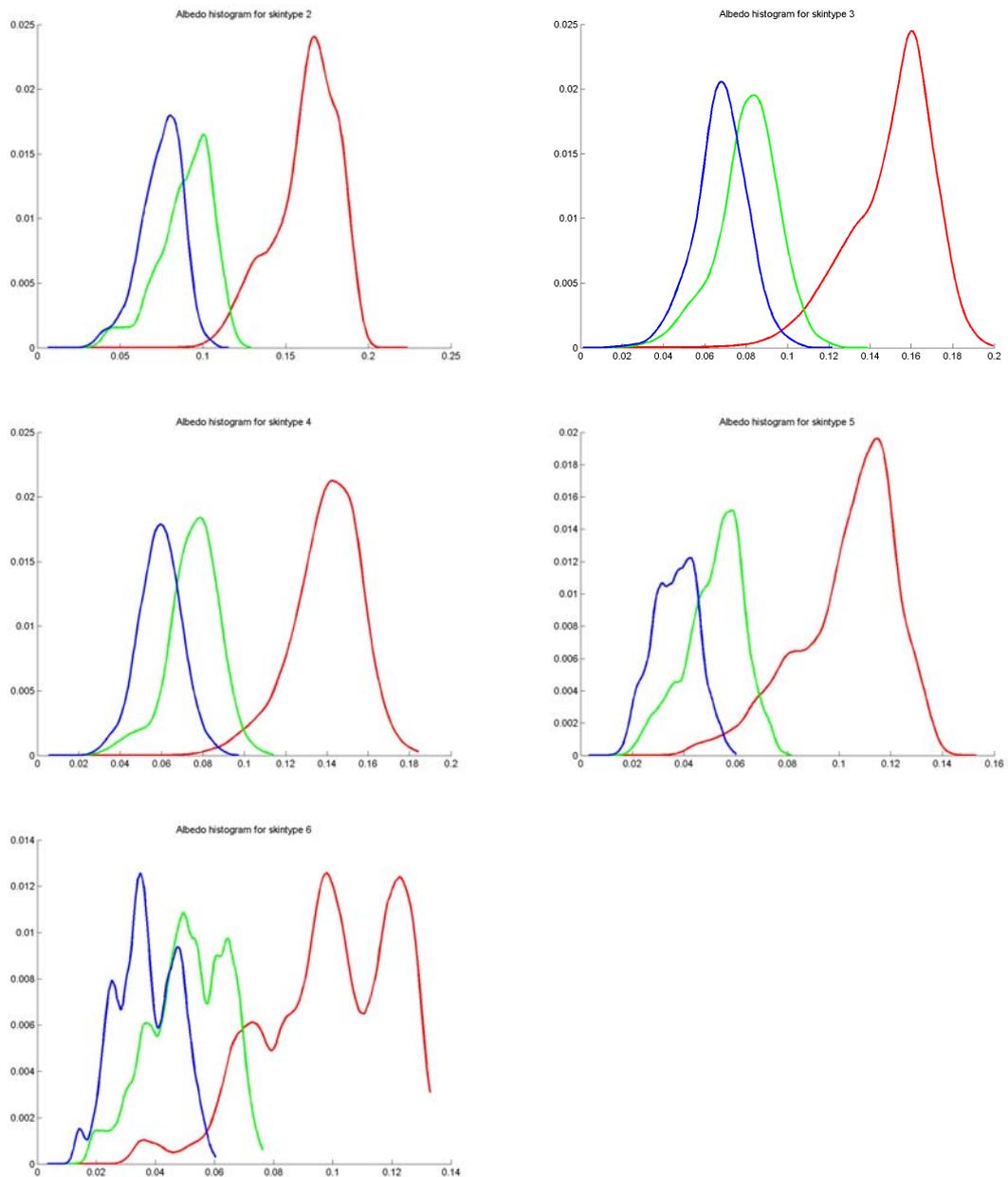


Figure 10.10: Albedo histograms for each skin type without distinguishing different face regions. Again, skin type I and II are combined in a single histogram.

10.1.4 Further Analysis

We further produced histogram plots of albedo maps, Blinn-Phong and Torrance-Sparrow BRDF parameters for all combinations of skin types and facial regions. Additionally, we determined the mean Blinn-Phong and Torrance-Sparrow BRDF parameters for all combinations of skin type, age group, gender, and face region. All information has been made publicly available at <http://www.merl.com/to/be/announced.html>.

Our model treats each component independently, allowing us to render or edit each one without affecting the other two. Consequently, we analyzed each component separately and did not consider a potential correlation between them. In the future we would like to perform a more in-depth statistical analysis that takes cross-correlations into account.

10.2 Face Editing

Most of our face reflectance model parameters are intuitive for editing, such as analytic BRDF parameters, average skin color, or detailed skin texture. To change the surface reflectance of a face model, we can either edit the parameters of the analytic BRDF model directly, or we can adjust them using the statistics from our database. This allows to change all appearance parameters based on, for example, different external conditions (e.g., from neutral to sweaty), or by classifications after skin type or age.

10.2.1 Trait Vectors

In case of the data-driven BRDF model, where intuitive parameters do not exist *per se*, we use mean differences to navigate the low-dimensional spaces of surface BRDFs (using their NMF basis), similar to [BV99, MPBM03]. We compute the average of the basis vectors in each complementary pair of clusters associated with a trait (i.e., those faces with, and those without). The differences between the complement averages provide *trait vectors* that can be used for navigation and interpolation between traits.

Specifically, we compute trait vectors for the NMF basis of surface BRDFs in the face database (see Section 10.1.1). To synthesize a desired BRDF appearance, we apply a linear combination of the trait vectors to the BRDF coefficients of a source face. Figure 10.11, top row, shows for example the application of the *lotion* trait vector derived from the experiment in Section 10.1.1.



Figure 10.11: Face editing by adding trait vectors to different source faces. Top: Changing the surface BRDF using the acquired lotion trait. The example interpolates from “neutral” to a two-fold exaggeration of this trait. Bottom: Changing translucency, albedo histogram, and surface BRDF from skin type IV to II. (Both models use spatially subsurface scattering parameters and a data-driven BRDF model. No micro-normals are used.)

10.2.2 Histogram Interpolation

A more general technique that can equally be applied to all parameters of the model is the histogram interpolation technique presented in [MZD05]. For each face region, we can match the parameter histogram of the source face with another parameter histogram using histogram interpolation, as shown in Figure 10.12. We



Figure 10.12: Face appearance transfer from source faces (upper row) using the statistics of a target to the results (lower row). Left: Transferring freckles using the method by Heeger and Bergen. Center: Transferring lotion by matching histograms of BRDF parameter textures. Right: Skin type transfer by matching histograms of the diffuse albedo texture. (All models use static diffusion, a modulation texture, and Torrance-Sparrow surface reflectance. The models contain micro-normals.)

can also transfer skin color histograms from one subject to another. This allows us to change, for example, the skin type (e.g., from II to V in Figure 10.12, center), but also the parameters of an analytic reflectance model, for example.

To transfer skin texture, we use the texture analysis technique of Heeger and Bergen [HB95]. We first transform the albedo map of each subject into a common and decorrelated color space using the method discussed in [HB95, Section 3.5] and process each transformed color channel independently. We then compute statistics (histograms) of the original albedo map at full resolution, and of filter responses at different orientations and scales organized as a steerable pyramid. We use seven pyramid levels with four oriented filters, and down-sample the albedo map by a factor of two at each level. Each histogram has 256 bins. The first level of the pyramid is a low-pass and high-pass filter computed on the full-resolution albedo map. The output of the low-pass filter is a histogram of average skin color.

The histograms of all pyramid filter responses (30 in total) and the histogram of the original albedo map are concatenated into a $256 \times 31 \times 3 = 23\,808$ element vector H . To transfer skin textures, we use the histogram vector H of a source and a target face and apply the histogram-matching technique of Heeger and Bergen [HB95]. We can either use the histograms of the whole albedo map or restrict the analysis/synthesis to certain facial regions. Note that Heeger and Bergen start their texture synthesis with a noise image as the source. In our case it makes more sense to start from the source albedo map. To allow for sufficient variation during histogram matching, we add some noise to the source albedo map before we compute its histogram vector H . A similar method was used by Tsumura et al. [TOS⁺03] for synthesis of 2D skin textures, and by Cula and Dana [CD02] to analyze bi-directional texture functions (BTFs) of skin patches. Figure 10.12 shows transfers of skin texture, BRDF, and skin color for three face models. In this example, the translucency values of the source and result models remain the same. Figure 10.11, bottom, shows a skin texture transfer from type IV to II, including translucency and data-driven BRDF interpolation along the respective skin type trait vector. Additional examples for albedo changes by the method of Heeger and Bergen can be found in Figure 10.13.



Figure 10.13: *Albedo changes. From left to right: Real photograph; rendering; making the face sun-burnt; adding hair follicles in the beard area; making the skin type darker.*

10.3 Discussion

In this chapter we analyzed specular and diffuse reflectance properties across a larger population of people. The pre-classification of subjects allowed us to relate certain parameter variations to individual traits, such as age, gender, skin type, and skin properties artificially induced by external conditions. We showed how the results can be used for appearance editing, which has been a major goal of this thesis. In addition, we believe that the representative parameters that we found for translucency, albedo, and BRDF parameters will prove to be useful for *ab initio* design of skin properties of artificial face models.

The face database is currently skewed toward Asian and Caucasian males. Nevertheless, it already covers sufficient variety to allow for an analysis of various traits. Future analysis may concentrate on a directed extension of the database with respect to female subjects, with particular interest in the aspects of facial makeup.

Conclusion and Outlook

This chapter provides a short summary of the thesis and points out possible directions for future work in this field.

11.1 Conclusion

The focus of this thesis has been the development of a practical skin model to represent the reflectance properties of human skin. An important feature of our model is that all its parameters can be robustly estimated from measurements. This reduces the large amount of measured data to a manageable size, facilitates editing, and enables face appearance changes. Images from our model come close to reproducing photographs of real faces for arbitrary illumination and pose.

The skin model is based on skin measurements that were taken from a large group of subjects. To this end, respective acquisition hardware and a suited data processing pipeline had to be developed. We contribute two novel acquisition devices and a fitting procedure to obtain respective parameters for our skin model. We introduce new methods for scan processing, motion compensation, and direct rendering of the acquired reflectance data. We provide a careful analysis of skin measurements to evaluate the plausibility of our modeling approach. An analysis across 149 different subjects provides insight into the variance of face reflectance parameters with respect to age, gender, or skin type. We show how the analysis re-

sults can be used to define intuitive editing operations to alter the skin appearance of a scanned subject. A detailed list of contributions can be found in Chapter 1.

11.2 Discussion and Future Work

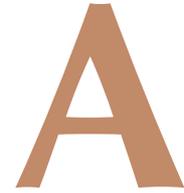
As discussed in Section 9.3, there is room for improvement in our model. For example, it would be interesting to measure absorption and scattering parameters over a dense spectrum, rather than using only three color channels. It would also be interesting to compare the results from the diffuse dipole approximation with a full Monte Carlo subsurface scattering simulation. Differences of the two models have been analyzed for extreme situations [LPT05], but it remains open how relevant the distinction is for facial geometry. Multi-layered skin models still bear the potential for increased realism [DJ05]. It may be worth to investigate in rapid acquisition techniques for *in vivo* measurements of the skin's layered structure.

Retrieving more densely sampled spatially varying subsurface scattering measurements may reveal a subtle variational distribution that could be taken into respect. Future work may address the construction of a respective measurements setup, eventually using spatially varying illumination, e.g., using eye-safe projectors and highly sensitive cameras.

Other important areas that require a different modeling approach are facial hair (eyebrows, eyelashes, mustaches, and beards), hair, ears, eyes, and teeth. Very fine facial hair also leads to asperity scattering and the important “velvety” look of skin near grazing angles [KP03]. Our model does not take this into account.

We captured face reflectance on static, neutral faces. Equally important are expressions and face performances. For example, it is well known that the blood flow in skin changes based on facial expressions. Our setup has the advantage that such reflectance changes could be captured in real-time.

A P P E N D I X



Glossary And Notation

A.1 Notation

\mathbb{R}, \mathbb{R}^3	The set of real numbers, three-dimensional space.
\mathcal{P}^3	Homogeneous three-space, specified by \mathbb{R}^4 vectors.
$s, t, u, v, x, y, z, w, \varepsilon, \varphi, \vartheta$	Scalar variables.
e, d	Scalar constants and operators.
x, p, u, v	Vectorial variables.
\mathbf{x}, \mathbf{p}	Homogeneous vectorial variables in \mathcal{P}^3 .
$\langle \mathbf{n}, \mathbf{h} \rangle$	Dot product.
$\mathbf{u} \times \mathbf{v}$	Cross product.
M, N	Matrices.
$M^\top, M^{-\top} := (M^\top)^{-1}$	Matrix transposition and inverse transposition.
A	A surface patch or its area.
Ω	Solid angle [sr].
Ω_+	Positive hemisphere over a surface point.
$\mathbf{n}, \mathbf{u}, \mathbf{v}$	Normal vector and two tangential vectors spanning a skewed local surface frame. $\mathbf{n} = (\mathbf{u} \times \mathbf{v}) / \ \mathbf{u} \times \mathbf{v}\ $.
$\mathbf{n}, \mathbf{t}_1, \mathbf{t}_2$	Normal vector and two tangential vectors spanning an orthonormal local surface frame.

- ω Hemispherical direction, represented as $\omega \in \mathbb{R}^3$.
Alternatively expressed as (ϑ, φ) .
- v, l View and lighting direction if no hemispherical parameterization exists.
- BRDF Bi-directional reflectance distribution function.
- BSSRDF Bi-directional surface scattering reflectance distribution function.
- DDBRDF Data-driven BRDF.
- SRF Surface reflectance field.

A.2 Radiometric Quantities

Table A.1 gives an overview over radiometric quantities used in this thesis. The face scanner processing pipeline strictly uses SI units. Before rendering, distances are converted to millimeters. For more information see Section 2.1.

