Example-based image relighting

Teng, Xiao

2014

Teng, X. (2014). Example-based image relighting. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/55404>

<https://doi.org/10.32657/10356/55404>

Example-based Image Relighting

Teng Xiao

School of Computer Engineering

*A thesis submitted to Nanyang Technological University
in fulfillment of the requirement for the degree of
Doctor of Philosophy*

January 2013

To my parents, Teng Yongxiang and Zhao Meiling

Abstract

Synthesizing photo-realistic images under different illumination conditions remains an important but challenging problem in both computer vision and graphics communities. Over the past few years, many researchers in both communities have tried inventing new methods to capture the complex light interactions existing in nature. However, most of them either rely on expensive custom-built equipment, or only focus on a class of objects, such as faces, to make the problem tractable with their approaches. Our research focuses on an example-based relighting framework for general objects.

In this thesis, a review of the representative existing works on image relighting is provided. To avoid explicit 3D model acquisition as required in geometry-based methods and tedious laboratory setups as required for existing image-based methods, an example-based framework is proposed and analyzed. In this framework, we use a database of reference objects captured under different illumination conditions. Given one or several input image(s) of a new object captured under illumination conditions present in the reference database, new images can be synthesized for the input object under novel illumination conditions. We start from a single image approach with spherical surface and Lambertian model assumptions. The analysis is extended to non-spherical surface and more general BRDF models. Finally, a multiple image approach is introduced. Given limited sample images of the input object, the example-based relighting method can synthesize realistic images under different lighting conditions. It is a general and purely data-driven method which does not require an analytical model of the surface geometry or reflectance. The framework is demonstrated on standard real image databases and our own collected data, with comparisons showing that it outperforms image-based model fitting methods.

This thesis concludes with a discussion of future work for example-based relighting. These include applying the framework to more complex images using sophisticated similarity estimation methods. Besides exploring physically correct solutions, we may also consider research methods for achieving perceptually appropriate relighting.

Acknowledgements

This thesis marks the completion of my postgraduate research which was started in August 2008 at the School of Computer Engineering, Nanyang Technological University (NTU), Singapore. During these years I have received considerable help from many people, but none more so than from my supervisor Dr. Cham Tat Jen who has been a most exceptional supervisor. I would like to take this chance to express my sincere gratitude for his patience and altruistic assistance on my research. His thoughtful guidance, invaluable suggestions and constant encouragement has always been my source of inspiration. I have learned a lot from his continual enthusiasm, passion and serious attitude towards research.

At the Centre for Multimedia and Network Technology (CeMNet), I was very fortunate to have many wonderful people to spend time with me. First, I deeply appreciate Gao Shenghua for his friendship and assistance during my graduate studies. I want to thank my seniors - Pham Minh Tri, Huang Yi, Duc Viet Dung and Li Gang, my juniors - Tan Wei Chian and Tu Trung Hieu and my labmates Duan Lixin, Wang Zhengxiang, Duan Qi, Lu Haifeng, Zhang Da, Yang Ming, Zhang Yu, Hu Yiqun, Wang Huan, Cheng Xiangang, Zhong Feng, Guo Tiantian, Wang Danqi, Yang Shengbo, Zhang Chi, Hu Zhiming, Xiang Tianyu, Xiang Liu, Shou Wei, Chen Lin, Li Wen, Xu Xinxin and all colleagues in CeMNet, for their accompany and generous help either in research or in life.

I also would like to take this opportunity to thank NTU for providing me the research scholarship, and CeMNet's technicians Ms Chua Poo Hua, Ms Teo Cheng Kee, Cindy and Mrs Ng-Siom Siew Ling for their help.

I am thankful to Dr. Song Mingli, Dr. Bu Jiajun, Dr. Chen Chun, Dr. Lin Yuxu

and all colleagues at Visual Perception Key Lab of Ministry of Education and Microsoft, Zhejiang University, for their guidance and support during my undergraduate studies. They helped me set the goal of pursuing a doctoral degree.

Lastly and most importantly, I wish to express my deepest gratitude to my parents Teng Yongxiang and Zhao Meiling, for their unfailing love, support and encouragement, and to my cousin Zhang Hao, who is also working on his Ph.D in computer vision, thank him for all helpful discussions.

Contents

Abstract	i
Acknowledgements	iii
List of Figures	viii
List of Tables	xiv
List of Abbreviations	xv
1 Introduction	1
1.1 Background and Motivation	1
1.2 Research Methodology and Scope	5
1.3 Thesis Organization	5
2 Literature Survey	8
2.1 Surface Reflectance Distribution Functions	8
2.1.1 BRDF	10
2.1.2 Illumination Equation	11
2.1.3 ABRDF	13
2.2 Geometry-based Relighting	14
2.3 Image-based Relighting	18
2.3.1 Relighting using Basis Functions	19
2.3.2 Reflectance Function-based Relighting	22
2.4 Example-based Approaches	26
2.5 Example-based Relighting	27

3	Single Image Relighting	33
3.1	Framework Overview	33
3.1.1	Preliminaries	35
3.1.2	Similarity Matching	36
3.1.3	Relationship to Direct Physics-based Emulation	45
3.2	Experiments	48
3.2.1	Synthetic Images	48
3.2.2	Real Images	50
3.3	Smoothness Constraints	51
3.3.1	Markov Random Field	51
3.3.2	Standard Poisson Solution	54
3.3.3	Screened Poisson Solution	57
3.4	Computational Efficiency	59
3.5	Conclusion	61
4	Surface Shape and Image Measurements	62
4.1	Curvature based Analysis	62
4.1.1	Surface Curvature	62
4.1.2	Classification of Surface Patch using Curvature	65
4.1.3	Relation of Curvature and Synthesis Errors	66
4.2	Local Shape Analysis with Image Measurements	69
4.2.1	Local Surface Shape	69
4.2.2	Image Measurements	79
4.3	Experiments	85
4.4	Conclusion	88
5	Multiple Image Relighting	89
5.1	Introduction	89
5.2	Framework Overview	91
5.2.1	Similarity Matching	93
5.2.2	Direct ABRDF Transfer	94
5.2.3	Modeled ABRDF Transfer	94

5.2.4	Comparison with Model Fitting Methods	98
5.2.5	Color Consistency and Intensity Normalization	99
5.3	Experiments	101
5.3.1	Results from the “Observed” Experiments	102
5.3.2	Results from the “Unobserved” Experiments	108
5.3.3	Color Consistency and Intensity Normalization	113
5.3.4	Analysis	114
5.4	Conclusion	116
6	Conclusion and Future Work	118
6.1	Summary of Research	118
6.2	Future Work	120
6.2.1	Single Image Relighting	120
6.2.2	Multiple Image Relighting	121
6.2.3	Overall	122
	Appendix A: Large Results of Multiple Image Relighting	124
	Bibliography	161

List of Figures

1.1	Manual Relighting. Figure from [31]. Even with a sophisticated software, PhotoShop, it is still not very handy to relit a dull and flat photograph on top-left with the lighting condition in another photo on top-right in order to spice it up as the result on bottom.	3
1.2	The illumination variation of a frontal face from the Yale Face Database B [30]. The images are divided into groups according to the angle between the lighting direction and optic axis. (From Top to bottom: up to 12° , 25° , 50° , 77°)	4
1.3	Summary. The middle part of the figure is from [58, 40]. From top to bottom, the whole diagram shows the flow and the connections of the research reported in this thesis. The cloud on top is the key of our approaches, <i>i.e.</i> the reference data which contains the images of different types of object under various lighting conditions. This example database is used to relight objects with a single image or multiple images. Reflectance and geometry of objects are analyzed as shown in the middle. Reflectance model can be either simple single-lobe or complex multiple-lobes functions, while geometry can be characterized by two principal curvatures. Lambertian reflectance (single lobe) and spherical shape (equal principal curvatures) are assumed first, as the limitation of the single image approach has been found in the analysis, a multiple image approach is introduced to relax these assumptions.	6
2.1	BRDF, the reflectance properties of a surface defined as a function of illumination direction ω_i and viewing direction ω_r relative to the local surface orientation N	11
2.2	Vectors in Blinn-Phong shading, the default shading model used in OpenGL and Direct3D's fixed-function pipeline. Halfway vector $\vec{H}$ is introduced to replace $\vec{R} \cdot \vec{V}$ with $\vec{N} \cdot \vec{H}$ for faster performance.	12
2.3	The relation of different components in the Phong model, figure from [15]. The total function is the sum of three components.	13
2.4	Figure from Marschner and Greenberg [49] (a-d) Relighting a diffuse, rigid object, (e-g) relighting a face for a consistent composite.	15

2.5	Relighting method of Pighin et al. [64]. A coarse approximation of 3D face model (middle) is used to render the greyscale change under a new synthetic light source.	16
2.6	The relighting method used in [57]. The light distribution or the light source direction was extracted from the eyes of subjects in the images. 3D model is estimated to enable relighting.	16
2.7	Photometric consistent superimposition, figure from [59]. The normal map of the subject was manually assigned by user.	17
2.8	Haeberli's Synthetic Lighting for Photography [33]. From left to right, the image are: with the ambient, with the left lamp on, with the right lamp on and the last one with the synthetical illumination which are modified and combined from the first three images.	20
2.9	Figure from [56], all 20 images are rendered as linear combinations of 9 basis images, the sequence of the images are corresponding to a sequence of sun positions.	21
2.10	Bayesian Relighting by Fuchs et al. [28] The best coefficients of a linear combination was estimated from probe ball images, then the same coefficients were applied to the training scene images to synthesize the relighting result.	22
2.11	Figure from [21], the mosaic of reflectance functions.	24
2.12	Figure from [21], face synthetically illuminated with novel lighting.	25
2.13	Figure from [48], rendering results with polynomial texture map (top) and conventional texture map (bottom).	25
2.14	Relighting comparison with different reflectance reconstruction methods by Masselus et al. [50].	25
2.15	Image Analogies, figure from [34]. Images A and A' are the example pair, images B and B' are the input and the result.	26
2.16	Artistic lighting transferring. The first row shows the masterpiece by Ansel Adams and the "flat" input photo, and the second row shows the results from direct histogram transfer and Bae et al. [6]	27
2.17	The Quotient Image for relighting, figure from [70]. (b) and (c) are the source image and its quotient image. The bootstrap set of three example objects is shown in (a), the synthesized images are shown in (d).	28
2.18	Face relighting, figure from [81]. The leftmost column shows the input images, the top row shows the example images.	29
2.19	Illumination transferring with known geometry, figure from [47].	30