<i>Name</i>	<i>Symbol</i>	<i>Definition</i>	<i>SI Unit</i>
radiant energy	Q	<i>basic unit of radiometry</i>	J
radiant flux	Φ	$\Phi = dQ/dt$	W
radiance	L	$L(\mathbf{x}, \omega) = \frac{d^2\Phi}{d\Omega dA_{\text{source}}}$	$\text{W sr}^{-1} \text{m}^{-2}$
radiant intensity	I	$I(\omega) = d\Phi/d\Omega$	W sr^{-1}
radiant exitance, radiosity	M, B	$M(\mathbf{x}_o) = d\Phi/dA_{\text{source}}$ $= \int_{\Omega_+} L(\mathbf{x}_o, \omega) n d\omega$	W m^{-2}
irradiance	E	$E(\mathbf{x}_i) = d\Phi/dA$ $= \int_{\Omega_+} L(\mathbf{x}_i, \omega) n d\omega$	W m^{-2}

Table A.1: Radiometric quantities and their definitions.

System Calibration

This chapter provides a short summary of the calibration procedure of the face scanner dome. For the calibration of the subsurface scattering measurement device, please refer to Section 6.2.3.

The face scanner consists of 16 cameras, 150 light sources, and a commercial 3D scanner. During an acquisition sequence, all light sources are subsequently triggered. The cameras are synchronized with the light sources, acquiring images of all generated illumination conditions. In our case this leads to $16 \times 150 = 2'400$ images, a *reflectance field* of the subject's face. Immediately before the reflectance field acquisition, the built-in 3D scanning system acquires the facial geometry of the scanned subject. Since its construction, the scanner is used for various different applications. Most of them require geometric and photometric calibration of all components. In the remainder we describe the necessary calibration steps.

B.1 Overview

The three main components of the scanner require the following calibrations:

- Camera System *Extrinsic and intrinsic calibration.* This includes lens distortion coefficients and a projective transformation from world to image coordinates.
Color Characteristic of each camera, describing its individual spectral response.
A radiometric calibration is implicitly performed during light source calibration.
- Light Sources *Positions* inside the dome construction.
Intensity distribution within each light cone, including light source color and light spot fall-off parameters.
- 3D Scanner *Registration* with the common world coordinate system. Other calibration steps required by the 3D scanner are performed by its proprietary software.

The remainder describes these calibrations in detail.

B.2 Camera System

All 16 cameras are identical models, using the same lens type. However, each lens and camera shows its individual geometric and colorimetric characteristic, requiring an extensive calibration procedure.

B.2.1 Geometric Calibration

Geometric camera calibration models each camera's imaging transformation from world to image coordinates. For a pinhole camera model this can be expressed in homogeneous coordinates as

$$\mathbf{p}_I = H\mathbf{p}_W, \quad \mathbf{p}_W = (x, y, z, 1)^\top \in \mathcal{P}^3, \quad (\text{B.1})$$

where \mathbf{p}_W denotes the point's world coordinates, and the Euclidean interpretation of \mathbf{p}_I provides its image coordinates.

A camera calibration is usually separated into two parts: extrinsic and intrinsic calibration. The former contains a rigid transformation, aligning world points with a canonic local camera coordinate system. It implicitly contains position and orientation of the camera in space. The latter models the camera's lens system, including focal length, and aspect ratio and placement of the CCD sensor at the end of the optical path.

H can accordingly be decomposed in two matrices $H = IE$. In contrast to the pinhole camera model, real-world cameras show lens distortion. This has to be modeled by a non-linear distortion function d :

$$\mathbf{p}_I = d(IE\mathbf{p}_W) . \quad (\text{B.2})$$

There is a variety of common distortion models, cf. for example [WCH92, Zha00].

In our system, we are using a standard implementation of the Intel OpenCV library [OCV] to determine camera distortion parameters for each single camera. The algorithm uses a set of images taken of a checker board etalon. It computes I and a set of radial and tangential distortion parameters k_1, k_2 , and p_1, p_2 , respectively. We are using the distortion parameters to undistort the camera images in all subsequent steps. The intrinsic matrix I is solely used as a starting guess for a global multi-camera calibration.

The multi-camera calibration uses subsequent images of a small LED that is swept through the working volume. A global, non-linear, Euclidean bundle adjustment over all camera parameters tries to find camera matrices H_I for each camera so that the different LED observations are consistent with a single 3D position for each frame. The resulting calibration is unique up to scale, rotation, and translation.

For a set of 10 to 16 cameras, a number of 200–300 images per camera turned out to be sufficient.

B.2.2 Photometric Calibration

This section describes the calibration procedure and the color/intensity modeling required to color correct camera images to obtain surface reflectance samples.

Black Images

Every raw image acquired by a camera should be shifted by subtracting the corresponding *black image*. A black image is an image that is acquired shortly before the data acquisition, either with a black cap in front of the lenses or with all illumination turned off, depending on the application. Its purpose is twofold:

- ▶ Subtraction of the DC offset in image sensor and AD converter.
- ▶ In case of a controlled lighting environment (as, e.g., in the face scanner dome): Additional subtraction of the ambient illumination.

We are capturing black images before acquiring images for the light cone calibration and before acquiring any reflectance field.

Black capped-lens-images also have to be subtracted from the images acquired for vignetting and color calibration. Even for the capped-lens-images it is important to acquire the respective black images right before the image acquisition, as the black image is highly temperature dependent.

Vignetting & Color Calibration

Vignetting and color calibration build a unit. They perform two tasks:

- ▶ Compensation for the radial intensity fall-off due to masking effects inside the lens system (including aperture).
- ▶ Relative sensor calibration of the cameras between each other.

Vignetting/color calibration “equalizes” all cameras, i.e., after vignetting/color correction, all cameras should produce identical images under an identical situation.

The calibration procedure determines for each camera i a vignetting/color correction function c_i . Given an acquired Image I_a and the corresponding black image I_b , the vignetting/color corrected image I_{cc} is

$$I_{cc} = c_i(I_a - I_b) . \quad (\text{B.3})$$

The color correction functions are determined in a way that all cameras i ($i \geq 1$) are corrected to match the characteristic of camera 0:

$$c_i(I_a^i - I_b^i) = I_a^0 - I_b^0 , \quad (\text{B.4})$$

provided that $I_{\{a,b\}}^i$ are images taken by cameras i under identical situations. (c_0 is by definition the identity mapping.) In our project, we use c_i of the following structure (let c be color sample at image position (x, y)):

$$c_i(c) = \frac{1}{I_V^i(x, y)} A c + t, \quad A \in \mathbb{R}^{3 \times 3}, \quad c, t \in \mathbb{R}^3. \quad (\text{B.5})$$

That is, we are using an affine color correction model. See [Dan92, FL00, Ami02] for a discussion of different color calibration models. The color calibration code determines A and t for each camera. Note that I_V can be given in arbitrary units, as the respective scale is compensated by the scale of A .

B.3 Light Sources

Surface reflectance analysis requires knowledge of the irradiance at each surface point. This requires a geometric model of the light sources, as well as a description of the angular intensity variation within the light cone.

B.3.1 Light Source Positions

We currently do not explicitly calibrate for the light source positions. They are rather derived from the extrinsic camera calibration, using the knowledge of the dome construction.

As the extrinsic camera calibrations already have been transformed to match their positions in the dome’s CAD plan, the light source positions can simply be taken from that plan.

B.3.2 Radiometric Calibration

The vignetting and color calibration equalized all cameras. However, there is still no *absolute* color calibration. Acquisition of color always has to take the light source spectrum into account. We are obtaining an absolute color correction by calibrating the light source colors of all 150 light sources using a standardized reflectance target. The target hemispherically reflects 99.9% of the incident power, equally over all light frequencies in the visible spectrum.

In our setting, light source color is not the only issue. The light sources themselves show an equivalent to vignetting. They show a significant intensity fall-off within their light cones.

For each light source we are taking images of the illuminated reflection target. The target is 12'' × 12'' in size. It is placed in the center of the dome, roughly covering the same area as a subject’s face will cover. We acquire between 15 and 30 complete reflectance fields of the target under varying orientations. The target is surrounded by 16 fiducials that allow for a pose reconstruction of each target acquisition.

The observed pixel intensities along the white surface delivers a slice through the light cone of the light source. The reflection target is built to show Lambertian reflection, and its position is known. Hence, the light source’s radiant intensity in the direction of each surface point can be determined by normalizing the observations by the cosine-term and the distance d to the light source:

$$I_{\text{cone}} = \frac{4\pi d^2}{\langle \mathbf{n}, \mathbf{l} \rangle} c(I_{\text{observed}} - I_{\text{black}}), \quad \mathbf{l}, \mathbf{n} \in \mathbb{R}^3, \quad (\text{B.6})$$

with \mathbf{l} the illumination direction and \mathbf{n} the target’s surface normal.

Similar to the vignetting calibration we fit a polynomial to the light source’s light cone: We assign a calibration matrix to each light source, so we can treat it like a camera. The light cone fall-off is then expressed as a second-order polynomial in light source “image” space.

Our approach ignores the *near field* of the light source, but this is acceptable for the actual distances between light sources and objects. The light cone is fitted for each color channel separately. Thus, we automatically obtain a white balance compensating the light source color.

B.4 3D Scanner Registration

The 3D scanning system is treated as a black-box, equipped with its proprietary calibration procedure. However, the acquired 3D scans have to be registered with the common dome coordinate system.

We acquire a calibration target both with the 3D scanner and with the calibrated camera setup. The calibration target is equipped with markers which can be identified in the 3D scan, as well as in the camera images.

A non-linear optimization is used to reconstruct 3D positions m_i of corresponding markers in the camera images. Extracting their locations b_i from the 3D geometry reduces the registration problem to finding a scale s and an affine mapping (Q, t) to register b_i with m_i . We find the respective transformation by minimizing

$$\sum_i (s(Qb_i + t) - m_i)^2 \tag{B.7}$$

using a modified Procrustes algorithm.

Color Plates

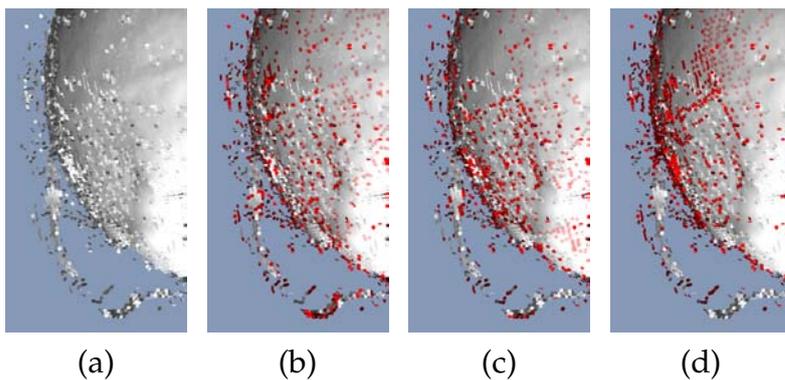


Figure 3.9, page 31: Three different outlier classifiers. Potential outliers marked in red. (a) Raw scanned geometry. (b) Classification using the miniball criterion. (c) Plane-fit criterion. (d) k -nearest-neighbor graph criterion. All criteria were thresholded to classify 7% of the surfels as outliers.

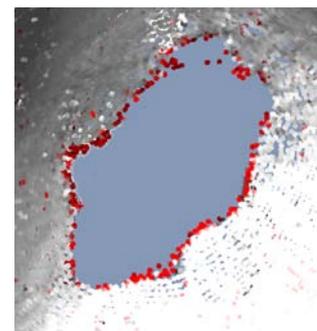


Figure 3.10, page 31: The nearest-neighbor outlier criterion performs poor around manifold borders.

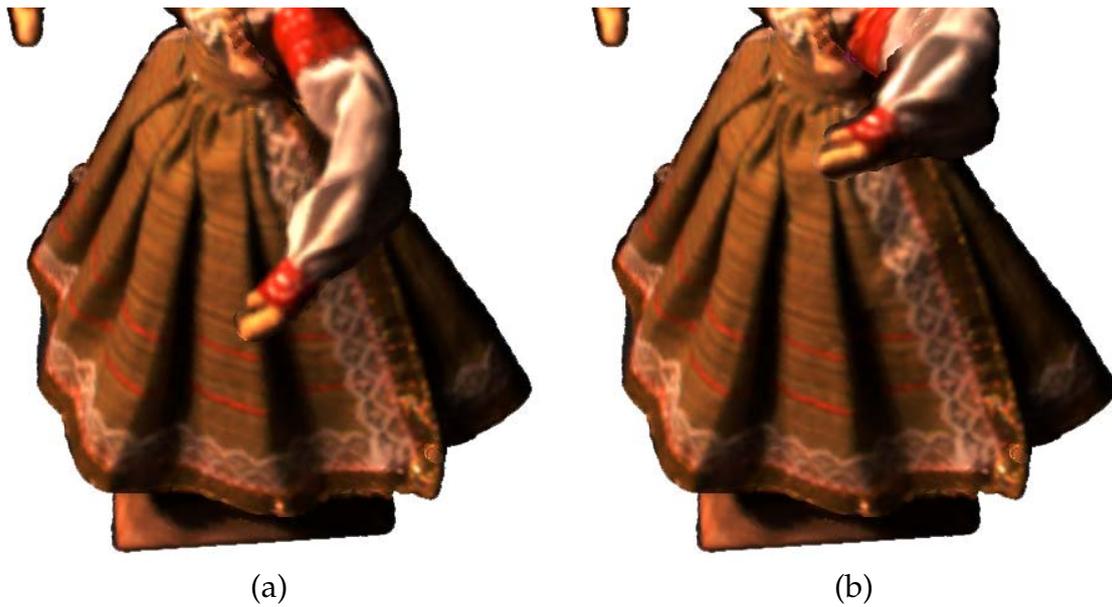


Figure 4.11, page 55: View extrapolation. (a) Original doll data set. (b) Occluded parts of the skirt become visible as one arm is lifted. The model is lit by two point light sources. (a) is chosen to be an original view to force the views used for view extrapolation to be at maximum distance from the current view.