2.20	Face relighting using reflectance transfer, figure from [61]. From left to right: the reference subject, the target subject, the warped quotient image, the target subject with similar illumination as the reference subject.	30
2.21	Face illumination transferring with locally constrained global optimization, figure from [17]. The first row shows reference images, the second row shows the target image to be relit and the relighting results.	31
2.22	Face illumination transferring with edge-preserving filters, figure from [18]. The first row shows reference images, the second row shows the target image to be relit and the relighting results.	31
3.1	Framework overview: for every pixel of the input image B , we search the best match by intensity and gradient in the example image A with the same illumination condition, after we find the best match, we do a lookup what it looks like in the example image A_p with another illumination condition, then assign that value to the result B_p	34
3.2	Surface normal is parameterized by its σ and its tilt τ	36
3.3	Tilt estimation error as a function of eccentricity, figure from [25]	47
3.4	Relighting sphere with sphere example. Images A and A_p are the reference pair, images B and B_p are the input and result. The ground truth and error images are shown in the middle and on the right of each subfigure respectively. The coordinate maps A_{coords} and B_{coords} are introduced to show the corresponding points between the reference and test objects, which are found by the algorithm and assigned the same color.	49
3.5	Relighting cylinder with sphere example. Similar as figure 3.4.	50
3.6	Relighting ellipsoid with sphere example. Similar as figure 3.4.	51
3.7	Relighting complex models with sphere example. Similar as figure 3.4.	53
3.8	Relighting real images with examples. Similar as figure 3.4 except the Poisson reconstructed image $B_p-recons$	54
3.9	Relighting real images with Poisson solutions. From left to right, the second row shows the intensity transfer result; the standard Poisson result from gradient transfer (equivalent to set $\lambda = 0$ in screened Poisson solution); the standard Poisson result from gradient transfer with optimal offset w.r.t. real lighting result B_p-real ; the screened Poisson result from intensity and gradient transfer (set $\lambda = 0.01$); and the screened Poisson result after it is histogram-matched to image B . The third row shows the differences between the results in the second row and the real lighting result. The numbers indicate the average intensity errors.	60

4.1	The principal curvatures of the 3d models. The values are indicated by the color bars.	67
4.2	Koenderink's shape index and curvedness, figure from [58]	68
4.3	Approximate linear relation relation of curvature and synthetic error. In "avg error", k_1 is the vertical axis, $[-0.5, 0.5]$ from top to bottom, and k_2 is the horizontal axis, $[-0.5, 0.5]$ from left to right.	73
4.4	Four types of surface shapes. Each shape corresponds to four rows in each table, the row numbers are listed in the caption of each shape. . .	87
5.1	Framework Overview. Given limited images of a new object as input, for each point on this object, the best match point in reference database is found by searching for the near neighbour in a vector space, where each dimension of the vectors is corresponding to the image intensity level of the points under one lighting condition. Then new images of the object can be synthesized through direct or modeled ABRDF transfer.	92
5.2	Comparison between Example-based Relighting and Model Fitting Results (Observed in Examples). Database: YalefaceB. "PF" [48], "T-F" [41] and "ER" means Polynomial Fitting, Tensor Fitting and Example-based Relighting, respectively. The number before them is the order of the model, and the number after is the number of input images for fitting or matching. Same abbreviations will be used in following figures.	103
5.3	Comparison between Example-based Relighting and Model Fitting Results (Observed in Examples): One case. Database: YalefaceB. The input images for fitting or matching are marked by crosses, and the test cases are marked by circles. For better visualization, the displayed error images are amplified by a factor of 3 and shown in negative gray-scale. For the abbreviations, please refer to figure 5.2. Large version of the results can be found in Appendix A.	105
5.4	Comparison between Example-based Relighting and Model Fitting Results (Observed in Examples): One case, with bad samples. Database: YalefaceB. Large version of the results can be found in Appendix A. . .	106
5.5	Comparison between Example-based Relighting and Model Fitting Results (Observed in Examples): One case. Database: BOLD. For the explanation of abbreviations and the meaning of error images, please refer to figure 5.2 and 5.3. Large version of the results can be found in Appendix A.	107
5.6	Comparison between Example-based Relighting and Model Fitting Results (Unobserved in Examples). Database: YalefaceB. Besides the abbreviations used in figure 5.2, "W" is short for "With".	109

5.7	Comparison between Example-based Relighting and Model Fitting Results (Unobserved in Examples): One Case. Database: YalefaceB. For the explanation of abbreviations and the meaning of error images, please refer to figure 5.2 and 5.3. Large version of the results can be found in Appendix A.	111
5.8	Comparison between Example-based Relighting and Model Fitting Results (Unobserved in Examples): One Case. Database: BOLD. Large version of the results can be found in Appendix A.	112
5.9	Color Consistency and Intensity Normalization. Besides the abbreviations used in figure 5.2, “L” means luminance and “N” means normalization. The displayed error images are amplified by a factor of 3 and shown in negative gray-scale.	113
5.10	PCA Analysis. Database: YalefaceB.	115
5.11	Relation between the Best Match Distance and Example-based Relighting (ER) Error. Warmer colors mean higher density of points. The corresponding points of the outliers are shown on the right. Database: YalefaceB.	116
5.12	Relation between Example-based Relighting (ER) Error and Polynomial Fitting (PF) Error, with (right) or without (left) sufficient example data. Warmer colors mean higher density of points. Database: YalefaceB. . .	116
A.1	Comparison between Example-based Relighting and Model Fitting Results (Observed in Examples): One Case with good samples. Database: YalefaceB. Large version of Figure 5.3. In this high resolution version, the results of our method may appear noisy, as it is a pointwise transfer. Poisson reconstruction methods mentioned in Chapter 3.3 are used to refine the results.	133
A.2	Comparison between Example-based Relighting and Model Fitting Results (Observed in Examples): One Case with bad samples. Database: YalefaceB. Large version of Figure 5.4. A poor choice of illumination conditions for the test object often led to poor model fitting. However, the example-based relighting method is much more robust to this choice of illumination conditions.	139
A.3	Comparison between Example-based Relighting and Model Fitting Results (Observed in Examples): One Case. Database: BOLD. Large version of Figure 5.5.	146
A.4	Comparison between Example-based Relighting and Model Fitting Results (Unobserved in Examples): One Case. Database: YalefaceB. Large version of Figure 5.7.	153

A.5 Comparison between Example-based Relighting and Model Fitting Results (Unobserved in Examples): One Case. Database: BOLD. Large version of Figure 5.8.	160
--	-----

List of Tables

4.1	Surface classification based on Gaussian and mean curvatures	66
4.2	Surface classification based on principal curvatures	66
4.3	Numerical results: solved parameters. There are 16 unique solutions found, one solution each row.	86
4.4	Numerical results: resulting derivatives. These solutions all result in the same combination of image measurements.	86
4.5	Numerical results: errors of resulting derivatives (10^{-015} *).	86
4.6	Numerical results: errors of the constraints used in nonlinear solvers (10^{-015} *)	87
5.1	Average errors for the methods mentioned in this chapter. Besides the abbreviations used before, “O”, and “U” are short for Observed and Unobserved.	110

List of Abbreviations

ABBREVIATIONS	FULL EXPRESSIONS
ABRDF	Apparent BRDF
BOLD	Birmingham Object Lighting Database
BRDF	Bidirectional Reflectance Distribution Function
BSSRDF	Bidirectional Surface Scattering Reflectance Distribution Function
BTF	Bidirectional Texture Function
CT	Cosine Transform
DSLR	Digital Single-Lens Reflex Camera
ER	Example-based Relighting
ICT	Inverse Cosine Transform
MAP	Maximum-A-Posteriori
PF	Polynomial Fitting or Plenoptic Function
PR	Poisson Reconstruction
$\mathbb{R}^n$	n-dimensional Euclidean space
SF	Scattering Function
SHBMM	Spherical Harmonic Basis Morphable Model
SVBRDF	Spatially-varying BRDF
SVD	Singular Value Decomposition
SI	International System of Units or Shape Index
SF	Scattering Function
TF	Tensor Fitting
YalefaceB	Yale Face Database B

Chapter 1

Introduction

1.1 Background and Motivation

Due to the ubiquity of digital cameras, encompassing both professional DSLRs as well as inexpensive cameras on mobile phones, digital photography is now widely prevalent. However, since the camera was invented many years ago, the user procedure of capturing still photographs has not changed a lot.

It is often the case that even with expensive and powerful equipment, most of the photos taken by amateurs are flat and dull compared to those taken by professionals. Inappropriate lighting is often responsible for poor quality photos, with the best photographers distinguished by their patience in waiting for perfect outdoor lighting conditions, or their ability in designing ideal illumination conditions in studio setups. Unfortunately, the optimal illumination is rarely available, often occurring only propitiously. It is usually both time-consuming and costly to deliberately setup the perfect lighting for each photography session. Alternatively, users may post-process the photos physically or digitally. There are online tutorials teaching how to spice up photos by

virtually altering lighting conditions in software photo-editing tools (see figure 1.1 [31]). However even with some commercial editing tools, it still requires tedious work to generate visually-pleasing results. In figure 1.1, the inconsistency between the sun position and the shading on the castle is still noticeable.

In order to make photography easier for everyday life, a variety of methods have been invented to allow *lazy* shots, *e.g.* Microsoft Photo Fuse can rescue a group portrait by merging the best parts of a series of shots, and the Lytro camera provides the functionality of changing focus and perspective after a photo is taken. Automatic image relighting is also an interesting research area, with the potential to enable a wide population of users to take great photographs without worrying about illumination conditions during acquisition.

In many computer vision tasks, such like object recognition, detection and tracking, the large image variation due to illumination, such as shown in figure 1.2 [30], is one of the major problems encountered as it is difficult to identify whether the intensity changes are due to surface properties, or illumination, or both. For example, previous research on face recognition [2] reveals that the changes induced by illumination can easily be larger than the variations due to the change of identity. The illumination problem has received considerable attention in recognition literature, with most of the methods trying to discard variations due to lighting. Intuitively, it may be helpful if we can first relight test images according to the illumination under which the training images are captured.

Image relighting potentially has many other practical applications. It will be valuable in areas as diverse as interior light design, video communication enhancement, augmented reality, and many other photo/video post-production tasks. We aim to create new methods that can relight the scene in an input image under alternative illumination conditions.

Figure 1.1: Manual Relighting. Figure from [31]. Even with a sophisticated software, PhotoShop, it is still not very handy to relit a dull and flat photograph on top-left with the lighting condition in another photo on top-right in order to spice it up as the result on bottom.

There is substantial research on synthesizing the images of an object or a scene under different lighting conditions, however, most of the research either rely on expensive custom-built equipment, or focus on a class of objects, such as faces, to make the problem tractable with the proposed approaches. Geometry-based relighting methods require 3D information of the object while traditional image-based relighting approach-

Figure 1.2: The illumination variation of a frontal face from the Yale Face Database B [30]. The images are divided into groups according to the angle between the lighting direction and optic axis. (From Top to bottom: up to 12° , 25° , 50° , 77°)

es need a laboratory setup for each object. Both are cumbersome, so we want to explore the possibility of an example-based relighting framework. In this thesis, we intend to transfer the missing information from a pre-collected reference dataset containing images of different generic objects under a variety of lighting conditions.

1.2 Research Methodology and Scope

As shown in figure 1.3, the research was conducted as follows. We started from a single image approach with strong assumptions, such as using a spherical geometry model and the Lambertian reflectance model. We found that such assumptions break down quickly on general surfaces. We extended the analysis to non-spherical surfaces, by locally modeling them with surface curvature. We found that even with the Lambertian model, the surface shape is not uniquely determined by image measurements up to second order. Therefore, a multiple image approach was proposed which can be applied on non-spherical and non-Lambertian surfaces. Given the input image(s), we did a lookup what the object contained in the image(s) looks like with other illumination conditions in a reference database, then relighting can be performed via transfer.

In this thesis, we assume that the scene contains only directional lighting, but do not require geometric structure of the scene to be known.

1.3 Thesis Organization

The rest of the thesis is organized as follows:

- **Chapter 2** provides background information on reflectance functions and discusses the previous works on image relighting, such as geometry-based methods,

Figure 1.3: Summary. The middle part of the figure is from [58, 40]. From top to bottom, the whole diagram shows the flow and the connections of the research reported in this thesis. The cloud on top is the key of our approaches, *i.e.* the reference data which contains the images of different types of object under various lighting conditions. This example database is used to relight objects with a single image or multiple images. Reflectance and geometry of objects are analyzed as shown in the middle. Reflectance model can be either simple single-lobe or complex multiple-lobes functions, while geometry can be characterized by two principal curvatures. Lambertian reflectance (single lobe) and spherical shape (equal principal curvatures) are assumed first, as the limitation of the single image approach has been found in the analysis, a multiple image approach is introduced to relax these assumptions.

traditional image-based methods and other example-based methods.

- **Chapter 3** presents a single image relighting method using an analogy-based framework. Theoretical proof is given for cases in which certain assumptions can be made, such as spherical surfaces and Lambertian reflectances. Poisson reconstruction methods are introduced to refine the results.
- **Chapter 4** provides an analysis of the errors of the single image approach with respect to surface curvature. The analysis is extended to non-spherical surfaces, and the limitation of using high order derivatives is discussed.
- **Chapter 5** presents a multiple image relighting method using an example-based framework. To relight the input objects, pointwise ABRDF exemplars are matched and transferred. There is no assumption of same shape or even topology between test and reference objects. This method can be extended to relight objects under illumination conditions not available in the reference data. It outperforms state-of-the-art polynomial-based model fitting methods.
- **Chapter 6** concludes the thesis and sketches out future research directions.

Chapter 2

Literature Survey

Surface reflectance models are a central concept in the task of image relighting as they define how light is reflected at the surface of interest. A brief review of surface reflectance distribution functions will be presented first, followed by a survey of existing relighting methods. These methods will be classified into different approaches of Geometry-based Relighting, Image-based Relighting and Example-based Relighting.

2.1 Surface Reflectance Distribution Functions

In general, when light interacts with a surface, different portions of the incident light will be reflected, absorbed or transmitted. For opaque materials, the incident light is mainly reflected or absorbed. To formally quantify the amount of incident and reflected light, the two standard measures of irradiance and radiance are used. Irradiance is defined as the power of light falling on a unit area of surface, with SI units of watts per squared meter. Radiance is defined as the power of light reflected from a surface point, per unit foreshortened surface area projected perpendicular to the direction of light, per unit solid angle. The SI units of radiance are watts per squared meter per

steradian.

One useful way to think about the distribution of light in general space is to consider the Plenoptic Function (PF) [1]:

$$PF(x, y, z, \theta, \phi, t, \lambda). \quad (2.1)$$

The function returns the radiance of the light at a particular wavelength λ , at a point in space (x, y, z) , at time t and travelling in the direction with angular coordinates (θ, ϕ) . If color is to be ignored, the wavelength of the light can be dropped from the function.

The surface reflectance distribution function is a function that describes the relation between incoming irradiance and outgoing radiance, it represents the property of a surface with respect to illumination. The most general form of the function is the 14D Scattering Function (SF) defined as:

$$SF(x_i, y_i, z_i, \theta_i, \phi_i, t_i, \lambda_i; x_r, y_r, z_r, \theta_r, \phi_r, t_r, \lambda_r), \quad (2.2)$$

which returns the ratio of the resultant scattered radiance to the causal incident irradiance at different points on and within a 3D object. The parameters associated with the incident light have a subscript i , while those associated with the scattered light have a subscript r . The scattering function may also be considered as a ratio of the scattered light plenoptic function to the incident light plenoptic function.

However, the full function is difficult to capture in practice. By assuming invariance to time t and wavelength λ , and only modeling the surface points instead of the whole space, the Bidirectional Surface Scattering Reflectance Distribution Function

(BSSRDF) is defined as an 8D function:

$$BSSRDF(x_i, y_i, \theta_i, \phi_i; x_r, y_r, \theta_r, \phi_r), \quad (2.3)$$

where light entering the surface may scatter internally and exit at another surface location.

Assume subsurface scattering can be ignored, the BSSRDF can be further simplified into the Spatially-Varying Bidirectional Reflectance Distribution Function (SVBRDF) which is a 6D function:

$$SVBRDF(x, y; \theta_i, \phi_i; \theta_r, \phi_r), \quad (2.4)$$

where light reaches and leaves the surface at the same location (x, y) .

2.1.1 BRDF

The Bidirectional Reflectance Distribution Function (BRDF) is the function that describes how much light is reflected, ignoring the variation across surface in SVBRDF or only considering the function at a single point. As shown in figure 2.1, it gives the reflectance properties of a surface as a function of illumination direction and viewing direction relative to the local surface orientation N . It is defined as the ratio of the reflected light L_r in direction $\omega_r(\theta_r, \phi_r)$ to the light E_i that reaches the surface from direction $\omega_i(\theta_i, \phi_i)$.

$$BRDF = \frac{dL_r}{dE_i} = \frac{dL_r}{L_i \cos \theta_i d\omega_i}, \quad (2.5)$$

where L_i is the light source intensity, θ_i is the angle between ω_i and the surface normal N . The cosine term represents foreshortening of the surface element in the direction of the light.

Figure 2.1: BRDF, the reflectance properties of a surface defined as a function of illumination direction ω_i and viewing direction ω_r relative to the local surface orientation N .

2.1.2 Illumination Equation

There are a lot of techniques developed to model the BRDF and incorporate into shading models. A widely used empirical shading model is the Phong model [63]. The model computes the intensity of a surface patch as observed at a known viewer location. It essentially treats the intensity as the sum of three components: a constant ambient component, a viewer-independent diffuse component and a viewer-dependent specular component. The formula is expressed as:

$$I = k_a I_a + I_i [k_d (\vec{N} \cdot \vec{L}) + k_s (\vec{R} \cdot \vec{V})^\alpha]. \quad (2.6)$$

I, I_a, I_i : the intensity of the reflected light, the ambient light and the light source,

k_a, k_d, k_s : ambient, diffuse and specular reflection coefficient of certain material,

$\vec{N}$: surface normal direction, $\vec{L}$: light direction, $\vec{R}$: reflection direction, $\vec{V}$: viewer

direction,

α : specular exponent.

For faster performance, instead of calculating the angle between R and V , Blinn [14] decided to replace $\vec{R} \cdot \vec{V}$ with $\vec{N} \cdot \vec{H}$, where $\vec{H}$ is the halfway vector, equals to $\frac{\vec{L} + \vec{V}}{|\vec{L} + \vec{V}|}$ as shown in figure 2.2. Though Blinn-Phong model is an approximation to the original Phong model, it has been experimentally validated that the results more accurately match real-world data [54].

The angle between $\vec{N}$ and $\vec{H}$ is smaller than the angle in Phong model (half if in the same plane), so we can find an exponent α' which makes $(\vec{N} \cdot \vec{H})^{\alpha'} \approx (\vec{R} \cdot \vec{V})^\alpha$, (2.6) can be rewritten as:

$$I = k_a I_a + I_i [k_d (\vec{N} \cdot \vec{L}) + k_s (\vec{N} \cdot \vec{H})^{\alpha'}]. \quad (2.7)$$

Figure 2.2: Vectors in Blinn-Phong shading, the default shading model used in OpenGL and Direct3D's fixed-function pipeline. Halfway vector $\vec{H}$ is introduced to replace $\vec{R} \cdot \vec{V}$ with $\vec{N} \cdot \vec{H}$ for faster performance.

The individual intensity contribution from the different components in the Phong model can be seen from figure 2.3, where the 3D plots indicate the relative intensities from different viewing angles. Diffuse reflection is usually considered as the dominant reflection component for many materials. In the Lambertian model, the image intensity is independent of the viewing direction and can be represented as:

$$I = k_d I_i (\vec{N} \cdot \vec{L}) = k_d I_i \cos \alpha = \text{constant} \cos \alpha, \quad (2.8)$$

which is exact the diffuse component of the Phong or Blinn-Phong model. The Lambertian model will be used in our single image approach.

Figure 2.3: The relation of different components in the Phong model, figure from [15]. The total function is the sum of three components.

2.1.3 ABRDF

The Bidirectional Texture Function (BTF) is introduced in [19]. Instead of representing the ratio between radiance and irradiance, it integrates the lighting conditions of a particular light source. In practice, images captured from different viewports under different lighting conditions are stored as samples. The benefit of BTF is that it accounts for effects induced by both reflectance and geometry, *e.g.* specular reflectance and shadowing. The BTF is parameterized similar to the SVBRDF, and at each surface

point it is called the Apparent BRDF (ABRDF),

$$ABRDF(\theta_i, \phi_i; \theta_r, \phi_r). \quad (2.9)$$

By locking the viewport, two degrees of freedom (θ_r, ϕ_r) are dropped,

$$ABRDF(\theta_i, \phi_i). \quad (2.10)$$

Even so, the data needed to represent ABRDF is huge, one sample per lighting direction. Model fitting methods, such as biquadratic polynomials [48] and 3rd order Cartesian tensors [7] were used to compactly model ABRDF, which then can be evaluated under given lighting conditions. Such methods are much faster than fitting a different nonlinear analytical model, like Phong, for every pixel.

All BRDF related functions are actually wavelength dependent, but in practice it has been ignored or approximated by independent BRDFs in all color channels, as color is a perceptual simplification for the spectral distribution of light.

2.2 Geometry-based Relighting

The illumination recorded in images mainly depends on three basic elements: scene geometry, surface reflectance and light source. Most existing works explicitly measure or estimate these properties to relight a scene. The solutions of generating new images based on traditional graphics techniques usually render the object under some artificial lighting conditions with a 3D representation of the scene or object obtained by special equipment or estimated from 2D image data.

With a laser range camera, Marschner and Greenberg [49] obtained a 3D model

of the scene while taking a photo, with the model used to estimate the lighting in the image. Once the light distribution was determined, two images were rendered under the existing illumination and the desired illumination based on the 3D model. The relighting can then be applied on the original photo by multiplying it with the pixelwise ratios of the two rendered images. Relighting examples of a watering pot and a face are shown in figure 2.4.

Figure 2.4: Figure from Marschner and Greenberg [49] (a-d) Relighting a diffuse, rigid object, (e-g) relighting a face for a consistent composite.

Using photogrammetric techniques, Pighin et al. [64] generated the coarse approximation of 3D face model, and rendered it using a new synthetic light source in greyscale as show in figure 2.5. The rendering was then used to scale the colors in the original images for a dramatic change in lighting.

There are a group of methods employing reflectance estimation from images and 3D models [80, 20, 16, 75]. For example, Boivin and Gagalowicz [16] iteratively approxi-

Figure 2.5: Relighting method of Pighin et al. [64]. A coarse approximation of 3D face model (middle) is used to render the greyscale change under a new synthetic light source.

mated the BRDF model, going from a Lambertian hypothesis to a complex reflectance model, by reducing the differences between the image rendered with Lambertian model of the known 3D geometry and the real image. Then the original scene can be visualized under novel illuminations with recovered reflectance properties of the surfaces.

In Nishino and Nayar’s work [57], the light distribution or the light source direction was extracted from the eyes of subjects in the images. Then virtual 3D objects and 3D face models (integrated from estimated surface normal maps) can be relit under novel lighting conditions as shown in figure 2.6.