Figure 4.12, page 56: Preservation of material properties Left: The original doll surface reflectance field. Center, Right: A sinusoidal warp is applied to the data set.



Figure 4.13, page 57: *Deformation of refractive objects. Different warps were applied to the mug data set. All glasses are rendered from the same viewpoint and lit by five point light sources (red, green, blue, 2×white). The effect of the deformations on specular highlights can easily be observed.*

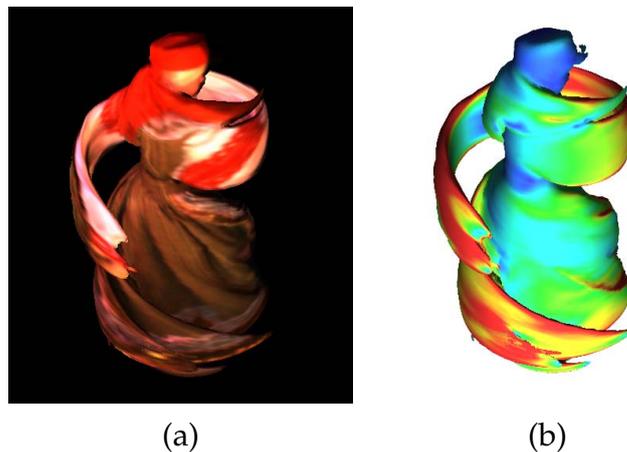


Figure 4.14, page 58: *Even under large deformations, the surfel renderer displays a closed surface. (a) The depicted warp anisotropically stretches parts of the doll data (b) The stretch anisotropy is color encoded showing the ratio of the principal axes of the elliptical surface splats. Splats in blue regions remain circular, green areas show a stretch ratio of 1:2, while red denotes a local stretch of more than 1:4.*



Figure 4.15, page 59: *Viewing ray divergence. Top: The doll's skirt was distorted. Left: The corresponding camera blending field displays the viewing ray divergence in distorted regions. Right: The light source blending field.*

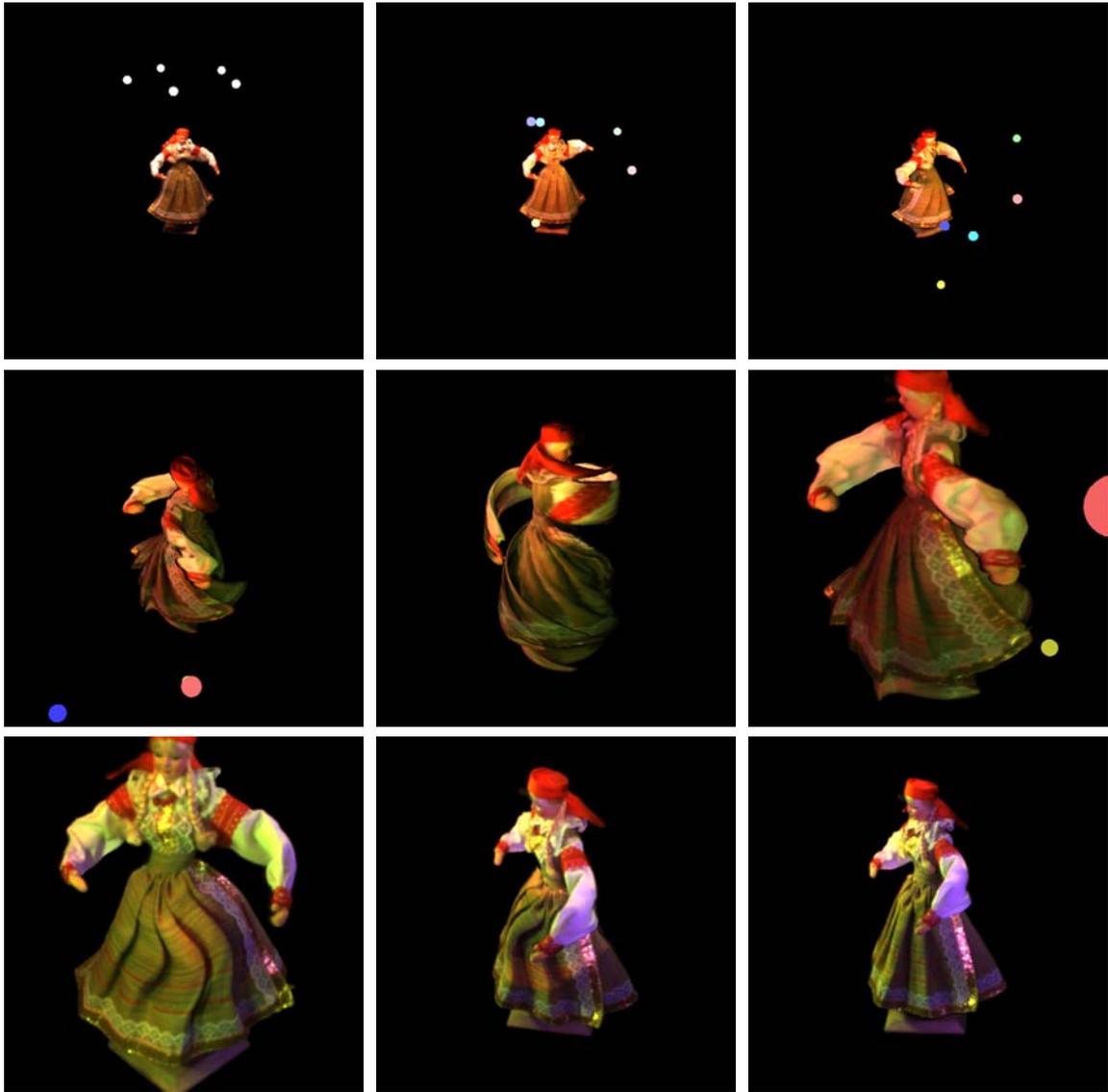


Figure 4.16, page 60: Selected frames of a surface reflectance field animation. The model is continuously deformed, while viewpoint and lighting conditions change. The model is lit by colored light sources, denoted by spheres of the light color.

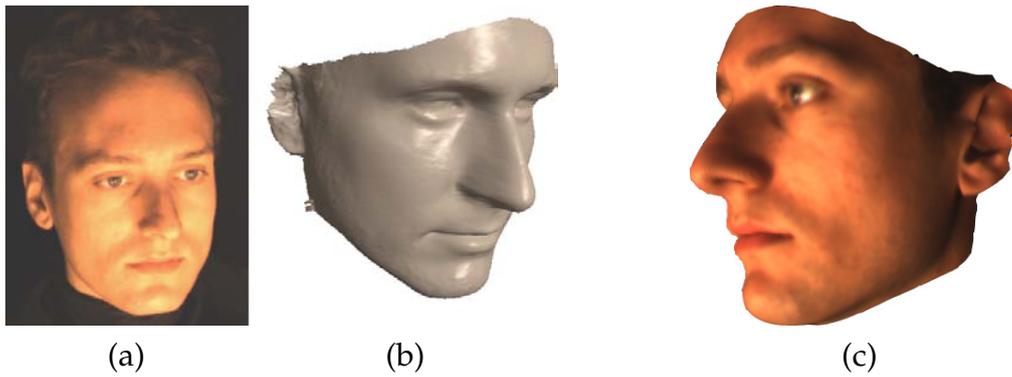


Figure 4.17, page 61: Face datasets obtained from silhouette fitting. (a,b) One of the input views and the fitted geometry. (c) A rendering using the deformable surface reflectance field renderer.

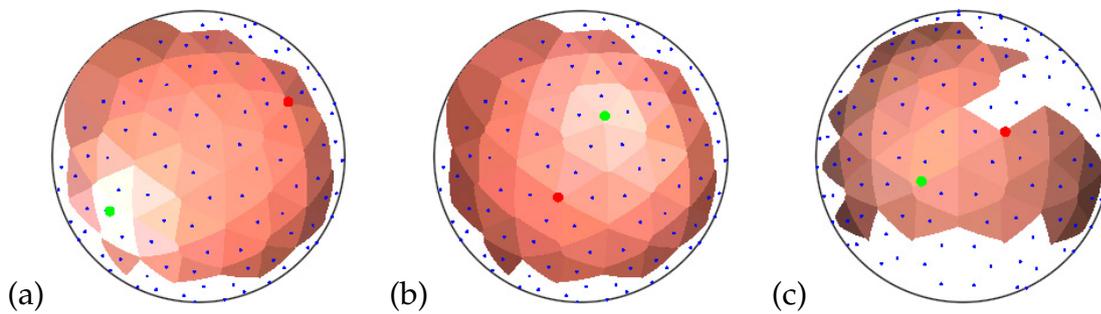


Figure 7.8, page 109: Visualization of two skin lumitexels for different viewpoints. Only the samples for a single camera are shown at the time. The red and green dots are the intersection of the view vector and reflection vector with the hemisphere, respectively. [see p. 109 for full caption]

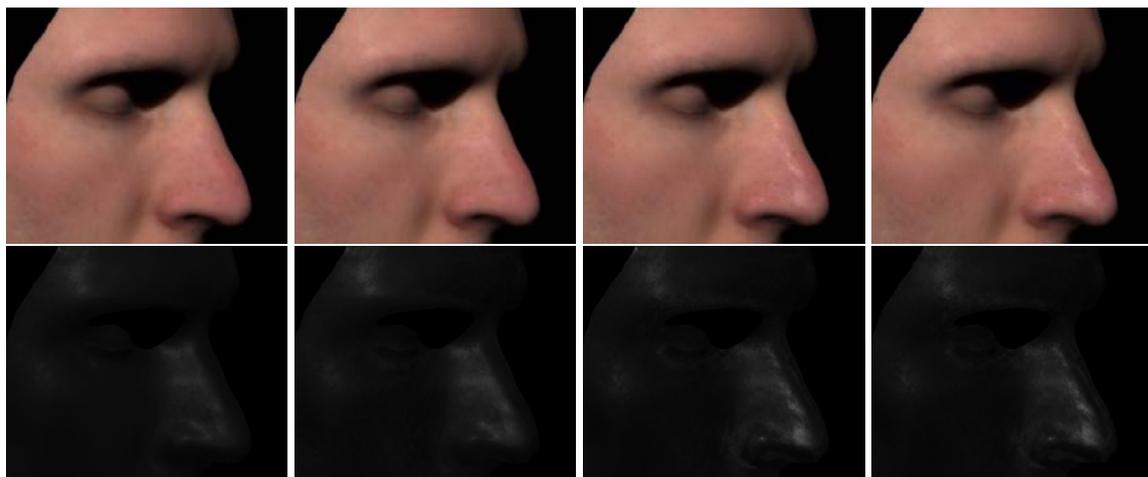


Figure 8.5, page 130: Synthetic images of a face using, from left to right, 1, 2, 4, and 24 NMF basis BRDFs. Top: Rendered images. Bottom: Surface reflection only. Four NMF basis BRDFs are sufficient to represent the surface reflectance of a single person.



Figure 7.11, page 114: Results of the normal estimation step. From left to right, top row: normals from Specular Direction, Helmholtz Stereopsis. Bottom row: photometric stereo with highlight removal, standard photometric stereo.

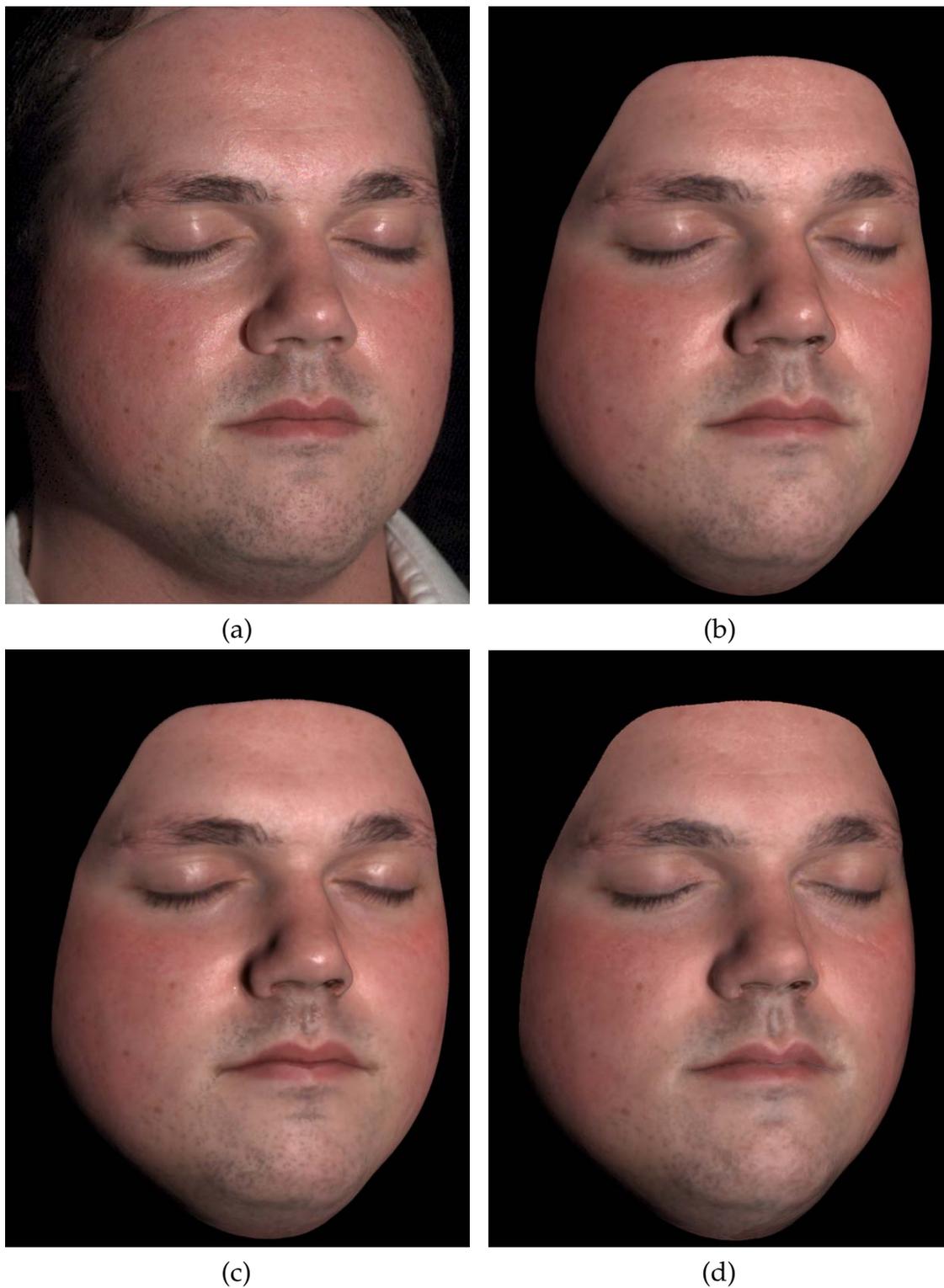


Figure 8.6, page 132: BRDF Models. (a) *Input photograph, compared to our face model with* (b) *spatially-varying Torrance-Sparrow; (c) spatially-varying Blinn-Phong; and (d) uniform Torrance-Sparrow BRDF models.*



Figure 9.1, page 137 (I): Comparison of real photographs of four subjects (top) with fits of our model (bottom).



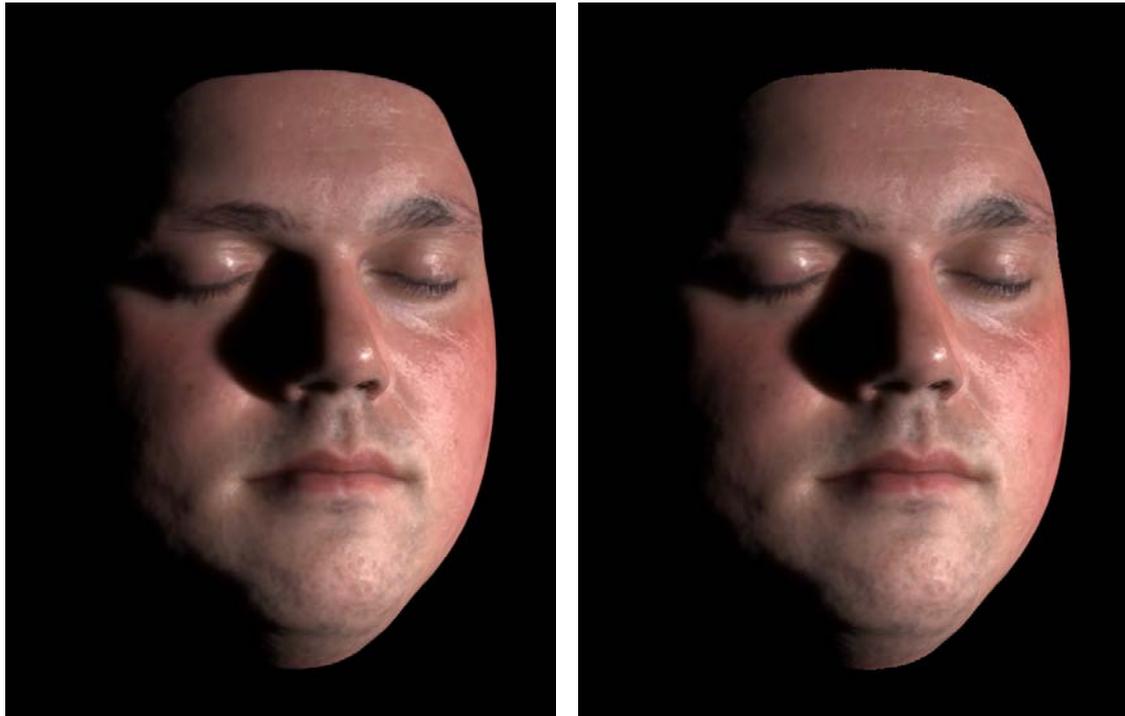
Figure 9.1, page 137 (II): Comparison of real photographs of four subjects (top) with fits of our model (bottom).



Figure 9.2, page 138: *Components of our model using a data-driven BRDF. (a) One of the input images with single light source illumination. (b) 3D surface scan shaded with the BRDF approximation of the diffuse subsurface term. The model is lit using the camera and light source calibration derived for view (a). (c) Subtracting the estimated diffuse reflection term (b) from (a) reveals surface reflection as used for the BRDF fit. (d) Reconstructed model. (e) Underlying dipole diffusion. Note the smoother shading and soft shadows in comparison to (b), while preserving texture detail. (f) Reconstructed surface reflection.*



Figure 9.3, page 139: Components of the model using the modulation texture and a Torrance-Sparrow BRDF. (a) 3D geometry. (b) 3D geometry and albedo texture; no diffusion process is shown. (c) Specular reflectance. (d) Color encoded Torrance-Sparrow coefficients m and ρ_s . (e) Rendering based on the complete model. (f) Original input image.



(a)

(b)



Albedo



μ'_s



μ_a

Figure 9.4, page 140: *Two approaches to model diffuse reflectance. (a) A constant average diffusion term is modulated by an albedo texture. (b) Spatially varying material properties μ'_s and μ_a control the diffusion process. The bottom row shows the respective coefficient textures.*

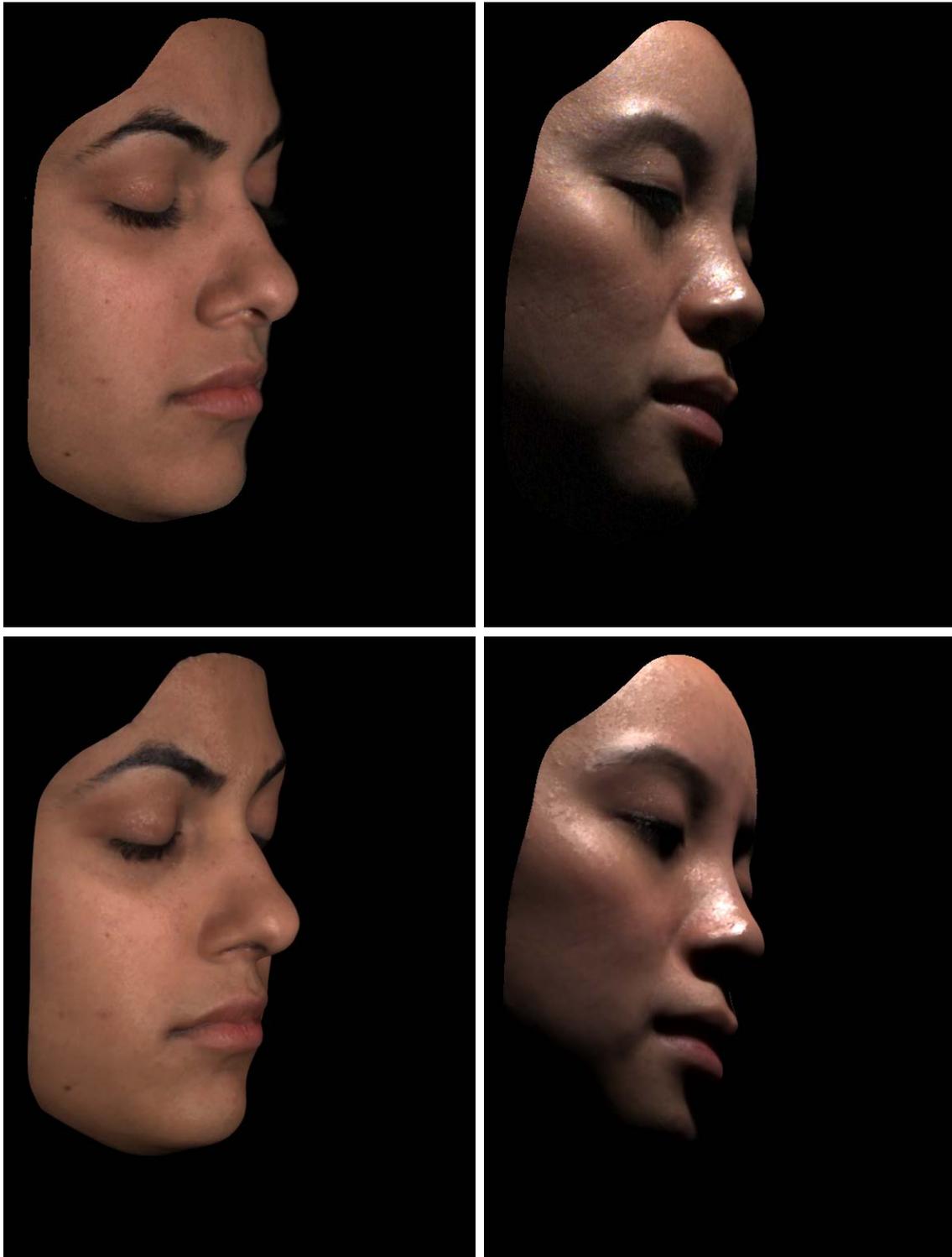


Figure 9.5, page 142 (I): Comparison of real photographs (top) to our model (bottom). All photographs were cropped according to the 3D model to remove distracting features.



Figure 9.5, page 142 (II): Comparison of real photographs (top) to our model (bottom). All photographs were cropped according to the 3D model to remove distracting features.

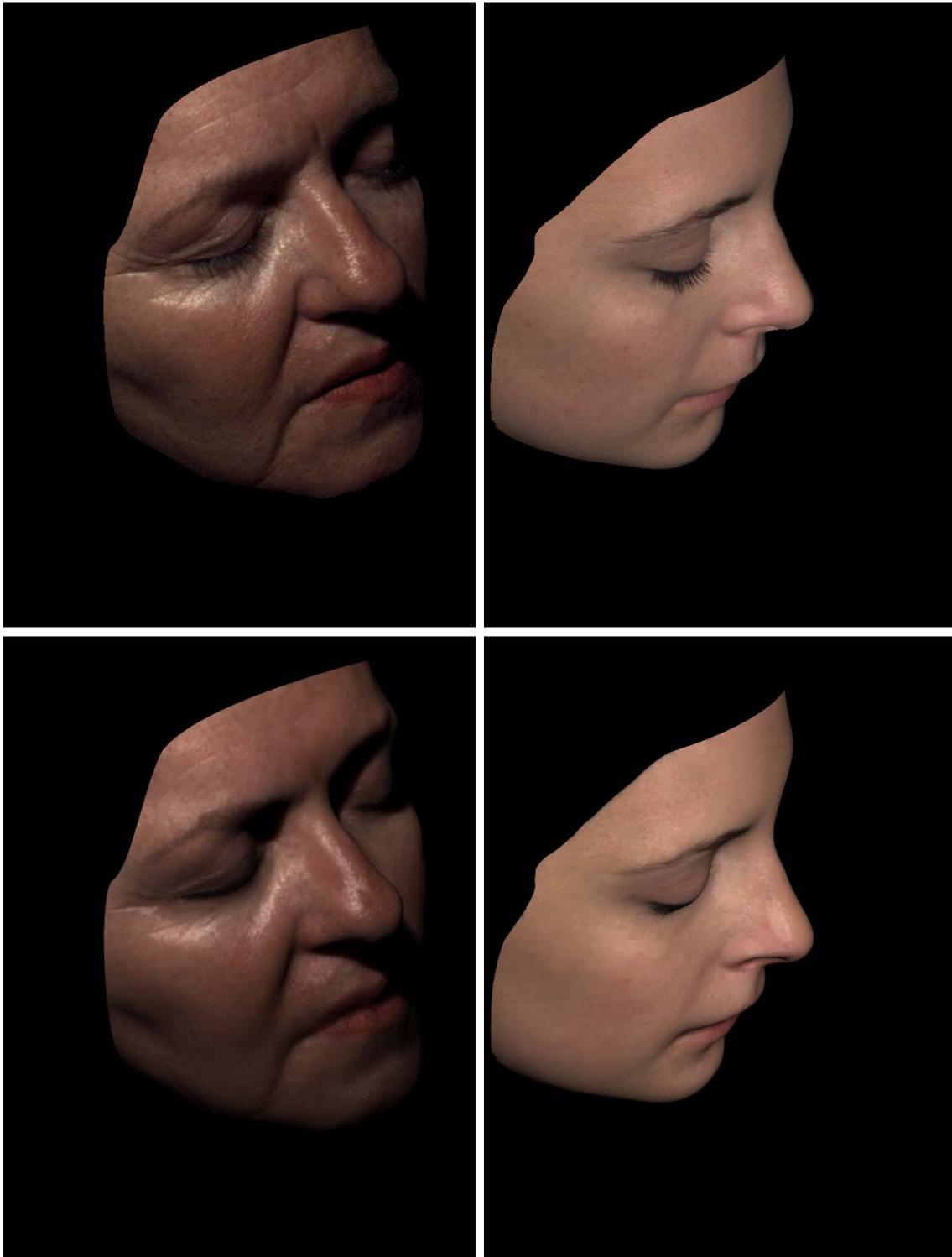


Figure 9.5, page 142 (III): Comparison of real photographs (top) to our model (bottom). All photographs were cropped according to the 3D model to remove distracting features.