Figure 2.6: The relighting method used in [57]. The light distribution or the light source direction was extracted from the eyes of subjects in the images. 3D model is estimated to enable relighting.

Additional methods use simple geometric representations such as an ellipsoid [9], or a generic 3D face model [77] that are aligned to faces in the image, then composite the new rendered light effect based on these approximations onto the original images. Such methods reduce the requirement to know the geometry accurately, but are limited to prior known object groups such as faces. Fuchs et al. [27] used a morphable model [13] to fit to images containing faces with different poses and illumination conditions. The geometry and reflectance were then estimated, which allow rendering new images with complex lighting conditions.

Okabe et al. [59] developed a system to manually assign an approximate normal map over a photo with pen-based user interface, then the scene can be re-rendered with a new lighting condition specified by users. Some superimposition results are shown in figure 2.7.

Figure 2.7: Photometric consistent superimposition, figure from [59]. The normal map of the subject was manually assigned by user.

Recently, Haber et al. [32] presented a method using an all-frequency relighting framework introduced by Ng et al. [53]. This framework can estimate the surface reflectance and illumination from image collections with known geometry. Haber et al. [32] combined it with multiview stereo reconstruction to recover all information (geometry, reflectance and illumination) simultaneously from images of a static scene, then relighting can be conducted for a novel view.

The primary difficulty with above approaches is that they all require some knowledge of 3D geometry. The geometry must be either obtained beforehand using 3D scanning equipment, estimated from image-based approximation methods such as photometric stereo [79], shape from motion [76] and passive stereo techniques [5], or assumed to be adequately defined by some generic 3D model. The benefit from known geometry is that most methods can recover complex BRDF and it is possible to re-render the object or scene from a novel view.

2.3 Image-based Relighting

Image-based methods relight real scenes without directly estimating 3D structure as in traditional geometry-based techniques, but require surface reflectance and illumination information which are implicitly captured in existing images. These approaches use recorded images to render new images, where image-based entities are the basic rendering primitives. They do not use any explicit geometric representations, and the rendering is independent of scene complexity as image pixels are directly manipulated.

2.3.1 Relighting using Basis Functions

Relighting methods using basis functions depend on illuminance calculations obeying the super-position principle. This means:

- If the intensities of the light sources are increased, the intensities of the rendered images are multiplied by the same factors.
- The image rendered with n light sources is the sum of the n images resulting from each light source independently.

Several methods first decompose the luminous intensity distributions into a series of basis luminance functions, then relight the image with novel illumination by computing a weighted summation of the rendered images, which are pre-recorded under each basis luminance function.

Haeberli [33] demonstrated (figure 2.8) how to relight a scene by super-position after it has been photographed. Several images of the same scene were captured under a set of basis lighting conditions. The contribution of each light source can then be modified and combined to achieve the desired lighting effect.

Nimeroff et al. [56] described some desirable properties of basis functions: the number of basis functions should be as small as possible to reduce rendering and storage requirements, yet sufficient so that the combination of basis functions can synthesize any arbitrary light source. Steerable functions are introduced in [56], which can be formed from a linear combination of rotated version of themselves. Such rotation invariant basis functions are useful (see figure 2.9) for synthesizing sunlight illumination, as the sun path is roughly on a sphere.

Legendre polynomials and spherical harmonic functions were introduced by Dobashi et al. [22] as basis functions. Luminous intensity distributions can be specified as a

Figure 2.8: Haeberli’s Synthetic Lighting for Photography [33]. From left to right, the image are: with the ambient, with the left lamp on, with the right lamp on and the last one with the synthetical illumination which are modified and combined from the first three images.

series of Legendre polynomials or spherical harmonic functions.

Georghiades et al. [10, 29, 30] theoretically proved the set of n -pixel images of an object is a convex cone in n -dimensional Euclidean space $\mathbb{R}^n$, which means the vector set is closed under linear combinations with positive coefficients. More specifically, the set of images is a convex polyhedral cone in $\mathbb{R}^n$ for convex objects rendered under the Lambertian model, so it often can be determined from three properly chosen basis vectors (images). The dimension of the cone is determined by the number of distinct surface normals.

Figure 2.9: Figure from [56], all 20 images are rendered as linear combinations of 9 basis images, the sequence of the images are corresponding to a sequence of sun positions.

Fuchs et al. [28] explored further on the linear nature of light using a Bayesian maximum-a-posteriori (MAP) approach. Every scene under some illumination condition was captured with a black snooker ball present, the image of the ball being a representation of the incident light distribution. Given the image of the probe ball under a novel illumination condition, MAP estimation was used to find the best coefficients to specify the image as a linear combination of training probe ball images.

The same coefficients then were applied to the training scene images to synthesize the relighting result as shown in figure 2.10.

Figure 2.10: Bayesian Relighting by Fuchs et al. [28] The best coefficients of a linear combination was estimated from probe ball images, then the same coefficients were applied to the training scene images to synthesize the relighting result.

2.3.2 Reflectance Function-based Relighting

Reflectance function-based methods explicitly measure the reflectance properties in all visible parts of the scene.

Debevec et al. [21] used the reflectance function $R_{xy}(\theta, \phi)$ to represent how surfaces transform incident into radiant illumination. First the reflectance field is described by an 8D BSSRDF as:

$$R = R(R_i; R_r) = R(u_i, v_i, \theta_i, \phi_i; u_r, v_r, \theta_r, \phi_r), \quad (2.11)$$

where (u, v) is a point on an unoccluded surface, (θ_i, ϕ_i) is the direction of light arriving at the surface point, and (θ_r, ϕ_r) is the direction of light leaving the surface point. To save acquisition time and storage, they assumed the incident illumination field originates far away from the surface, and reduce the field to:

$$R = R(R_i; R_r) = R(\theta_i, \phi_i, u_r, v_r, \theta_r, \phi_r). \quad (2.12)$$

At each pixel, they formed a slice of the reflectance field which is the so called reflectance function $R_{xy}(\theta, \phi)$, which represents the amount of light reflected to the camera by pixel (x, y) with the illumination from direction (θ, ϕ) . Assuming $L_i(\theta, \phi)$ is the i th original light source, relighting was computed by summing the results of multiplying R_{xy} with each novel incident illumination L_i .

$$L(x, y) = \sum_{\theta, \phi} R_{xy}(\theta, \phi) L_i(\theta, \phi). \quad (2.13)$$

The reflectance function was densely sampled as shown in figure 2.11, and images of the face synthetically illuminated with novel lighting are shown in figure 2.12.

Polynomial Texture Maps, used in Malzbender et al. [48], store the lighting information as polynomials. Instead of a reflectance function image for each texel, only a few coefficients are needed, thus improving the compactness of the representation. During the relighting process as shown in figure 2.13, the polynomials are evaluated to generate new images under various illumination conditions.

Masselus et al. [50] compared a set of methods to reconstruct the continuous reflectance function from sparse reflectance function samples recorded in images with directional illumination. These methods are variously based on zero-order hold [21], interpolation (linear & IDW), biquadric polynomial [48], spherical harmonics [78], wavelets (5/3 LeGall & 9/7 Daubechies), and B-splines. They found that multilevel B-splines achieve the best reconstruction result. A comparison of the relighting results with these reconstructed reflectance functions are shown in figure 2.14.

Sen et al. [69] captured the light-field similar as the work by Debevec et al. [21], however, they described a method to elegantly encoded the 4D light transport. Based on this, they demonstrated how various types of scene relighting could be performed.

Figure 2.11: Figure from [21], the mosaic of reflectance functions.

Figure 2.12: Figure from [21], face synthetically illuminated with novel lighting.

Figure 2.13: Figure from [48], rendering results with polynomial texture map (top) and conventional texture map (bottom).

Figure 2.14: Relighting comparison with different reflectance reconstruction methods by Masselus et al. [50].

A recent work by Barmpoutis et al. [7] used anti-symmetric tensor splines for the estimation of the Apparent BRDF field for human faces from at least nine images. These account for specularities and cast shadows, without requiring explicit 3D information. Essentially, the images under novel lighting conditions are interpolated or extrapolated from observed ones. The method performed better when the observed illumination directions were more uniformly distributed. The impact of the distribution on tensor-based method was discussed in [41].

2.4 Example-based Approaches

In [11], Beymer and Poggio developed a framework to generate new images of the same object or objects of a known class from example images. For instance, the pose and expression parameters of face images are learned then transferred to novel face images.

Efros and Leung [24] proposed an example-based texture synthesis approach, in which the input example texture image is sampled, and the new texture image grown with constraints of maintaining neighborhood coherence. Efros and Freeman [23] performed texture transfer on arbitrary images by matching correspondence maps.

Figure 2.15: Image Analogies, figure from [34]. Images A and A' are the example pair, images B and B' are the input and the result.

Hertzmann et al. [34] learned filtering and painting operations from example input-output pairs, which were then applied to new inputs, as shown in figure 2.15. Similar

ideas were applied to example-based super-resolution [26] and example-based photometric stereo [35].

2.5 Example-based Relighting

Example-based artistic relighting is an interesting topic in computational photography. To reduce the tedious effort of manual touching, as shown in figure 2.16, Bae et al. [6] proposed a method to transfer a particular photographic look from a model photo to a “flat” photo.

Figure 2.16: Artistic lighting transferring. The first row shows the masterpiece by Ansel Adams and the “flat” input photo, and the second row shows the results from direct histogram transfer and Bae et al. [6]

However in this thesis, we will mainly focus on the relighting effect caused by the changing of the illumination direction.

Figure 2.17: The Quotient Image for relighting, figure from [70]. (b) and (c) are the source image and its quotient image. The bootstrap set of three example objects is shown in (a), the synthesized images are shown in (d).

Riklin-Raviv and Shashua [68, 70] introduced quotient (ratio) images for image-based relighting of objects in the same class. When given three images of each object in the same class, and another image of a new object in the same class, novel images of the object with new illumination conditions can be synthesized, as shown in figure 2.17. The three images of example objects are rendered under three different illumination conditions. All the example images are called the bootstrap set.

Zhang et al. [81] introduced Spherical Harmonic Basis Morphable Model (SHBMM) for face relighting. From a single image, they showed it is possible to estimate the model parameters of face geometry, spherical harmonic bases and illumination coefficients. The face texture (albedo) can then be computed from the spherical harmonic bases. Given one face image, it can be relit by providing an example face image under the desired novel illumination condition, as shown in figure 2.18.

Ramamoorthi and Hanrahan [66, 67, 46, 47] developed a theoretical framework

Figure 2.18: Face relighting, figure from [81]. The leftmost column shows the input images, the top row shows the example images.

from the frequency domain point of view. They converted the reflection operator as a directional convolution. It can be used for analyzing the reflected light field of a curved convex homogeneous surface under distant illumination. They showed that (figure 2.19) when given photos of two objects of known geometry, the image of one object under a new lighting condition can be rendered from three images. The three images are the image of another object under the new lighting condition, and the images of both objects rendered under a different lighting condition.

Peers et al. [61] proposed a face relighting approach to transfer the desired illumination on an example subject to a new subject, using quotient images and an image warping. The example images with target illumination and uniform illumination are generated from the reflectance dataset by matching the face orientation [37] of the target and reference subjects. The quotient image of the source subject is warped to

Figure 2.19: Illumination transferring with known geometry, figure from [47].

the target subject, then multiplied to the uniform illuminated target image to achieve the relighting. An example is shown in figure 2.20.