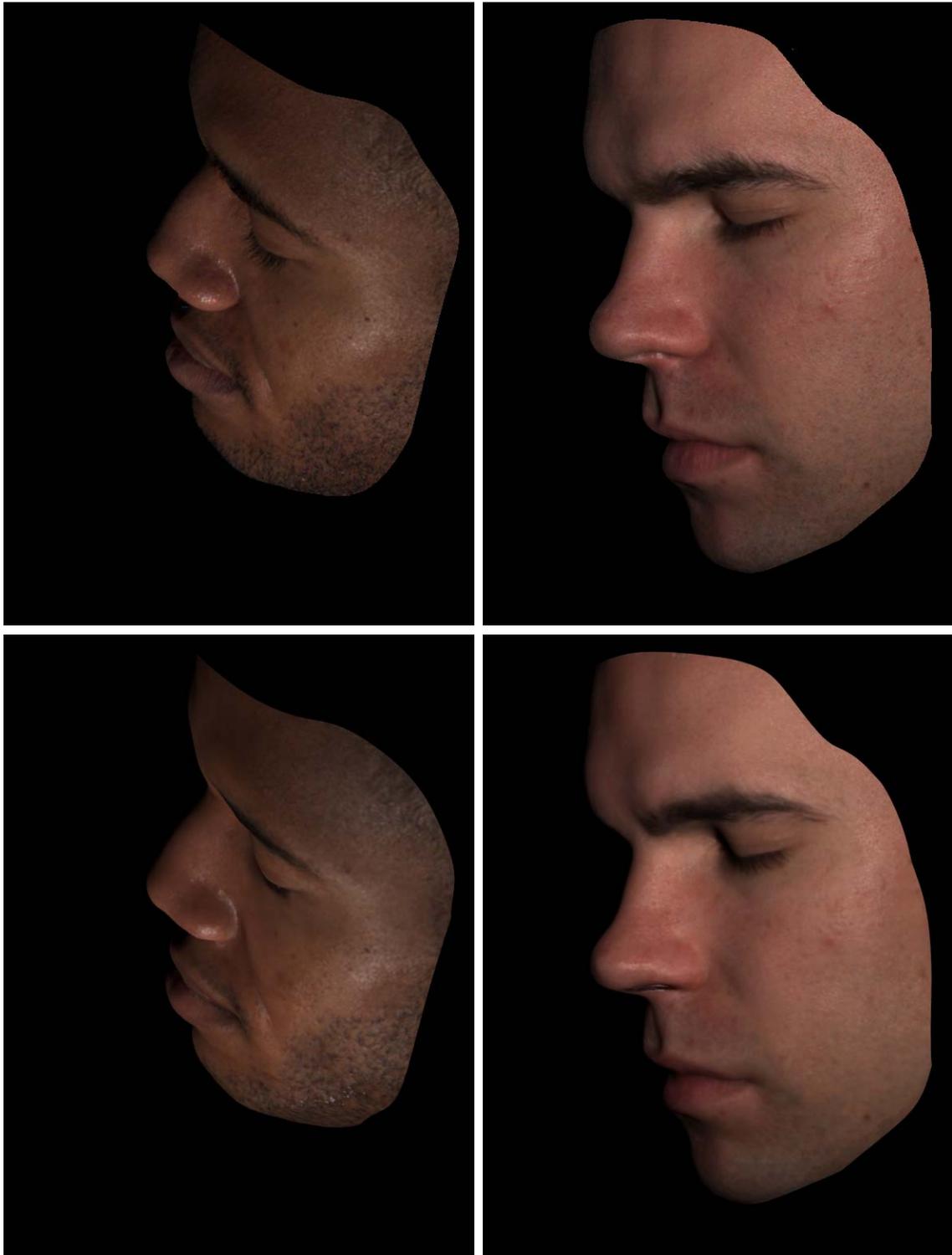


Figure 9.5, page 142 (IV): Comparison of real photographs (top) to our model (bottom). All photographs were cropped according to the 3D model to remove distracting features.



Figure 9.5, page 142 (V): Comparison of real photographs (top) to our model (bottom). All photographs were cropped according to the 3D model to remove distracting features.

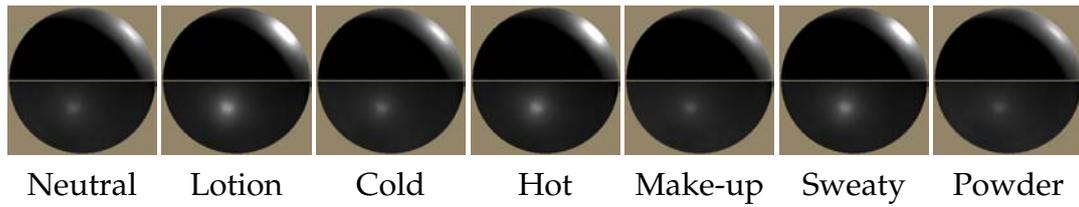


Figure 10.3, page 148: *Visualization of the average surface BRDF for dataset A. Top hemispheres: Back lighting from grazing angle. Bottom hemispheres: Front lighting from lower left.*

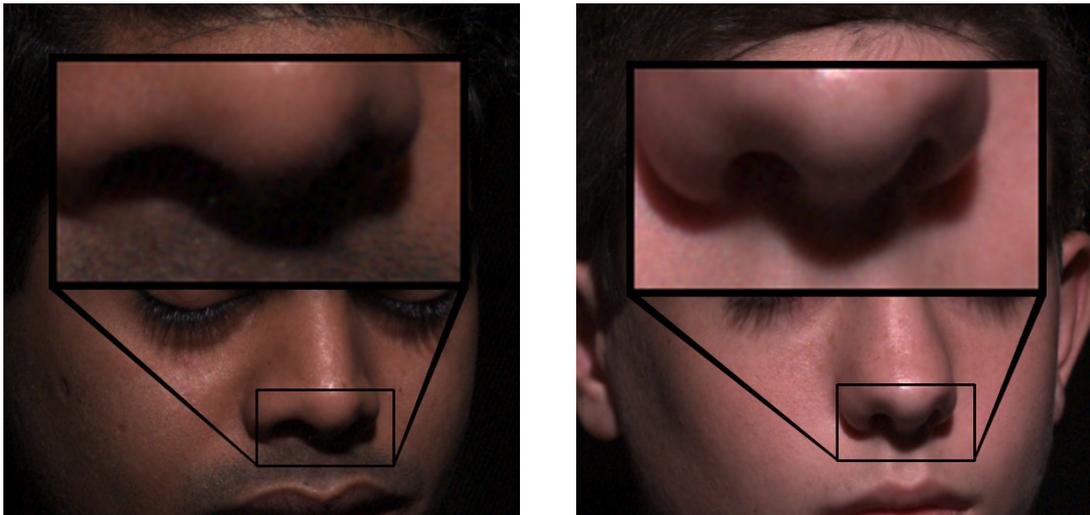


Figure 10.7, page 153: *Photographs of subjects with minimum (left) and maximum (right) translucency values in dataset B. The differences at shadow boundaries are subtle.*



Figure 10.8, page 153: *Renderings with minimum (left) and maximum (right) translucency.*



Figure 10.11, page 157: Face editing by adding trait vectors to different source faces. Top: Changing the surface BRDF using the acquired lotion trait. The example interpolates from “neutral” to a two-fold exaggeration of this trait. Bottom: Changing translucency, albedo histogram, and surface BRDF from skin type IV to II. (Both models use spatially subsurface scattering parameters and a data-driven BRDF model. No micro-normals are used.)

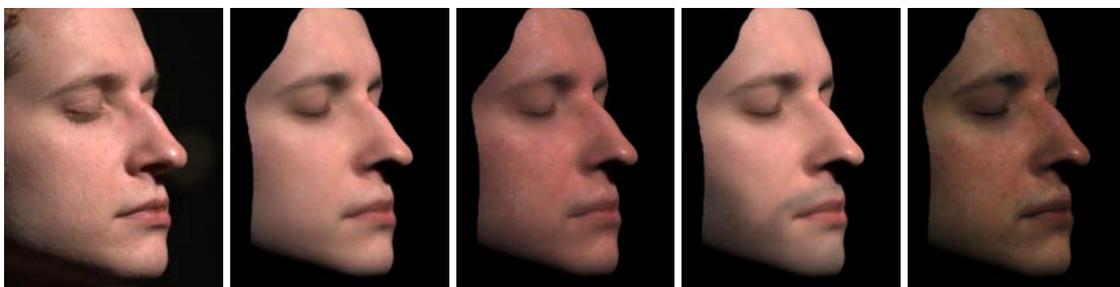


Figure 10.13, page 159: Albedo changes. From left to right: Real photograph; rendering; making the face sun-burnt; adding hair follicles in the beard area; making the skin type darker.



Figure 10.12, page 158: Face appearance transfer from source faces (upper row) using the statistics of a target to the results (lower row). Left: Transferring freckles using the method by Heeger and Bergen. Center: Transferring lotion by matching histograms of BRDF parameter textures. Right: Skin type transfer by matching histograms of the diffuse albedo texture. (All models use static diffusion, a modulation texture, and Torrance-Sparrow surface reflectance. The models contain micro-normals.)

Curriculum Vitae

Dipl.-Inform. Tim Alexander Weyrich



Personal Data

Sep. 26, 1974 Born in Gross-Gerau, Germany,
son of Christel and Prof. Dr. Dr. h.c. Wolf Weyrich.
Nationality German

Education

Jul. 10, 2006 Ph.D. defense.
Aug. 2001 – Apr. 2006 Research assistant and Ph.D. student at the Computer Graphics
Laboratory of the Swiss Federal Institute of Technology (ETH)
Zurich, Prof. Markus Gross.
Jun. 30, 2001 Diploma degree in Computer Science.
Oct. 1995 – Jun. 2001 Diploma Studies of Computer Science, Technical University of
Karlsruhe, Germany. Specialization: Knowledge-based Systems
and Computer Graphics.; Complementary studies: Biomedical
Engineering.
Jun. 15, 1994 Abitur from Heinrich-Suso-Gymnasium Constance, Germany.

Awards & Fellowships

Sep. 2006 – Aug. 2009 Postdoctoral Teaching Fellowship of the Council of Science &
Technology of Princeton University.

- Jan. 2000 Academic award “Deutscher Studienpreis” of the Körber Foundation, Germany. A third prize was awarded for the semester thesis “Evaluation and Optimization of Incisions In Soft Tissue”.
- 1991 / 1992 Reached the finals of the German Federal Computer Science Competition (Bundeswettbewerb Informatik).

Scientific Publications

- T. Weyrich, W. Matusik, H. Pfister, B. Bickel, C. Donner, C. Tu, J. McAndless, J. Lee, A. Ngan, H. W. Jensen, M. Gross. Analysis of Human Faces using a Measurement-Based Skin Reflectance Model. In *ACM Transactions on Graphics (TOG), ACM SIGGRAPH 2006 Papers*, pp. 1013–1024, Boston, MA, USA, July 2006.
- B. Bickel, T. Weyrich, W. Matusik, H. Pfister, C. Donner, C. Tu, J. McAndless, J. Lee, A. Ngan, H. W. Jensen, M. Gross. Processing And Editing of Faces Using A Measurements-Based Skin Reflectance Model. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Sketches*, p. 168, Boston, MA, USA, July 2006.
- Ch. Sigg, T. Weyrich, M. Botsch, M. Gross. GPU-Based Ray-Casting of Quadratic Surfaces. In *Proceedings of Eurographics Symposium on Point-Based Graphics*, pp. 59–65, July 29–30, 2006, Boston, MA.
- T. Weyrich, W. Matusik, H. Pfister, J. Lee, A. Ngan, H. W. Jensen, M. Gross. A Measurement-Based Skin Reflectance Model for Face Rendering and Editing. *Mitsubishi Electric Research Laboratories (MERL) Technical Report*, TR2005-071, July 2005.
- F. Sadlo, T. Weyrich, R. Peikert, M. Gross. A Practical Structured Light Acquisition System for Point-Based Geometry and Texture. In *Proceedings of Eurographics Symposium on Point-Based Graphics 2005*, pp. 89–98, Stony Brook, USA, June 21–22.
- T. Weyrich, H. Pfister, M. Gross. Rendering Deformable Surface Reflectance Fields. In *IEEE Transactions on Visualization and Computer Graphics*, 11(1), pp. 48–58, Jan–Feb 2005.
- D. Cotting, T. Weyrich, M. Pauly, M. Gross. Robust Watermarking of Point-Sampled Geometry. In *Proceedings of International Conference on Shape Modeling and Applications 2004 (SMI04)*, pp. 233–242, Genova, Italy, June 7–9, 2004.
- T. Weyrich, M. Pauly, R. Keiser, S. Heinzle, S. Scandella, M. Gross. Post-processing of Scanned 3D Surface Data. In *Proceedings of Eurographics Symposium on Point-Based Graphics 2004*, pp. 85–94, Zurich, Switzerland, Jun 2-4, 2004.
- B. Giesler, T. Salb, T. Weyrich, R. Dillmann. Using a Panoramic Camera for 3D Head Tracking in an AR Environment. In *Proceedings of 9th IEEE Int. Conference on Mechatronics and Machine Vision in Practice*, Chiang Mai, Thailand, Sep 10–12, 2002.
- T. Gockel, T. Salb, T. Weyrich, R. Dillmann. Interactive simulation of teeth cleaning. *CARS 2001, 7th Computed Maxillofacial Imaging Congress (CMI)*, June 2001, Berlin.
- T. Salb, T. Weyrich, O. Burgert, R. Dillmann. Haptic mediation of risk potentials over a soft tissue data set. *The First PHANToM Users Research Symposium*, May 1999, Heidelberg.
- T. Salb, T. Weyrich, R. Dillmann. Preoperative Planning and Training Simulation for Risk Reducing Surgery. *International Training and Simulation Conference (ITEC)*, April 1999, The Hague, The Netherlands. (Best Paper Award.)