Figure 2.20: Face relighting using reflectance transfer, figure from [61]. From left to right: the reference subject, the target subject, the warped quotient image, the target subject with similar illumination as the reference subject.

Illumination transfer problem has also attracted a lot of attention recently. However, many of these still focus on a class of object, *e.g.* human faces. By assuming radiance to be locally homogeneous, a global optimization is proposed in [17] to generate a smooth image, while texture is preserved by local constraints during illumination

Figure 2.21: Face illumination transferring with locally constrained global optimization, figure from [17]. The first row shows reference images, the second row shows the target image to be relit and the relighting results.

Figure 2.22: Face illumination transferring with edge-preserving filters, figure from [18]. The first row shows reference images, the second row shows the target image to be relit and the relighting results.

transfer as shown in figure 2.21. In [18], they assume illumination effects are mostly captured by the large-scale layer of the image. After face alignment, edge-preserving filters are used to decompose the images into large-scale and detail layers. Relighting is achieved by combining the detail layer of the target image containing facial features and the large-scale layer of reference image catching the illumination effects, as shown in figure 2.22.

Chapter 3

Single Image Relighting

3.1 Framework Overview

These are some key considerations made in choosing the example-based approach:

- We intend to avoid explicit 3D model acquisition as required in geometry-based relighting methods.
- We also intend to avoid tedious laboratory setup as required for existing image-based relighting methods, such as those used in the movie industry.
- Some example-based relighting research has been carried out; however there remains much that has not yet been done.

We start by building the framework for a single input image. As limited information is provided by one image, some assumptions have to be made:

- A single directional light source.
- Each surface point can be locally approximated by a spherical patch.

- The surface reflectance function can be adequately modeled as Lambertian.

Inspired by image analogies [34], our method is illustrated in figure 3.1. While the image

Figure 3.1: Framework overview: for every pixel of the input image B , we search the best match by intensity and gradient in the example image A with the same illumination condition, after we find the best match, we do a lookup what it looks like in the example image A_p with another illumination condition, then assign that value to the result B_p .

analogies work transfers filtering operations, our method adopts the focus on image relighting. It is presented in a computer vision fashion (with focus on accuracy) versus a graphics fashion (with focus on visual results). We provide the mathematical and experimental analyses on the proposed method against the physics-based simulation.

3.1.1 Preliminaries

Let l, l' represent two different lighting conditions. Suppose we have an image I_B of a scene B captured under illumination condition l , and we want to synthesize image I'_B of the same scene illuminated under lighting condition l' . Suppose we also have example images $\{I_{A1}, I_{A2}, \dots\}$ of different scenes $\{A_1, A_2, \dots\}$ but with the same lighting condition as in image I_B and corresponding images $\{I'_{A1}, I'_{A2}, \dots\}$ of these scenes taken under illumination l' . Intuitively, if we have enough examples, we can find identical elements of I_B in $\{I_{A1}, I_{A2}, \dots\}$, then the result I'_B can be synthesized from corresponding related elements in $\{I'_{A1}, I'_{A2}, \dots\}$. The primary problems are: how many example pairs do we need, and how can we find the correct correspondence between the images under the same illumination condition? We will show that even one example pair can be useful, provided that the previously mentioned assumptions hold.

Formally, we establish the following equations:

$$\begin{aligned} I_A(u_a, v_a) &= \lambda \rho_A N_A \cdot L & I'_A(u_a, v_a) &= \lambda' \rho_A N_A \cdot L' \\ I_B(u_b, v_b) &= \lambda \rho_B N_B \cdot L & I'_B(u_b, v_b) &=? \end{aligned} \quad (3.1)$$

Where λ and ρ are illumination intensity and reflectivity coefficient respectively, N is the unit vector of surface normal, L is unit vector in lighting direction and (u, v) is a particular pixel position in image I . The subscripts A and B represent two different scenes, while the two lighting conditions are distinguished by the presence/absence of the prime symbol. The goal is to compute I'_B given the three observable images of I_A , I'_A and I_B . The main challenge in this problem is that the parameters of λ , ρ , N and L are all unknown for both scenes and both lighting conditions.

3.1.2 Similarity Matching

The following two sections will show how the similarities between two images under the same illumination condition can be captured. With the assumptions which we made in the previous section, correspondences between scene patches in images I_A and I_B are searched for by matching image intensities and image gradients. The surface normal of a patch is parameterized by its slant σ and its tilt τ as shown in figure 3.2, where slant is defined as the angle between the surface normal and the optical axis (viewing direction), and tilt is defined as the angle between the projection of the surface normal on the plane perpendicular to the optical axis and one horizontal axis on that plane. Two pixels (in the two images) are considered matched if their feature vectors are Euclidean nearest neighbors to each other in a 3D feature space representing image intensity and gradient. Following the approach of Lee and Rosenfeld [44], analysis is carried out in light source coordinate system.

Figure 3.2: Surface normal is parameterized by its σ and its tilt τ .

Image Intensity

We will show that by matching image intensity, we can find the corresponding patches which have the same surface slant in the two scenes. Formally, we are considering the case when we have found (u_a^*, v_a^*) such that $I_B(u_b, v_b) = I_A(u_a^*, v_a^*)$.

Under the assumption of Lambertian surface:

$$I = \lambda \rho N \cdot L, \quad (3.2)$$

where I is the image intensity or measured image irradiance, λ and ρ are the scalars representing illumination intensity and reflectivity coefficient (assume the surface has uniform reflectivity), N is the unit surface normal vector, and L is the unit vector pointing in the light source direction.

The surface slant σ is the angle between the surface normal and the viewing direction; in the illumination coordinate system. It can be calculated as $\cos^{-1}(N \cdot L)$.

As $I = \lambda \rho N \cdot L$,

$$\sigma = \cos^{-1} \left(\frac{I}{\lambda \rho} \right). \quad (3.3)$$

When we match the intensities, $I_A = I_B$, so $\frac{I_A}{\lambda \rho} = \frac{I_B}{\lambda \rho}$, assuming the $\lambda \rho$ is a constant for different objects. However we also consider the possibility of matching normalized intensities $\frac{I_A}{\lambda \rho_A} = \frac{I_B}{\lambda \rho_B}$, assuming the brightest pixel values in the two images are $\lambda \rho_A$ and $\lambda \rho_B$ respectively and use these gray levels as normalization factors for the images of different objects. Hence

$$\sigma_A = \cos^{-1} \left(\frac{I_A}{\lambda \rho_A} \right) = \cos^{-1} \left(\frac{I_B}{\lambda \rho_B} \right) = \sigma_B, \quad (3.4)$$

as σ_A and $\sigma_B \in [0, \frac{\pi}{2}]$, we implicitly match the surface slants in the illumination coor-

dinate system.

The relation of intensity derivatives and surface tilt is much more complicated. Here, surface tilt is defined as the angle between the projection of the surface normal on the plane perpendicular to the optical axis and one horizontal axis on that plane. It is shown in Pentland [62] that if the local patch can be approximated by an umbilical patch for which the two principal curvatures are same, then the tilt τ is given by:

$$\tau = \tan^{-1} \left(\frac{-\left(\frac{\partial^2 I}{\partial x^2} - \frac{\partial^2 I}{\partial y^2}\right) + \sqrt{\left(\frac{\partial^2 I}{\partial x^2} - \frac{\partial^2 I}{\partial y^2}\right)^2 + 4\left(\frac{\partial^2 I}{\partial x \partial y}\right)^2}}{2\frac{\partial^2 I}{\partial x \partial y}} \right). \quad (3.5)$$

However high order derivatives are quite sensitive to noise, so we follow Lee and Rosenfeld's work [44] to investigate matching of image gradients, which are first order derivatives.

Image Gradient

We will show here that by matching image gradients, the corresponding patches will have the same surface tilt in the two scenes. Formally, we are considering the case when we have found (u_a^*, v_a^*) such that $\nabla I_B(u_b, v_b) = \nabla I_A(u_a^*, v_a^*)$.

Assume each surface patch can be approximated by a spherical patch. Then

$$z = \sqrt{R^2 - x^2 - y^2}, \quad (3.6)$$

$$\frac{\partial z}{\partial x} = -x \frac{1}{\sqrt{R^2 - x^2 - y^2}} = -\frac{x}{z}, \quad (3.7)$$

$$\frac{\partial z}{\partial y} = -y \frac{1}{\sqrt{R^2 - x^2 - y^2}} = -\frac{y}{z}, \quad (3.8)$$

where R is the radius of the approximated spherical patch. For convenience, the positive

solution of z is chosen here. The relation between the sign of z and the surface type (convex/concave) will be shown later.

The surface normal is perpendicular to the surface patch, and can be determined by the depth gradient $\left(\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right)^T$. Consider a step on the surface from (x, y, z) to some other nearby point $(x + \delta x, y + \delta y, z + \delta z)$ on the surface:

$$\delta z \approx \frac{\partial z}{\partial x} \delta x + \frac{\partial z}{\partial y} \delta y. \quad (3.9)$$

Hence

$$\begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial z}{\partial x} \\ \frac{\partial z}{\partial y} \\ -1 \end{bmatrix} \approx 0. \quad (3.10)$$

The vector $\left(\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}, -1\right)^T$ is perpendicular to the surface and in the direction of the surface normal.

The unit normal vector is:

$$N = \frac{1}{\sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 + 1}} \begin{bmatrix} \frac{\partial z}{\partial x} \\ \frac{\partial z}{\partial y} \\ -1 \end{bmatrix} = \frac{z}{R} \begin{bmatrix} -x\frac{1}{z} \\ -y\frac{1}{z} \\ -1 \end{bmatrix} = \frac{1}{R} \begin{bmatrix} -x \\ -y \\ -z \end{bmatrix}. \quad (3.11)$$

Let the illumination direction L be parameterized with slant and tilt (σ_L, τ_L) .

So (3.2) becomes:

$$\begin{aligned}
 I &= \lambda \rho N \cdot L \\
 &= \lambda \rho \frac{1}{R} \begin{bmatrix} -x \\ -y \\ -z \end{bmatrix} \cdot \begin{bmatrix} \sin \sigma_L \cdot \cos \tau_L \\ \sin \sigma_L \cdot \sin \tau_L \\ \cos \sigma_L \end{bmatrix} \\
 &= \frac{\lambda \rho}{R} (-x \sin \sigma_L \cdot \cos \tau_L - y \sin \sigma_L \cdot \sin \tau_L - z \cdot \cos \sigma_L).
 \end{aligned} \tag{3.12}$$

To transform the original coordinate system to the image plane, under assumption of weak perspective projection, we apply:

$$u = f \frac{x}{z_0}, \tag{3.13}$$

$$v = f \frac{y}{z_0}, \tag{3.14}$$

where f is the focal length, z_0 is the constant average depth which is large compared to the object's dimensions.