List of Figures

2.1	Sample basis BRDFs	11
2.2	The dipole model	12
2.3	Surfel definition	18
3.1	Typical artifacts of raw scanner data	22
3.2	Bridging the gap between acquisition and modeling	22
3.3	Tools of different complexity	24
3.4	The volumetric brush	24
3.5	Outlier classification	27
3.6	The plane fit criterion	28
3.7	The miniball criterion	29
3.8	Nearest-neighbor graph	29
3.9	Three different outlier classifiers	31
3.10	The nearest-neighbor criterion	31
3.11	The MLS smoother tool	32
3.12	Selective noise removal using the MLS smoother	32
3.13	Volumetric diffusion	34
3.14	Manual hole filling using the MLS spray can tool	35
3.15	Removal of an undesired bump	36
3.16	Robustness of the volumetric diffusion tool	37
3.17	Egyptian sculpture scanned at the British Museum	37

List of Figures

4.1	Example reflectance image	41
4.2	Relighting using a reflectance image	41
4.3	Rendering a deformed surface reflectance field requires switching between reference frames	42
4.4	Reflectance queries must be mapped into acquisition space	43
4.5	A viewing ray in object space intersects the object	44
4.6	The bi-directional reflection distribution function (BRDF)	45
4.7	Local coordinate systems	48
4.8	Close-ups of a surfel geometry under deformation	51
4.9	Viewing ray divergence due to back-projection	53
4.10	The presented models	54
4.11	View extrapolation	55
4.12	Preservation of material properties	56
4.13	Deformation of refractive objects	57
4.14	Surfel geometry under large deformations	58
4.15	Viewing ray divergence	59
4.16	Selected frames of a surface reflectance field animation	60
4.17	Face datasets obtained from silhouette fitting	61
5.1	Face images under differently spaced light sources	67
5.2	Color Kinetics iColor™ light source	67
5.3	Dome light source and camera type	68
5.4	Acquisition dome simulator	70
5.5	3D reconstruction testbed	71
5.6	Second expansion stage of the scanner	72
5.7	The final dome construction	73
5.8	Close-ups of the dome construction	74
5.9	Raw reflectance images	75
5.10	Pseudo random order of light sources for motion compensation	77
5.11	Flow field relations	78
5.12	Flow masking	78
5.13	Exemplary head trajectories	80
5.14	Motion compensation	81
6.1	Contact-less measurement apparatus	87
6.2	Experimental subsurface scattering measurement devices	89
6.3	Subsurface scattering measurement device	90
6.4	BSSRDF sensor fiber layout	91
6.5	The subsurface scattering probe	91
6.6	Usage of the BSSRDF sensor	92
6.7	BSSRDF gun calibration procedure	93
6.8	Subsurface scattering model fit	95
6.9	Spatial translucency variance measurements	96
6.10	Anisotropy measurements for subsurface scattering in human skin	97

7.1	Overview over the processing pipeline	100
7.2	Two alternative geometry processing pipelines	101
7.3	A sample skin mask	104
7.4	Cleaned 3QTech geometry data	104
7.5	Desbrun parameterization of a face model	105
7.6	Comparison of face model parameterizations	106
7.7	Depth discontinuity images	108
7.8	Visualization of two skin lumitexels for different viewpoints	109
7.9	Helmholtz reciprocity setup	111
7.10	Approximate Helmholtz reciprocity	112
7.11	Comparison of normal maps	114
7.12	Geometry refinement work flow	116
7.13	Geometry refinement result	118
8.1	Physiology of skin	120
8.2	Model constituents	121
8.3	Micro-normals	123
8.4	Block diagram of the model parameter estimation process	124
8.5	Synthetic images of a face using NMF bases of different dimensionality	130
8.6	BRDF Models	132
9.1	Comparison of real photographs of four subjects with fits of our model	137
9.2	Components of our model using a data-driven BRDF	138
9.3	Components of the model using the modulation texture and a Torrance-Sparrow BRDF	139
9.4	Two approaches to model diffuse reflectance	140
9.5	Comparison of real photographs to our model	142
10.1	Encoding of face regions	146
10.2	Representative subjects in dataset B	147
10.3	Visualization of the average surface BRDF for dataset A	148
10.4	Reconstruction error as a function of the number of NMF basis BRDFs	149
10.5	Variation of Torrance-Sparrow coefficients	151
10.6	Canonical correlation analysis of the Torrance-Sparrow BRDF parameters	152
10.7	Photographs of subjects with minimum and maximum translucency values	153
10.8	Synthetic images with minimum and maximum translucency values	153
10.9	Plot of absorption and scattering parameters	154
10.10	Albedo histograms	155
10.11	Face editing by adding trait vectors to different source faces	157
10.12	Face appearance transfer	158

List of Figures

10.13 Albedo changes 159

List of Tables

4.1	Local impostor parameterization allows for the determination of the look-up function \mathcal{L} during rendering	47
4.2	A shading operation represents a basic reflectance field look-up . . .	53
10.1	The Fitzpatrick skin type system	146
10.2	Age distribution	146
10.3	Mean and variance of μ_{tr} for dataset B.	151
A.1	Radiometric quantities and their definitions	164

Bibliography

- [ABCO⁺01] Marc Alexa, Johannes Behr, Daniel Cohen-Or, Shachar Fleishman, David Levin, and Cláudio T. Silva. Point set surfaces. In *Proceedings of IEEE Visualization*, pages 21–28, San Diego, CA, October 2001.
- [ACDL00] N. Amenta, S. Choi, T. K. Dey, and N. Leekha. A simple algorithm for homeomorphic surface reconstruction. In *Proceedings of 16th Annual Symposium on Computational Geometry*, pages 213–222, 2000.
- [Alt89] Simon L. Altman. Hamilton, grassmann, rodrigues, and the quaternion scandal—what went wrong with one of the major mathematical discoveries of the nineteenth century? *A Mathematical Association of America journal*, December 1989.
- [AM93] S. Arya and D. M. Mount. Algorithms for fast vector quantization. In *Data Compression Conference*, pages 381–390. IEEE Computer Society Press, 1993.
- [AMD01] Elli Angelopoulou, Rana Molana, and Kostas Daniilidis. Multispectral skin color modeling. In *CVPR (2)*, pages 635–642. IEEE Computer Society, 2001.

Bibliography

- [Ami02] Isaac Amidror. Scattered data interpolation methods for electronic imaging systems: a survey. *Journal of Electronic Imaging*, 11(2):157–176, April 2002.
- [BBM⁺01] Chris Buehler, Michael Bosse, Leonard McMillan, Steven J. Gortler, and Michael F. Cohen. Unstructured lumigraph rendering. In *Computer Graphics, SIGGRAPH 2001 Proceedings*, pages 425–432, Los Angeles, CA, 2001.
- [Ben85] J. L. Bentley. Multidimensional binary search trees used for associative searching. *Communications of the ACM*, pages 509–517, 1985.
- [Bic06] Bernd Bickel. A measurement-based skin reflectance model. Master's thesis, Eidgenössisch Technische Hochschule Zürich (ETH), Switzerland, and Mitsubishi Electric Research Laboratories (MERL), USA, March 2006.
- [BL03] George Borshukov and J. P. Lewis. Realistic human face rendering for "the matrix reloaded". In *ACM SIGGRAPH 2003 Conference Abstracts and Applications (Sketch)*, page 1, New York, NY, USA, 2003. ACM Press.
- [Bli77] James F. Blinn. Models of light reflection for computer synthesized pictures. In *SIGGRAPH '77: Proceedings of the 4th annual conference on Computer graphics and interactive techniques*, pages 192–198, New York, NY, USA, 1977. ACM Press.
- [BP01] Svetlana Barsky and Maria Petrou. Colour photometric stereo: Simultaneous reconstruction of local gradient and colour of rough textured surfaces. In *ICCV*, pages 600–605, 2001.
- [BSBK02] Mario Botsch, Stephan Steinberg, Stephan Bischoff, and Leif Kobbelt. OpenMesh — a generic and efficient polygon mesh data structure. In *OpenSG Symposium 2002*, Darmstadt, Germany, January 2002.
- [BV99] Volker Blanz and Thomas Vetter. A morphable model for the synthesis of 3d faces. In *SIGGRAPH '99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 187–194, New York, NY, USA, 1999. ACM Press/Addison-Wesley Publishing Co.
- [CBC⁺01] J. C. Carr, R. K. Beatson, J.B. Cherrie, T. J. Mitchell, W. R. Fright, B. C. McCallum, and T. R. Evans. Reconstruction and representation of 3D objects with radial basis functions. In *Computer Graphics, SIGGRAPH 2001 Proceedings*, pages 67–76, Los Angeles, CA, August 2001.

- [CBCG02] W-C. Chen, J-Y. Bouguet, M. H. Chu, and R. Grzeszczuk. Light Field Mapping: Efficient Representation and Hardware Rendering of Surface Light Fields. *ACM Transactions on Graphics (SIGGRAPH 2002)*, 21(3):447–456, July 2002.
- [CCSS01] Yung-Yu Chuang, Brian Curless, David H. Salesin, and Richard Szeliski. A bayesian approach to digital matting. In *Proceedings of IEEE CVPR 2001*, volume 2, pages 264–271. IEEE Computer Society, December 2001.
- [CD02] Oana G. Cula and Kristin J. Dana. Image-based skin analysis. In *Texture 2002*, The Second International Workshop on Texture Analysis and Synthesis, pages 35–42, Copenhagen, Denmark, June 2002.
- [CDMR04] O. Cula, K. Dana, F. Murphy, and B. Rao. Bidirectional imaging and modeling of skin texture. *IEEE Transactions on Biomedical Engineering*, 51(12):2148–2159, December 2004.
- [CDMR05] O. Cula, K. Dana, F. Murphy, and B. Rao. Skin texture modeling. *International Journal of Computer Vision*, 62(1–2):97–119, April–May 2005.
- [CT82] Robert L. Cook and Kenneth E. Torrance. A reflection model for computer graphics. *ACM Transactions On Graphics*, 1(1):7–24, January 1982.
- [CWPG04] Daniel Cotting, Tim Weyrich, Mark Pauly, and Markus Gross. Robust watermarking of point-sampled geometry. In *Proceedings of International Conference on Shape Modeling and Applications 2004*, pages 233–242, Genova, Italy, June 2004. IEEE Computer Society Press.
- [Dan92] James L. Dannemiller. Spectral reflectance of natural objects: how many basis functions are necessary? *Journal of the Optical Society of America A*, 9(4):507–515, April 1992.
- [DEL⁺99] Julie Dorsey, Alan Edelman, Justin Legakis, Henrik Wann Jensen, and Hans K ohling Pedersen. Modeling and rendering of weathered stone. In *Computer Graphics, SIGGRAPH 99 Proceedings*, pages 225–234, Los Angeles, CA, August 1999.
- [DG01] Tamal K. Dey and Joachim Giesen. Detecting undersampling in surface reconstruction. In *Proceedings of 17th Annual Symposium on Computational Geometry*, pages 257–263, 2001.
- [DHT⁺00] Paul Debevec, Tim Hawkins, Chris Tchou, Haarm-Pieter Duiker, Westley Sarokin, and Mark Sagar. Acquiring the reflectance field of a human face. In *Computer Graphics, SIGGRAPH 2000 Proceedings*, pages 145–156, July 2000.

Bibliography

- [DJ05] Craig Donner and Henrik Wann Jensen. Light diffusion in multi-layered translucent materials. *ACM Transactions on Graphics (SIGGRAPH 2005)*, 24(3):1032–1039, 2005.
- [DM97] Paul Debevec and Jitendra Malik. Recovering high dynamic range radiance maps from photographs. In *Computer Graphics, SIGGRAPH 97 Proceedings*, pages 369–378, Los Angeles, CA, 1997.
- [DMA02] M. Desbrun, M. Meyer, and P. Alliez. Meshes and parametrization: Intrinsic parameterizations of surface meshes. *Computer Graphics Forum*, 21(3), September 2002.
- [DMGL02] James Davis, Steven R. Marschner, Matt Garr, and Marc Levoy. Filling holes in complex surfaces using volumetric diffusion. In *First International Symposium on 3D Data Processing, Visualization, and Transmission*, pages 428–861, June 2002.
- [DMSB99] Mathieu Desbrun, Mark Meyer, Peter Schröder, and Alan H. Barr. Implicit fairing of irregular meshes using diffusion and curvature flow. In *Computer Graphics, SIGGRAPH 99 Proceedings*, pages 317–324, Los Angeles, CA, August 1999.
- [DTM96] Paul Debevec, Camillo J. Taylor, and Jitendra Malik. Modeling and rendering architecture from photographs: A hybrid geometry- and image-based approach. In *Computer Graphics, SIGGRAPH 96 Proceedings*, pages 11–20, August 1996.
- [DvGNK99] Kristin J. Dana, Bram van Ginneken, Shree K. Nayar, and Jan J. Koenderink. Reflectance and texture of real world surfaces. *ACM Transactions on Graphics*, 1(18):1–34, 1999.
- [DWT⁺02] Paul Debevec, Andreas Wenger, Chris Tchou, Andrew Gardner, Jamie Waese, and Tim Hawkins. A lighting reproduction approach to live-action compositing. *ACM Transactions on Graphics (SIGGRAPH 2002)*, 21(3):547–556, July 2002.
- [DYB98] Paul Debevec, Yizhou Yu, and George Boshokov. Efficient view-dependent image-based rendering with projective texture-mapping. In *Proceedings of the 9th Eurographics Workshop on Rendering*, pages 105–116, Vienna, Austria, June 1998.
- [EVNT78] G. Eason, A. Veitch, R. Nisbet, and F. Turnbull. The theory of the backscattering of light by blood. *Journal of Physics D: Applied Physics*, 11:1463–1479, 1978.
- [FBF77] J. H. Friedman, J. L. Bentley, and R. A. Finkel. An algorithm for finding best matches in logarithmic expected time. *ACM Transactions on Mathematical Software*, pages 209–226, 1977.