The first derivatives of the image intensity on the image plane are:

$$\frac{\partial I}{\partial u} = \frac{z_0}{f} \frac{\partial I}{\partial x} = \frac{z_0}{f} \frac{\lambda \rho}{R} \left(-\sin \sigma_L \cdot \cos \tau_L + \frac{x}{z} \cos \sigma_L \right), \tag{3.15}$$

$$\frac{\partial I}{\partial v} = \frac{z_0}{f} \frac{\partial I}{\partial y} = \frac{z_0}{f} \frac{\lambda \rho}{R} \left(-\sin \sigma_L \cdot \sin \tau_L + \frac{y}{z} \cos \sigma_L \right). \tag{3.16}$$

Having introduced the key parameters in this framework, we now describe their status. The main unknown parameter we want to compute is the surface tilt τ , or equivalently the surface normal N , part of which encodes the surface slant σ recovered via image intensity matching in Section 3.1.2. The observables here are the image intensity I and the image gradient $\left(\frac{\partial I}{\partial u}, \frac{\partial I}{\partial v} \right)^T$, which we will assume can be reasonably

measured from the images. Subsequently, we will show that τ can be computed without the need to recover the unknown light parameters $(\sigma_L, \tau_L, \lambda)$, the other surface parameters of ρ and R , the object distance z_0 , or the focal length f .

First consider a special case here where the scene is illuminated by a head light, i.e. the illumination direction is align with the viewing direction. Then $\sin \sigma_L = 0, \cos \sigma_L = 1$, so

$$\frac{\partial I}{\partial v} / \frac{\partial I}{\partial u} = \frac{\frac{z_0}{f} \frac{\partial I}{\partial y}}{\frac{z_0}{f} \frac{\partial I}{\partial x}} = \frac{\frac{y}{z}}{\frac{x}{z}} = \frac{y}{x}, \quad (3.17)$$

$$\tau = \tan^{-1} \left(\frac{n_y}{n_x} \right) = \tan^{-1} \left(\frac{-\frac{y}{R}}{-\frac{x}{R}} \right) = \tan^{-1} \frac{y}{x} = \tan^{-1} \left(\frac{\partial I}{\partial v} / \frac{\partial I}{\partial u} \right). \quad (3.18)$$

If two surface points have the same image gradient:

$$\begin{bmatrix} \frac{\partial I_A}{\partial u} \\ \frac{\partial I_A}{\partial v} \end{bmatrix} = \begin{bmatrix} \frac{\partial I_B}{\partial u} \\ \frac{\partial I_B}{\partial v} \end{bmatrix}, \quad (3.19)$$

then

$$\tau_A = \tan^{-1} \left(\frac{\partial I_A}{\partial u} / \frac{\partial I_A}{\partial v} \right) = \tan^{-1} \left(\frac{\partial I_B}{\partial u} / \frac{\partial I_B}{\partial v} \right) = \tau_B. \quad (3.20)$$

The tilts of the two surface patches are the same in the illumination coordinate system which is aligned with the viewer coordinate system for head light.

For general cases, let the transformation from the viewer coordinate system to the illumination coordinate system be:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{pmatrix} \cos \sigma_L \cdot \cos \tau_L & \cos \sigma_L \cdot \sin \tau_L & -\sin \sigma_L \\ -\sin \tau_L & \cos \tau_L & 0 \\ \sin \sigma_L \cdot \cos \tau_L & \sin \sigma_L \cdot \sin \tau_L & \cos \sigma_L \end{pmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad (3.21)$$

where the illumination coordinate system is defined by axes x', y', z' :

$$\mathbf{x}' = \begin{bmatrix} \cos \sigma_L \cdot \cos \tau_L \\ \cos \sigma_L \cdot \sin \tau_L \\ -\sin \sigma_L \end{bmatrix}, \mathbf{y}' = \begin{bmatrix} -\sin \tau_L \\ \cos \tau_L \\ 0 \end{bmatrix}, \mathbf{z}' = \begin{bmatrix} \sin \sigma_L \cdot \cos \tau_L \\ \sin \sigma_L \cdot \sin \tau_L \\ \cos \sigma_L \end{bmatrix}. \quad (3.22)$$

Note that the z' axis is in the illumination direction L .

The surface normal $N'(n'_x, n'_y, n'_z)$ in the illumination coordinate system is:

$$\begin{bmatrix} n'_x \\ n'_y \\ n'_z \end{bmatrix} = \begin{pmatrix} \cos \sigma_L \cdot \cos \tau_L & \cos \sigma_L \cdot \sin \tau_L & -\sin \sigma_L \\ -\sin \tau_L & \cos \tau_L & 0 \\ \sin \sigma_L \cdot \cos \tau_L & \sin \sigma_L \cdot \sin \tau_L & \cos \sigma_L \end{pmatrix} \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix}. \quad (3.23)$$

Let u', v' be the axes of the image plane in the illumination coordinate system, then the image gradient becomes:

$$\begin{aligned} \begin{bmatrix} \frac{\partial I}{\partial u'} \\ \frac{\partial I}{\partial v'} \end{bmatrix} &= \frac{z_0}{f} \begin{bmatrix} \frac{\partial I}{\partial x'} \\ \frac{\partial I}{\partial y'} \end{bmatrix} \\ &= \begin{pmatrix} \cos \sigma_L \cdot \cos \tau_L & \cos \sigma_L \cdot \sin \tau_L \\ -\sin \tau_L & \cos \tau_L \end{pmatrix} \frac{z_0}{f} \begin{bmatrix} \frac{\partial I}{\partial x} \\ \frac{\partial I}{\partial y} \end{bmatrix} \\ &= \begin{pmatrix} \cos \sigma_L \cdot \cos \tau_L & \cos \sigma_L \cdot \sin \tau_L \\ -\sin \tau_L & \cos \tau_L \end{pmatrix} \begin{bmatrix} \frac{\partial I}{\partial u} \\ \frac{\partial I}{\partial v} \end{bmatrix}. \end{aligned} \quad (3.24)$$

The ratio of the first derivatives is:

$$\begin{aligned}
\frac{\partial I}{\partial v'} / \frac{\partial I}{\partial u'} &= \frac{-\sin \tau_L \frac{\partial I}{\partial u} + \cos \tau_L \frac{\partial I}{\partial v}}{\cos \sigma_L \cdot \cos \tau_L \frac{\partial I}{\partial u} + \cos \sigma_L \cdot \sin \tau_L \frac{\partial I}{\partial v}} \\
&= \frac{-\sin \tau_L \frac{\partial I}{\partial x} + \cos \tau_L \frac{\partial I}{\partial y}}{\cos \sigma_L \cdot \cos \tau_L \frac{\partial I}{\partial x} + \cos \sigma_L \cdot \sin \tau_L \frac{\partial I}{\partial y}} \\
&= \frac{-\sin \tau_L \cos \sigma_L \frac{x}{z} + \cos \tau_L \cos \sigma_L \frac{y}{z}}{-\cos \sigma_L \sin \tau_L + \cos^2 \sigma_L \cos \tau_L \frac{x}{z} + \cos^2 \sigma_L \sin \tau_L \frac{y}{z}} \\
&= \frac{-x \sin \tau_L + y \cos \tau_L}{-z \sin \sigma_L + x \cos \sigma_L \cos \tau_L + y \cos \sigma_L \sin \tau_L}.
\end{aligned} \tag{3.25}$$

By definition, the surface tilt τ in the illumination coordinate system is:

$$\begin{aligned}
\tau &= \tan^{-1} \left(\frac{n'_y}{n'_x} \right) \\
&= \tan^{-1} \frac{\frac{x}{R} \sin \tau_L - \frac{y}{R} \cos \tau_L}{-\frac{x}{R} \cos \sigma_L \cos \tau_L - \frac{y}{R} \cos \sigma_L \sin \tau_L + \frac{z}{R} \sin \sigma_L} \\
&= \tan^{-1} \frac{x \sin \tau_L - y \cos \tau_L}{-x \cos \sigma_L \cos \tau_L - y \cos \sigma_L \sin \tau_L + z \sin \sigma_L} \\
&= \tan^{-1} \left(\frac{\partial I}{\partial v'} / \frac{\partial I}{\partial u'} \right).
\end{aligned} \tag{3.26}$$

Note that the arctangent function leads to two possible tilts, which are π apart such that one corresponds to a convex surface and the other to a concave surface.

In this instance, because we can only measure the image gradients in the viewer coordinate system and we do not assume to know the illumination direction, we can not compute the surface tilt directly.

In terms of image gradient matching, if two surface points have same first image intensity derivatives:

$$\begin{bmatrix} \frac{\partial I_A}{\partial u} \\ \frac{\partial I_A}{\partial v} \end{bmatrix} = \begin{bmatrix} \frac{\partial I_B}{\partial u} \\ \frac{\partial I_B}{\partial v} \end{bmatrix}, \tag{3.27}$$

as the illumination parameters (σ_{LA}, τ_{LA}) and (σ_{LB}, τ_{LB}) are assumed to be the same,

from (3.24), the image gradients in the illumination coordinate system are also the same:

$$\begin{bmatrix} \frac{\partial I_A}{\partial u'} \\ \frac{\partial I_A}{\partial v'} \end{bmatrix} = \begin{bmatrix} \frac{\partial I_B}{\partial u'} \\ \frac{\partial I_B}{\partial v'} \end{bmatrix}. \quad (3.28)$$

Then

$$\tau_A = \tan^{-1} \frac{\partial I_A}{\partial u'} / \frac{\partial I_A}{\partial v'} = \tan^{-1} \frac{\partial I_B}{\partial u'} / \frac{\partial I_B}{\partial v'} = \tau_B, \quad (3.29)$$

which means the tilts of the two surface patches are the same in the illumination (and in fact any other) coordinate system, even if we are unable to compute the tilt directly. However, as mentioned before, τ_A and τ_B may be π apart due to the convex/concave ambiguity. This problem will be analyzed in Section 4.1. Here we assume the surfaces are convex.

In summary, two surface points having the same image intensity (normalized intensity) implies that they have the same surface slant σ , and if the two surface points have the same image gradient, it implies that they share the same surface tilt τ . As the surface normal direction is determined by these two parameters σ and τ , two object patches with same σ and τ have same surface normal direction.

Note that the equations are derived in their own illumination coordinate system. However, as the illumination conditions of I_A and I_B are the same, the geometric transformation from the illumination coordinates to the viewer coordinates are also the same for the two images. The equality of the surface normal directions is therefore invariant to the coordinate system used.

3.1.3 Relationship to Direct Physics-based Emulation

As listed in Section 3.1.1, the initial setup of the example-based framework can be represented as follows:

$$\begin{aligned} I_A(u_a, v_a) &= \lambda \rho_A N_A \cdot L & I'_A(u_a, v_a) &= \lambda' \rho_A N_A \cdot L' \\ I_B(u_b, v_b) &= \lambda \rho_B N_B \cdot L & I'_B(u_b, v_b) &=? \end{aligned} \quad (3.30)$$

where the three observed images are: I_A and I_B of the two scenes A and B captured under the identical illumination condition $l(\lambda, L)$, and image I'_A of scene A captured under another illumination condition $l'(\lambda', L')$. The image I'_B of an unobserved scene B under the illumination condition l' is to be synthesized from these three images.