- [Fee04] Catherine Feeny. Servers, spydercams and 'Spider-Man 2'. In *VFX Pro*, http://www.uemedia.net/CPC/vfxpro/printer_9050.shtml, July 2004.
- [FGCS05] Christian Fuchs, Michael Goesele, Tongbo Chen, and Hans-Peter Seidel. An empirical model for heterogeneous translucent objects. Research Report MPI-I-2005-4-006, Max-Planck-Institut für Informatik, Stuhlsatzenhausweg 85, 66123 Saarbrücken, Germany, May 2005.
- [Fit88] T. Fitzpatrick. The validity and practicality of sun-reactive skin types I through VI. *Arch. Dermatology*, 124:869–871, 1988.
- [FKIS02] Ryo Furukawa, Hiroshi Kawasaki, Katsushi Ikeuchi, and Masao Sakauchi. Appearance based object modeling using texture database: Acquisition, compression and rendering. In *Proceedings of the 13th Eurographics Workshop on Rendering*, Pisa, Italy, June 2002.
- [FL00] Brian V. Funt and Benjamin C. Lewis. Diagonal versus affine transformations for color correction. *Journal of the Optical Society of America A*, 17(11):2108–2112, November 2000.
- [FLS05] Martin Fuchs, Hendrik P. A. Lensch, and Hans-Peter Seidel. Reflectance from images: A model-based approach for human faces. *IEEE Transactions on Visualization and Computer Graphics*, 11(3):296–305, 2005. Member-Volker Blanz.
- [FP03] Sarah F. Frisken and Ronald N. Perry. Simple and efficient traversal methods for quadtrees and octrees. *Journal of Graphics Tools*, 7(3), May 2003.
- [FPW92] Thomas J. Farrell, Michael S. Patterson, and Brian C. Wilson. A diffusion theory model of spatially resolved, steady-state diffuse reflectance for the noninvasive determination of tissue optical properties in vivo. *Medical Physics*, 19(4):879–888, 1992.
- [GBK99] Athinodoros S. Georghiades, Peter N. Belhumeur, and David J. Kriegman. Illumination-based image synthesis: Creating novel images of human faces under differing pose and lighting. In *Proceedings of the IEEE Workshop on Multi-View Modeling and Analysis of Visual Scenes (MVIEW '99)*, pages 47–54, Washington, DC, USA, April 1999. IEEE Computer Society.
- [Geo03] Athinodoros S. Georghiades. Recovering 3-D shape and reflectance from a small number of photographs. In *Proceedings of the 14th Eurographics workshop on Rendering*, pages 230–240, Aire-la-Ville, Switzerland, Switzerland, 2003. Eurographics Association.

Bibliography

- [GGSC96] Steven J. Gortler, Radek Grzeszczuk, Richard Szeliski, and Michael F. Cohen. The lumigraph. In *Computer Graphics, SIGGRAPH 96 Proceedings*, pages 43–54, New Orleans, LS, August 1996.
- [GLL⁺04] Michael Goesele, Hendrik P. A. Lensch, Jochen Lang, Christian Fuchs, and Hans-Peter Seidel. DISCO—Acquisition of translucent objects. *ACM Transactions on Graphics (SIGGRAPH 2004)*, 24(3):835–844, 2004.
- [GP06] Markus Gross and Hanspeter Pfister, editors. *Point-Based Graphics*. Elsevier, to appear in 2006.
- [Gra03] Graphite. <http://www.loria.fr/~levy/Graphite/index.html>, 2003.
- [HB95] David J. Heeger and James R. Bergen. Pyramid-based texture analysis/synthesis. In *Proceedings of SIGGRAPH 95, Computer Graphics Proceedings, Annual Conference Series*, pages 229–238, August 1995.
- [HCD01] Tim Hawkins, Jonathan Cohen, and Paul Debevec. A photometric approach to digitizing cultural artifacts. In *2nd International Symposium on Virtual Reality, Archaeology, and Cultural Heritage, Glyfada, Greece*, November 2001.
- [Hec89] Paul S. Heckbert. Fundamentals of texture mapping and image warping. Master’s thesis, University of California at Berkeley, Department of Electrical Engineering and Computer Science, June 17 1989.
- [Hei04] Simon Heinzle. Development of a pointshop3d-plugin for post-processing of 3d-scans. Semester thesis, Eidgenössisch Technische Hochschule Zürich (ETH), Switzerland, 2004.
- [Hér03] Christophe Héry. Face cloning at ILM. In *SIGGRAPH 2003 Course ‘Digital Face Cloning’*, 2003.
- [Hér05] Christophe Héry. Face cloning at ILM. In *SIGGRAPH 2005 Course ‘Digital Face Cloning’*, 2005.
- [HK93] Pat Hanrahan and Wolfgang Krueger. Reflection from layered surfaces due to subsurface scattering. In *Computer Graphics, SIGGRAPH 93 Proceedings*, pages 165–174, Anaheim, CA, August 1993.
- [HS99] Wolfgang Heidrich and Hans-Peter Seidel. Realistic, hardware-accelerated shading and lighting. In *Computer Graphics, SIGGRAPH 99 Proceedings*, pages 171–178, Los Angeles, CA, August 1999.

- [HTSG91] Xiao D. He, Kenneth E. Torrance, François X. Sillion, and Donald P. Greenberg. A comprehensive physical model for light reflection. *Computer Graphics*, 25(Annual Conference Series):175–186, 1991.
- [Hub02] Sven Hubschmid. 3D reconstruction by structured light. Semester thesis, Eidgenössisch Technische Hochschule Zürich (ETH), Switzerland, 2002.
- [HWT⁺04] Tim Hawkins, Andreas Wenger, Chris Tchou, Andrew Gardner, Fredrik Göransson, and Paul Debevec. Animatable facial reflectance fields. In *Rendering Techniques '04 (Proceedings of the Second Eurographics Symposium on Rendering)*, pages 309–320, Norrköping, Sweden, June 2004.
- [INN05] Takanori Igarashi, Ko Nishino, and Shree K. Nayar. The appearance of human skin. Technical report, Department of Computer Science, Columbia University CUCS-024-05, June 2005.
- [INS] The institute for study and implementation of graphical heritage techniques (INSIGHT). <http://www.insightdigital.org/>.
- [Jai89] Anil K. Jain. *Fundamentals of Digital Image Processing*. Prentice Hall, Englewood Cliffs, NJ, 1989.
- [JB02] Henrik Wann Jensen and Juan Buhler. A rapid hierarchical rendering technique for translucent materials. In *Computer Graphics, SIGGRAPH 2002 Proceedings*, pages 576–581, Los Angeles, CA, July 2002.
- [JD99] Henrik Wann Jensen and Julie Dorsey. Rendering of wet materials. In *Rendering Techniques '99*, pages 273–282, Granada, Spain, June 1999. Springer, Wien.
- [JDJ06] Neel Joshi, Craig Donner, and Henrik Wann Jensen. Noninvasive measurement of scattering anisotropy in turbid materials by nonnormal incident illumination. *Optics Letters*, 31:936–938, 2006.
- [JGB⁺05] Andrew Jones, Andrew Gardner, Mark Bolas, Ian McDowall, and Paul Debevec. Performance geometry capture for spatially varying relighting. *ACM SIGGRAPH 2005 Conference Abstracts and Applications (Sketch)*, 2005.
- [JLW05] Zhongping Ji, Ligang Liu, and Guojin Wang. A global laplacian smoothing approach with feature preservation. In *Proceedings of Ninth International Conference on Computer Aided Design and Computer Graphics (CAD-CG'05)*, pages 269–274, 2005.

Bibliography

- [JMLH01] Henrik Wann Jensen, Steven R. Marschner, Marc Levoy, and Pat Hanrahan. A practical model for subsurface light transport. In *Computer Graphics, SIGGRAPH 2001 Proceedings*, pages 511–518, Los Angeles, CA, August 2001.
- [KB04a] Leif Kobbelt and Mario Botsch. A survey of point-based techniques in computer graphics. *Computers & Graphics*, 28:801–814, 2004.
- [KB04b] Aravind Krishnaswamy and Gladimir V. G. Baranoski. A biophysically-based spectral model of light interaction with human skin. *Computer Graphics Forum*, 23(3):331–340, September 2004.
- [KBMK01] M. Koudelka, P. Belhumeur, S. Magda, and D. Kriegman. Image-based modeling and rendering of surfaces with arbitrary BRDFs. In *Proc. of Computer Vision and Pattern Recognition*, page in press, Kauai, HI, December 2001.
- [KM99] Jan Kautz and Michael D. McCool. Interactive rendering with arbitrary BRDFs using separable approximations. In *Rendering Techniques '99 (Proceedings of the Tenth Eurographics Workshop on Rendering)*, pages 281–292, New York, NY, June 1999. Springer Wien.
- [KP03] J. Koenderink and S. Pont. The secret of velvety skin. *Machine Vision and Application*, 14:260–268, 2003. Special Issue on Human Modeling, Analysis, and Synthesis.
- [LC87] W. E. Lorensen and H. E. Cline. Marching cubes: A high resolution 3D surface construction algorithm. In *Computer Graphics, SIGGRAPH '87 Proceedings*, pages 163–169, San Francisco, CA, October 1987.
- [Len04] Hendrik P. A. Lensch. *Efficient, Image-Based Appearance Acquisition of Real-World Objects*. Cuvillier Verlag, 2004. PhD thesis.
- [Lev01] David Levin. Mesh-independent surface interpolation. In *Advances in Computational Mathematics*, 2001.
- [LFTG97] Eric P. F. Lafortune, Sing-Choong Foo, Kenneth E. Torrance, and Donald P. Greenberg. Non-linear approximation of reflectance functions. In *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 117–126, New York, NY, USA, 1997. ACM Press/Addison-Wesley Publishing Co.
- [LH96] Marc Levoy and Pat Hanrahan. Light field rendering. In *Computer Graphics, SIGGRAPH 96 Proceedings*, pages 31–42, New Orleans, LS, August 1996.

- [LK81] B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proceedings of 7th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 674–679, 1981.
- [LKG⁺01] Hendrik P. A. Lensch, Jan Kautz, Michael Goesele, Wolfgang Heidrich, and Hans-Peter Seidel. Image-based reconstruction of spatially varying materials. In *Proceedings of the 12th Eurographics Workshop on Rendering*, pages 104–115, June 2001.
- [LMPM03] Jinho Lee, Baback Moghaddam, Hanspeter Pfister, and Raghu Machiraju. Silhouette-based 3d face shape recovery. In *Proceedings of Graphics Interface*, pages 21–30, 2003.
- [LMPM04] Jinho Lee, Baback Moghaddam, Hanspeter Pfister, and Raghu Machiraju. Finding optimal views for 3D face shape modeling. In *Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FGRS04)*, pages 31–36. IEEE Computer Society, May 2004.
- [Loo87] Charles Loop. Smooth subdivision surfaces based on triangles. Master’s thesis, Department of Mathematics, University of Utah, August 1987.
- [LPT05] Hongsong Li, Fabio Pellacini, and Kenneth E. Torrance. A hybrid monte carlo method for accurate and efficient subsurface scattering. In *Proceedings of the 15th Eurographics Symposium on Rendering*, pages 283–290. Eurographics Association, 2005.
- [LRR04] Jason Lawrence, Szymon Rusinkiewicz, and Ravi Ramamoorthi. Efficient BRDF importance sampling using a factored representation. *ACM Transactions on Graphics (SIGGRAPH 2004)*, 23(3):496–505, August 2004.
- [MAA01] Michael D. McCool, Jason Ang, and Anis Ahmad. Homomorphic factorization of BRDFs for high-performance rendering. *Computer Graphics*, 35(Annual Conference Series):171–178, 2001.
- [MBR⁺00] Wojciech Matusik, Chris Buehler, Ramesh Raskar, Steven J. Gortler, and Leonard McMillan. Image-based visual hulls. In *Computer Graphics, SIGGRAPH 2000 Proceedings*, pages 369–374, Los Angeles, CA, July 2000.
- [MGR00] Stephen R. Marschner, Brian K. Guenter, and Sashi Raghupathy. Modeling and rendering for realistic facial animation. In *Proceedings of the 11th Eurographics Workshop on Rendering*, pages 231–242, London, UK, June 2000. Springer-Verlag.

Bibliography

- [MGW01] Thomas Malzbender, Dan Gelb, and Hans Wolters. Polynomial texture maps. In *Computer Graphics, SIGGRAPH 2001 Proceedings*, pages 519–528, Los Angeles, CA, 2001.
- [MPBM03] Wojciech Matusik, Hanspeter Pfister, Matthew Brand, and Leonard McMillan. A data-driven reflectance model. *ACM Transactions on Graphics (SIGGRAPH 2003)*, 22(3):759–770, July 2003.
- [MPN⁺02] Wojciech Matusik, Hanspeter Pfister, Addy Ngan, Paul A. Beardsley, Remo Ziegler, and Leonard McMillan. Image-based 3D photography using opacity hulls. *ACM Transactions on Graphics (SIGGRAPH 2002)*, 21(3):427–437, July 2002.
- [MPZ⁺02] Wojciech Matusik, Hanspeter Pfister, Remo Ziegler, Addy Ngan, and Leonard McMillan. Acquisition and rendering of transparent and refractive objects. In *Proceedings of the 13th Eurographics Workshop on Rendering*, Pisa, Italy, June 2002.
- [MRP98] Gavin S. P. Miller, Steven M. Rubin, and Dulce B. Ponceleon. Lazy decompression of surface light fields for precomputed global illumination. In *Proceedings of the 9th Eurographics Workshop on Rendering*, pages 281–292, Vienna, Austria, June 1998.
- [MTB⁺05] Ankit Mohan, Jack Tumblin, Bobby Bodenheimer, Cindy Grimm, and Reynold Bailey. Table-top computed lighting for practical digital photography. In *Rendering Techniques '05 (Proceedings of the Third Eurographics Symposium on Rendering)*, pages 165–172, Constance, Germany, June 2005.
- [MWL⁺99] Steven R. Marschner, Steven H. Westin, Eric P. F. Lafortune, Kenneth E. Torrance, and Donald P. Greenberg. Image-based BRDF measurement including human skin. In *Proceedings of the 10th Eurographics Workshop on Rendering*, pages 139–152, Granada, Spain, June 1999.
- [MWLT00] Steven R. Marschner, Steven H. Westin, Eric P. F. Lafortune, and Kenneth E. Torrance. Image-based measurement of the Bidirectional Reflectance Distribution Function. *Applied Optics*, 39(16):2592–2600, June 2000.
- [MZD05] Wojciech Matusik, Matthias Zwicker, and Frédo Durand. Texture design using a simplicial complex of morphable textures. *ACM Transactions on Graphics*, 24(3):787–794, 2005.
- [MZKB05] Satya P. Mallick, Todd E. Zickler, David J. Kriegman, and Peter N. Belhumeur. Beyond lambert: Reconstructing specular surfaces using color. In *CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* —