A physics-based formulation to compute I'_B is simply:

$$I'_B(u, v) = \lambda' \rho_B N(\sigma_B, \tau_B) \cdot L', \quad (3.31)$$

where the different parameters either have to be known, measured or computed. We compare this to the analogy-based formulation which is essentially an indirect lookup operation, expressed as:

$$I'_B(u, v) = I'_A(u^*, v^*), \quad (3.32)$$

where (u^*, v^*) is the image point in I'_A corresponding to the image point (u, v) in I'_B , as obtained via analogy to the matched pixels of (u^*, v^*) in I_A and (u, v) in I_B . As described in Section 3.1.2, the matched point (u^*, v^*) is found such that the corresponding image intensities and gradients in I_A and I_B are identical, i.e.

$$I_B(u, v) = I_A(u^*, v^*), \quad (3.33)$$

$$\frac{\partial I_B}{\partial u} = \frac{\partial I_A}{\partial u} \bigg|_{\substack{u=u^* \\ v=v^*}}, \quad (3.34)$$

$$\frac{\partial I_B}{\partial v} = \frac{\partial I_A}{\partial v} \bigg|_{\substack{u=u^* \\ v=v^*}}. \quad (3.35)$$

The goal is to show the situations when the physics-based formulation and the analogy-based formulation lead to the same result.

If image intensities of two pixels are equal:

$$\begin{aligned} I_A = I_B &\implies \lambda \rho_A N_A \cdot L = \lambda \rho_B N_B \cdot L \\ &\implies \rho_A \cos \sigma_A = \rho_B \cos \sigma_B. \end{aligned} \quad (3.36)$$

If image gradients of two pixels are equal, then with the aid of (3.28) and (3.29), we get:

$$\begin{aligned} \left. \begin{aligned} \frac{\partial I_A}{\partial u} &= \frac{\partial I_B}{\partial u} \\ \frac{\partial I_A}{\partial v} &= \frac{\partial I_B}{\partial v} \end{aligned} \right\} &\implies \frac{\partial I_A}{\partial v'} / \frac{\partial I_A}{\partial u'} = \frac{\partial I_B}{\partial v'} / \frac{\partial I_B}{\partial u'} \\ &\implies \tau_A = \tau_B. \end{aligned} \quad (3.37)$$

Note that the surface point parameters between I_A and I'_A are identical, since only the illumination condition has changed, likewise for I_B and I'_B .

Using (3.36) and (3.37) in (3.32), it follows that:

$$\begin{aligned} I'_B &= I'_A(u^*, v^*) \\ &= \lambda' \rho_A N(\sigma_A, \tau_A) \cdot L' \\ &= \lambda' \rho_B \frac{\cos \sigma_B}{\cos \sigma_A} N \left(\cos^{-1} \left(\frac{\rho_B}{\rho_A} \cos \sigma_B \right), \tau_B \right) \cdot L'. \end{aligned} \quad (3.38)$$

If the surface albedos are equal, $\rho_A = \rho_B$, then from (3.36), $\cos \sigma_A = \cos \sigma_B$. Since

$\sigma \in [0, \frac{\pi}{2}]$ for outward normals, $\sigma_A = \sigma_B$. Therefore (3.38) can be further written as:

$$\begin{aligned} I'_B &= \lambda' \rho_B N(\cos^{-1}(\cos \sigma_B), \tau_B) \cdot L' \\ &= \lambda' \rho_B N(\sigma_B, \tau_B) \cdot L'. \end{aligned} \quad (3.39)$$

As (3.31) and (3.39) are identical, we prove that the analogy-based approach leads to physics-based results.

Note that the derivation is based on an assumption that the surface can be approximated by a spherical patch in order to get surface tilt. A more general model by Ferrie and Levine [25] states that the intensity gradient direction will be the same as the tilt vector in illumination coordinate if and only if

$$\frac{(\frac{\partial f}{\partial x})^2 - (\frac{\partial f}{\partial y})^2}{\frac{\partial f}{\partial x} \frac{\partial f}{\partial y}} = \frac{\frac{\partial^2 f}{\partial x^2} - \frac{\partial^2 f}{\partial y^2}}{\frac{\partial^2 f}{\partial x \partial y}}, \quad (3.40)$$

where the surface is parameterized as $z = f(x, y)$. They related the deviations between actual surface tilt and the surface tilt estimated from image intensities to the deviation of the surface from a spherical model, as shown in figure 3.3. In Section 4.1, we will further analyze the relation of the relighting error and local surface shape.

Figure 3.3: Tilt estimation error as a function of eccentricity, figure from [25]

3.2 Experiments

3.2.1 Synthetic Images

Our framework is initially demonstrated on a few simple shapes, namely a sphere, a cylinder and an ellipsoid in figures 3.4, 3.5 and 3.6 respectively. In each experiment, we used the spheres as the analogy exemplars. Three illumination conditions with directional lights were considered: the light was either from the rear of the viewer and aligned with the viewing direction (“head lighted”), or the light was from the left or right and perpendicular to the viewing direction (“side-lighted”). The experiments involved converting one illumination condition to another through analogy-based relighting.

Images A and B are each associated with a coordinate map. It illustrates the best per-pixel match between image A and image B , essentially the corresponding points between the reference and test objects, which are the same color. In the coordinate map for image B , the color of each pixel encodes the corresponding pixel position in image A . The level of the green component in the color represents the x coordinate (low-to-high levels mapping to left-to-right positions), while the level of the red component represents the y coordinate (low-to-high levels mapping to top-to-bottom positions). Image A_p is the reference image with the alternate illumination condition, under which image B_p is generated through the analogy-based approach, while image B_{p-real} is synthesized using a physics-based approach. An average absolute error image is also shown. The results show that the analogy-based approach can relight these simple objects well enough that the average difference between the physics-based lighting results and example-based relighting results are under 10 intensity levels out of 255 levels with only very few noticeable artifacts.

Figure 3.4: Relighting sphere with sphere example. Images A and Ap are the reference pair, images B and Bp are the input and result. The ground truth and error images are shown in the middle and on the right of each subfigure respectively. The coordinate maps $Acoords$ and $Bcoords$ are introduced to show the corresponding points between the reference and test objects, which are found by the algorithm and assigned the same color.

We also applied the method on three human faces for which 3D models are available. Previous face relighting research usually assume the face model is mostly convex and smooth. However, faces at higher levels of detail will contain more than just convex patches; not only concave surface patches, but also saddle surface patches. The relation of the local shape and the relighting error is discussed in Section 4.1.

Although the size or the radius of the sphere is not important from the earlier theoretical analysis, quantization errors are present in the image intensities and estimation of image gradients. We can alleviate this problem by using a few example pairs in different scales, as we can see from in figure 3.7. Figures 3.7b, 3.7d and 3.7f show

Figure 3.5: Relighting cylinder with sphere example. Similar as figure 3.4.

results of experiments in which multiple scaled versions of image A and image A_p are provided as reference images. The corresponding matched scale for each pixel in image B is encoded in a scale level image, in which a brighter level maps to a larger scale.

3.2.2 Real Images

We also tested the framework on real images from Yale Face Database B [30]. The tests cover cases of relighting the face of a person in different poses, and relighting the faces of different people. As the real images contain textures and noises, surface patches with the same surface slants and tilts will not always be correctly matched, due to deviations from our framework which assumes constant albedo. As can be seen from the B_p results in figure 3.8, the direct transfer results are not quite smooth and are less visually pleasing. We introduce a gradient domain reconstruction stage described

Figure 3.6: Relighting ellipsoid with sphere example. Similar as figure 3.4.

in Section 3.3 to refine the results.

3.3 Smoothness Constraints

In order to improve the visual quality of the synthesis, we use Markov random field and Poisson reconstruction.

3.3.1 Markov Random Field

Similar to other image synthesis works, like [24], a Markov random field model is assumed. So the conditional distribution of a pixel only depends on its local neighbour pixels. Given all the neighbours synthesized so far, it can be estimated by querying the

(a) Relight male image with single-scale example image

(b) Relight male image with multi-scales example images

(c) Relight female image with single-scale example image

Figure 3.7

(d) Relight female image with multi-scales example images

(e) Relight female2 image with single-scale example image

(f) Relight female2 image with multi-scales example images

Figure 3.7: Relighting complex models with sphere example. Similar as figure 3.4.

Figure 3.8: Relighting real images with examples. Similar as figure 3.4 except the Poisson reconstructed image $Bp\text{-recons}$.

example image under the same lighting condition and finding similar neighbours. While finding the best match of intensity and derivatives for each pixel, we also included the similar distances between the neighbours of that pixel in the objective function to be minimized.

3.3.2 Standard Poisson Solution

Psychology-based research [42] on the human visual system shows that humans barely notice very gradual changes in image intensities, even if the total intensity change is large. Conversely, humans are very sensitive to small but sharp intensity changes that correspond to large gradients in the image. Computationally, we can relate this to the Laplacian operator ∇^2 on the image, which suppresses slow changes of intensity and retains large image gradients.

In the analogy-based method proposed in Section 3.1.2, correspondences are found between the exemplar pair of images I_A and I_B (with different scenes but having the same lighting condition) by matching both image intensities and gradients. However, the method reconstructs the desired image I'_B by directly copying intensities via mapped lookup from I'_A in a point-wise manner, without consideration for the resultant image gradients in I'_B . Because gradients are first order derivatives and sensitive to noise, small errors in the intensities may lead to large gradient noise, to which humans are sensitive.

Alternatively, one may consider a slightly modified approach where rather than copying intensities directly from image I'_A to I'_B , the image *gradients* of I'_A are copied to form a gradient field for I'_B instead. An attempt may then be made to compute image I'_B by integrating the gradient field, and since the gradients in I'_B are now deliberately specified, we can expect there would be less visually-sensitive gradient noise. However, this gradient field for I'_B , created via point-by-point lookup from I'_A , is very unlikely to result in a 2D-consistent gradient field that is integrable.

A method to circumvent this problem is to find an optimal image such that its derived (and hence integrable) gradient field is as close as possible to the target gradient field generated from lookup via I'_A . It turns out that a solution may be found by solving a Poisson equation, wherein the Laplacian of the image must be equal to the divergence of the target gradient field. Details are provided below.

As the intensity of an image is a function $I : R^2 \rightarrow R$, the gradient of I is defined by its two first partial derivatives:

$$(G_u, G_v) = \nabla I = \left(\frac{\partial I}{\partial u}, \frac{\partial I}{\partial v} \right), \quad (3.41)$$

where G_u and G_v are the gradients in u direction and v direction respectively.

As described in Section 3.1.3, we generate the relighting image via an indirect lookup operation $I'_B(u, v) = I'_A(u^*, v^*)$. Similarly, a 2D gradient field G'_B can be also obtained by assigning the corresponding gradient from G'_A , formally:

$$G'_B(u, v) = G'_A(u^*, v^*), \quad (3.42)$$

where G'_A is the gradient field of image I'_A .

Then we can reconstruct an image from this gradient field G via integration. However, as G has been constructed in a point-wise manner, in general it cannot be integrated back into a 2D scalar field that is the image. Hence the goal is to recover a 2D scalar function I such that the gradient of I , ∇I , best approximates G in a least squares manner:

$$I^* = \operatorname{argmin}_I \int \int \|\nabla I - G\|^2 dudv. \quad (3.43)$$

According to the Euler-Lagrange equation, the optimal I^* must satisfy the Poisson equation:

$$\nabla^2 I = \nabla \cdot G, \quad (3.44)$$

$$\frac{\partial^2 I}{\partial u^2} + \frac{\partial^2 I}{\partial v^2} = \frac{\partial G}{\partial u} + \frac{\partial G}{\partial v}. \quad (3.45)$$

It can be rapidly solved by the Fourier Cosine Transform [65] as:

$$I(u, v) = \text{ICT}(\hat{I}_{p,q}) = \text{ICT} \left(\frac{\text{CT} \left(\frac{\partial G}{\partial u} + \frac{\partial G}{\partial v} \right)}{2 \left(\cos \left(\frac{\pi p}{W} \right) + \cos \left(\frac{\pi q}{H} \right) - 2 \right)} \right), \quad (3.46)$$

where CT and ICT are the cosine transform and inverse cosine transform respectively, applied to the image of width W and height H , while p and q are spatial frequencies in u and v direction respectively.