- Volume 2, pages 619–626, Washington, DC, USA, 2005. IEEE Computer Society.
- [NDM05] Addy Ngan, Frédo Durand, and Wojciech Matusik. Experimental analysis of brdf models. In *Proceedings of the 15th Eurographics Symposium on Rendering*, pages 117–226. Eurographics Association, 2005.
- [Nef02] Marco Nef. Development of a simulation system of a multi-camera based face scanner. Diploma thesis, Eidgenössisch Technische Hochschule Zürich (ETH), Switzerland, March 2002.
- [NHE⁺00] Stephan Nickell, Marcus Hermann, Matthias Essenpreis, Thomas J. Farrell, Uwe Krämer, and Michael S. Patterson. Anisotropy of light propagation in human skin. *Phys. Med. Biol.*, 45:2873–2886, 2000.
- [NL01] Carmen So-Ling Ng and Ling Li. A multi-layered reflection model of natural human skin. In *CGI '01: Computer Graphics International 2001*, pages 249–256, Washington, DC, USA, 2001. IEEE Computer Society.
- [NRDR05] Diego Nehab, Szymon Rusinkiewicz, James Davis, and Ravi Ramamoorthi. Efficiently combining positions and normals for precise 3d geometry. *ACM Transactions on Graphics (SIGGRAPH 2005)*, 24(3):536–543, 2005.
- [NRH⁺77] F. E. Nicodemus, J. C. Richmond, J. J. Hsia, I. W. Ginsberg, and T. Limperis. Geometric considerations and nomenclature for reflectance. Monograph 161, National Bureau of Standards (US), October 1977.
- [NSI99a] Ko Nishino, Yoichi Sato, and Katsushi Ikeuchi. Appearance compression and synthesis based on 3D model for mixed reality. In *Proceedings of IEEE ICCV '99*, pages 38–45, September 1999.
- [NSI99b] Ko Nishino, Yoichi Sato, and Katsushi Ikeuchi. Eigen-texture method: Appearance compression based on 3D model. In *Proc. of Computer Vision and Pattern Recognition*, pages 618–624, June 1999.
- [OBA⁺03] Yutaka Ohtake, Alexander G. Belyaev, Marc Alexa, Greg Turk, and Hans-Peter Seidel. Multi-level partition of unity implicits. In *Computer Graphics, Siggraph 2003 Proceedings*, pages 463–470, San Diego, Ca, July 2003.
- [OCV] Open source computer vision library. <http://opencvlibrary.sourceforge.net/>.
- [OS88] S. Osher and J. A. Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on hamilton–jacobi formulations. *Journal of Computational Physics*, 79:12–49, 1988.

Bibliography

- [Pau03] M. Pauly. *Point Primitives for Interactive Modeling and Processing of 3D Geometry*. PhD thesis, Department of Computer Science, ETH Zurich, 2003.
- [PCD⁺97] K. Pulli, M. Cohen, T. Duchamp, H. Hoppe, L. Shapiro, and W. Stuetzle. View-based rendering: Visualizing real objects from scanned range and color data. In *Eurographics Rendering Workshop 1997*, pages 23–34, June 1997.
- [PGK02] Mark Pauly, Markus Gross, and Leif P. Kobbelt. Efficient simplification of point-sampled surfaces. In *Proceedings of the conference on Visualization '02*, pages 163–170. IEEE Computer Society, 2002.
- [PHL⁺98] Frédéric Pighin, Jamie Hecker, Dani Lischinski, Richard Szeliski, and David H. Salesin. Synthesizing realistic facial expressions from photographs. In *Computer Graphics*, volume 32 of *SIGGRAPH 98 Proceedings*, pages 75–84, 1998.
- [PKKG03] Mark Pauly, Richard Keiser, Leif Kobbelt, and Markus Gross. Shape modeling with point-sampled geometry. In *Computer Graphics*, *Siggraph 2003 Proceedings*, pages 641–650, San Diego, CA, July 2003.
- [Poi] Pointshop3D. <http://www.pointshop3d.com/>.
- [PSQ03] Sylvain Paris, François X. Sillion, and Long Quan. Lightweight face relighting. In *Proceedings of Pacific Graphics*, pages 41–50, Canmore, Canada, October 2003.
- [PZvBG00] Hanspeter Pfister, Matthias Zwicker, Jeroen van Baar, and Markus Gross. Surfels: Surface elements as rendering primitives. In *Computer Graphics (Proceedings of ACM SIGGRAPH 2000)*, pages 335–342, 2000.
- [Rus98] Szymon Rusinkiewicz. A new change of variables for efficient BRDF representation. In G. Drettakis and N. Max, editors, *Rendering Techniques '98 (Proceedings of Eurographics Rendering Workshop '98)*, pages 11–22, New York, NY, 1998. Springer Wien.
- [Sch03] Tobias Schwägli. 3D reconstruction using photometric and disparity-based stereo techniques. Diploma thesis, Eidgenössisch Technische Hochschule Zürich (ETH), Switzerland, 2003.
- [SGM04] P. Sander, D. Gosselin, and J. Mitchell. Real-time skin rendering on graphics hardware. *SIGGRAPH 2004 Sketch*, 2004.
- [Sig06] Christian Sigg. *Representation And Rendering of Implicit Surfaces*. PhD thesis, Department of Computer Science, ETH Zurich, 2006.

- [SLMB05] Alla Sheffer, Bruno Lévy, Maxim Mogilnitsky, and Alexander Bogomyakov. ABF++: Fast and robust angle based flattening. *ACM Transactions on Graphics — In print*, April 2005.
- [Sta95] Jos Stam. Multiple scattering as a diffusion process. In *Proceedings of the 6th Eurographics Workshop on Rendering*, pages 51–58, Dublin, Ireland, June 1995.
- [Sta01] Jos Stam. An illumination model for a skin layer bounded by rough surfaces. In *Proceedings of the 12th Eurographics Workshop on Rendering Techniques*, pages 39–52, Vienna, Austria, London, UK, June 2001. Springer-Verlag.
- [SWI97] Yoichi Sato, Mark D. Wheeler, and Katsushi Ikeuchi. Object shape and reflectance modeling from observation. In *Computer Graphics, SIGGRAPH 97 Proceedings*, pages 379–387, 1997.
- [SWPG05] Filip Sadlo, Tim Weyrich, Ronny Peikert, and Markus Gross. A practical structured light acquisition system for point-based geometry and texture. In *To appear in Proceedings of the Eurographics Symposium on Point-Based Graphics 2005*, Stony Brook, USA, June 2005.
- [SZW90] J. M. Schmitt, G. X. Zhou, and E. C. Walker. Multilayer model of photon diffusion in skin. *Journal of the Optical Society of America A*, 7(11):2141–2153, 1990.
- [Tau95] G. Taubin. A signal processing approach to fair surface design. In *Computer Graphics, SIGGRAPH 95 Proceedings*, pages 351–358, August 1995.
- [THM⁺03] Matthias Teschner, Bruno Heidelberger, Matthias Müller, Danat Pomeranerts, and Markus Gross. Optimized spatial hashing for collision detection of deformable objects. In *Proc. Vision, Modeling, Visualization VMV*, pages 47–54, 2003.
- [TOS⁺03] Norimichi Tsumura, Nobutoshi Ojima, Kayoko Sato, Mitsuhiro Shiraishi, Hideto Shimizu, Hirohide Nabeshima, Syuuichi Akazaki, Kimihiko Hori, and Yoichi Miyake. Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin. *ACM Transactions on Graphics (SIGGRAPH 2003)*, 22(3):770–779, 2003.
- [TS67] Kenneth E. Torrance and Ephraim M. Sparrow. Theory for off-specular reflection from roughened surfaces. *JOSA*, 57(9):1105–1114, September 1967.
- [Tuc00] V. Tuchin. Tissue optics: Light scattering methods and instruments for medical diagnosis. *Optical Engineering*, TT38, 2000.

Bibliography

- [Tur92] Greg Turk. Re-tiling polygonal surfaces. In *Proceedings of the 19th annual conference on Computer graphics and interactive techniques*, pages 55–64. ACM Press, 1992.
- [TWL⁺05] Xin Tong, Jiaping Wang, Stephen Lin, Baining Guo, and Heung-Yeung Shum. Modeling and rendering of quasi-homogeneous materials. *ACM Transactions on Graphics (SIGGRAPH 2005)*, 24(3):1054–1061, 2005.
- [VCBS03] J. Verdera, V. Caselles, M. Bertalmío, and G. Sapiro. Inpainting surface holes. In *Proceedings of IEEE International Conference on Image Processing (ICIP)*, pages II: 903–906, September 2003.
- [VPM⁺03] Daniel Vlasic, Hanspeter Pfister, Sergey Molinov, Radek Grzeszczuk, and Wojciech Matusik. Opacity light fields: interactive rendering of surface light fields with view-dependent opacity. In *SI3D '03: Proceedings of the 2003 symposium on Interactive 3D graphics*, pages 65–74, New York, NY, USA, 2003. ACM Press.
- [WAA⁺00] Daniel N. Wood, Daniel I. Azuma, Ken Aldinger, Brian Curless, Tom Duchamp, David H. Salesin, and Werner Stuetzle. Surface light fields for 3D photography. In *Computer Graphics, SIGGRAPH 2000 Proceedings*, pages 287–296, Los Angeles, CA, July 2000.
- [War92] G. Ward. Measuring and modeling anisotropic reflection. *Computer Graphics*, 26(Annual Conference Series):265–273, 1992.
- [WCH92] Juyang Weng, Paul Cohen, and Marc Herniou. Camera calibration with distortion models and accuracy evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(10):965–980, October 1992.
- [Wel91] E. Welzl. *Smallest enclosing disks (balls and ellipsoids)*, volume 555 of *Lecture Notes Comput. Sci.* Springer-Verlag, 1991.
- [WGT⁺05] Andreas Wenger, Andrew Gardner, Chris Tchou, Jonas Unger, Tim Hawkins, and Paul Debevec. Performance relighting and reflectance transformation with time-multiplexed illumination. *ACM Transactions on Graphics*, 24(3):756–764, 2005.
- [WHD03] Andreas Wenger, Tim Hawkins, and Paul Debevec. Optimizing color matching in a lighting reproduction system for complex subject and illuminant spectra. In *Proceedings of the Eurographics Symposium on Rendering: 14th Eurographics Workshop on Rendering*, pages 249–259. Eurographics Association, June 2003.

- [Wil78] Lance Williams. Casting curved shadows on curved surfaces. In *SIGGRAPH '78: Proceedings of the 5th annual conference on Computer graphics and interactive techniques*, pages 270–274, New York, NY, USA, 1978. ACM Press.
- [Wil05] Lance Williams. Case study: The gemini man. In *SIGGRAPH 2005 Course 'Digital Face Cloning'*, 2005.
- [WMP⁺05] Tim Weyrich, Wojciech Matusik, Hanspeter Pfister, Jinho Lee, Addy Ngan, Henrik Wann Jensen, and Markus Gross. A measurement-based skin reflectance model for face rendering and editing. Technical report, Mitsubishi Electric Research Laboratories, TR2005-071, Dec 2005.
- [WMP⁺06] Tim Weyrich, Wojciech Matusik, Hanspeter Pfister, Bernd Bickel, Craig Donner, Chien Tu, Janet McAndless, Jinho Lee, Addy Ngan, Henrik Wann Jensen, and Markus Gross. Analysis of human faces using a measurement-based skin reflectance model. *To appear in ACM Transactions on Graphics (SIGGRAPH 2006)*, 2006.
- [WPG05] Tim Weyrich, Hanspeter Pfister, and Markus Gross. Rendering deformable surface reflectance fields. In *IEEE Transactions on Computer Graphics and Visualization*, 11:48–58, January–February 2005.
- [WPK⁺04] Tim Weyrich, Mark Pauly, Richard Keiser, Simon Heinzle, Sascha Scandella, and Markus Gross. Post-processing of scanned 3D surface data. In *Proceedings of Eurographics Symposium on Point-Based Graphics 2004*, pages 85–94, Zurich, Switzerland, June 2004.
- [YDMH99] Yizhou Yu, Paul Debevec, Jitendra Malik, and Tim Hawkins. Inverse global illumination: Recovering reflectance models of real scenes from photographs. In *Computer Graphics, SIGGRAPH 99 Proceedings*, pages 215–224, Los Angeles, CA, August 1999.
- [ZBK02] Todd Zickler, Peter N. Belhumeur, and David J. Kriegman. Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction. In *ECCV '02: Proceedings of the 7th European Conference on Computer Vision-Part III*, pages 869–884, London, UK, 2002. Springer-Verlag.
- [Zha00] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.
- [ZPBG01] Matthias Zwicker, Hanspeter Pfister, Jeroen Van Baar, and Markus Gross. Surface splatting. In *Computer Graphics (Proceedings of ACM SIGGRAPH 2001)*, pages 371–378, Los Angeles, CA, July 2001.

Bibliography

- [ZPKG02] Matthias Zwicker, Mark Pauly, Oliver Knoll, and Markus Gross. Pointshop 3D: An interactive system for point-based surface editing. In *Computer Graphics, SIGGRAPH 2002 Proceedings*, pages 322–329, San Antonio, TX, July 2002.
- [ZPvBG02] Matthias Zwicker, Hanspeter Pfister, Jeroen van Baar, and Markus Gross. Ewa splatting. *IEEE Transactions on Visualization and Computer Graphics*, 8(3):223–238, 2002.
- [ZSS96] Denis Zorin, Peter Schröder, and Wim Sweldens. Interpolating subdivision for meshes with arbitrary topology. In *Proceedings of ACM SIGGRAPH 96*, pages 189–192. ACM Press, 1996.
- [ZWGS02] Zhunping Zhang, Lifeng Wang, Baining Guo, and Heung-Yeung Shum. Feature-based light field morphing. *ACM Transactions on Graphics (SIGGRAPH 2002)*, 21(3):457–464, July 2002.