The solution via cosine transform imposes a boundary condition of zero gradient,

i.e. the Neumann boundary condition. This is acceptable for the test data as the background of the images are mainly homogeneous and the gradients are quite small there.

We tested the framework on real images from the Yale Face Database B and the Extended Yale Face Database B [30]. In figure 3.8, images A and A_p are the example pair, image B is the image to be relit, image B_p is the direct result from analogy, image $B_p\text{-recons}$ is the Poisson reconstructed result, and image $B_p\text{-real}$ is the real image under the same illumination condition as image A_p . We can see there are a few differences between the synthesized results and the real images. However, the synthesized results can still be considered as reasonably reflecting the change in the illumination. Fidelity to the desired intensities is less important, as psychologically humans are more sensitive to gradient artifacts which are suppressed during Poisson reconstruction.

3.3.3 Screened Poisson Solution

We demonstrate that the gradient field computed via analogy can be used to reconstruct visually pleasing results with the standard Poisson solution as images $B_p\text{-R}$ in figure 3.9. However, the raw intensities computed via analogy can further be accounted for in the screened Poisson solution [12]. We can get a better solution by keeping the result close to both the raw intensities and the gradient field as:

$$I^* = \operatorname{argmin}_I \iint \lambda(I - D)^2 + \|\nabla I - G\|^2 dudv, \quad (3.47)$$

where λ is the trade-off coefficient and D is the data function. In our case, D represents the raw intensities obtained via direct transfer. The optimal I^* must satisfy:

$$\lambda I - \nabla^2 I = \lambda D - \nabla \cdot G. \quad (3.48)$$

Similar to (3.46), the Fourier domain solution is:

$$I(u, v) = \text{ICT}(\hat{I}_{p,q}), \quad (3.49)$$

where

$$\hat{I}_{p,q} = \frac{\text{CT}(\lambda D - (\frac{\partial G}{\partial u} + \frac{\partial G}{\partial v}))}{\lambda - 2(\cos(\frac{\pi p}{W}) + \cos(\frac{\pi q}{H}) - 2)}. \quad (3.50)$$

In (3.46), $\hat{I}_{0,0}$ is undefined, lead to an unknown offset, however, in screened Poisson solution (3.50):

$$\hat{I}_{0,0} = \hat{D}_{0,0}. \quad (3.51)$$

The problem of the unknown offset in standard Poisson is solved even with a very small positive λ , in our experiments, where it is usually set to 0.01.

The comparison of the results is shown in figure 3.9. As the real images contain textures and noises, surface patches with the same surface slants and tilts will not always be correctly matched, duo to deviations from our framework which assume constant albedo. As can be seen from the image B_p , the direct result is not quite smooth and visually pleasing. As mentioned previously, this is the result of noise due to non-convex patches, quantization errors and albedo variation. By applying the standard Poisson solution, we can generate a more realistic result as image B_p-R . Neumann boundary condition is reasonable satisfied in some images, *e.g.* the uncropped ones from Yale Face Database B, where the subject is surround by dark background, there are only small gradients at the boundary. However, in general, the screened

Poisson solution further improves the result, with respect to real image $B_p\text{-real}$, the mean intensity error of image $B_p\text{-SR}$ is reduced even compared to the mean error of image $B_p\text{-R}$ with an optimal offset ($B_p\text{-RwO}$). It is well known that both approaches can result in color shifting, however, we can prevent this problem by only applying these methods on luminance channel of the images. Because we focus on the effects caused by the change of lighting direction, the albedo of the object (color-opponent channels) can be assumed unchanged. We can see there are some differences between the synthesized results and the real images. However, the synthesized results can still be considered as reasonably reflecting the change in illumination. The final result $B_p\text{-SRH2B}$ is produced by matching the histogram of image $B_p\text{-SR}$ to the histogram of image B , so as to overcome albedo variation and thus improve the perception quality, as surface reflectance perception is correlated with image statistics [51].

3.4 Computational Efficiency

The most computationally expensive part of this method is nearest neighbor search in high-dimensional space. Tree-based methods are relatively efficient in our case, as we only need intensity and gradient of every single pixel, and thus the feature space is low dimensional.

With minor loss in accuracy, approximate algorithms provide large speedups. Arya and Mount [4] modified the original kd-tree algorithm for approximate matching. A point $p \in X$ is a so called ϵ -approximate nearest neighbor of a query point $q \in X$, if $\text{dist}(p, q) \leq (1+\epsilon)\text{dist}(p^*, q)$, where p^* is the true nearest neighbor. They also proposed to use a priority queue to accelerate the search by visiting tree nodes according to their distance to the query point.

Recently, Silpa-Anan and Hartley [72] proposed to utilize multiple randomized kd-

Figure 3.9: Relighting real images with Poisson solutions. From left to right, the second row shows the intensity transfer result; the standard Poisson result from gradient transfer (equivalent to set $\lambda = 0$ in screened Poisson solution); the standard Poisson result from gradient transfer with optimal offset w.r.t. real lighting result $B_p\text{-real}$; the screened Poisson result from intensity and gradient transfer (set $\lambda = 0.01$); and the screened Poisson result after it is histogram-matched to image B . The third row shows the differences between the results in the second row and the real lighting result. The numbers indicate the average intensity errors.

trees to speed up approximate nearest neighbor search, and Muja and Lowe [52] set up a framework to automatically determine the best algorithm between multiple randomized kd-trees and hierarchical k-means tree, as well as the best parameter values.

The current speed of a mixed Matlab and C implementation is around 1s for 150

by 150 sized images on a workstation with a 2.66GHz Intel Core2 Processor.

3.5 Conclusion

In this chapter, we presented our initial investigation on the analogy framework with assumptions of Lambertian model, uniform albedo and spherical patch. It has been shown to be feasible, both theoretically and experimentally. Our approach finds the best reference patches from the example image for target image. Both image intensities and image gradients have theoretically been shown to be useful for finding the correspondence in the context of analogy-based relighting. By applying Poisson-derived methods, problems in relighting real images are alleviated. While this initial approach is capable of handling predominantly convex surfaces, the relation between surface shape and image measurements will be further investigated in Chapter 4.

Chapter 4

Surface Shape and Image Measurements

4.1 Curvature based Analysis

In the previous chapter, we assume the local region of the surface at each point can be approximated by a spherical patch. As we can see from the experiments (figure 3.7) on the more complex models, the framework breaks down quickly on non-spherical surface patches. A further study on the relation of local shape and relighting error is necessary.

4.1.1 Surface Curvature

Local shape can be described by surface curvatures, such as the principal curvatures, mean curvature and Gaussian curvature. The principal curvatures at each point p of a differentiable surface, denoted by k_1 and k_2 , are the maximum and minimum values of the curvature. The product k_1k_2 of the two principal curvatures is the Gaussian curvature, K , and the average $(k_1 + k_2)/2$ is the mean curvature, H .

An explicit parametric representation of a surface S in 3D Euclidean space E^3 :

$$S(u, v) = \begin{bmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{bmatrix}, \quad (4.1)$$

with respect to a known coordinate system.

The surface curvatures can be derived from two so-called fundamental forms. The first fundamental form is given by:

$$\begin{aligned} I(du, dv) &= dS \cdot dS \\ &= (S_u du + S_v dv) \cdot (S_u du + S_v dv) \\ &= Edu^2 + 2Fdudv + Gdv^2, \end{aligned} \quad (4.2)$$

where E , F , and G are the coefficients of the first fundamental form such that $E = S_u \cdot S_u$, $F = S_u \cdot S_v$, and $G = S_v \cdot S_v$.

The second fundamental form is given by:

$$\begin{aligned} II(du, dv) &= -dS \cdot dn \\ &= -(S_u du + S_v dv) \cdot (n_u du + n_v dv) \\ &= Ldu^2 + 2Mdudv + Ndv^2, \end{aligned} \quad (4.3)$$

where L , M , and N are the coefficients of the second fundamental form, and n is the surface normal at $S(u, v)$ such that $L = S_{uu} \cdot n$, $M = S_{uv} \cdot n$, $N = S_{vv} \cdot n$, and $n = \frac{S_u S_v}{|S_u S_v|}$.

Gaussian curvature and mean curvature can be calculated from the coefficients of the two fundamental forms as follows:

$$K = \frac{LN - M^2}{EG - F^2}, \quad (4.4)$$

$$H = \frac{LG - 2MF + NE}{2(EG - F^2)}. \quad (4.5)$$

In our framework, the surface models are represented as depth maps from a single view. The definition of the surface S can therefore be simplified to

$$S(x, y) = \begin{bmatrix} x \\ y \\ z(x, y) \end{bmatrix}, \quad (4.6)$$

where $z(x, y)$ is a depth or range value at some point (x, y) in a given range image.

Then the coefficients of the fundamental forms become

$$\begin{aligned} E &= S_x \cdot S_x = 1 + z_x^2, \\ F &= S_x \cdot S_y = z_x \cdot z_y, \\ G &= S_y \cdot S_y = 1 + z_y^2, \\ L &= S_{xx} \cdot n = \frac{z_{xx}}{\sqrt{1 + z_x^2 + z_y^2}}, \\ M &= S_{xy} \cdot n = \frac{z_{xy}}{\sqrt{1 + z_x^2 + z_y^2}}, \\ N &= S_{yy} \cdot n = \frac{z_{yy}}{\sqrt{1 + z_x^2 + z_y^2}}, \\ n &= \frac{S_x S_y}{|S_x S_y|} = \frac{1}{\sqrt{1 + z_x^2 + z_y^2}}(-z_x, -z_y, 1). \end{aligned} \quad (4.7)$$

The Gaussian curvature and mean curvature can be calculated as:

$$\begin{aligned} K &= \frac{LN - M^2}{EG - F^2} \\ &= \frac{z_{xx} \cdot z_{yy} - z_{xy}^2}{(1 + z_x^2 + z_y^2)^2}, \end{aligned} \quad (4.8)$$

$$\begin{aligned}
H &= \frac{LG - 2MF + NE}{2(EG - F^2)} \\
&= \frac{z_{xx} + z_{yy} + z_{xx} \cdot z_y^2 + z_{yy} \cdot z_x^2 - 2z_x \cdot z_y \cdot z_{xy}}{2(1 + z_x^2 + z_y^2)^{\frac{3}{2}}}.
\end{aligned} \tag{4.9}$$

As $K = k_1 k_2$ and $H = \frac{k_1 + k_2}{2}$, it follows that k_1 and k_2 are the two roots of the quadratic equation:

$$k^2 - 2Hk + K = 0. \tag{4.10}$$

If K and H are known at each point of the surface, the two principal curvatures can be determined by:

$$k_{1,2} = H \pm \sqrt{H^2 - K}. \tag{4.11}$$

If $H^2 = K$ at a surface point, the point is referred to as an umbilical point to denote that the principal curvatures are equal, *i.e.* $k_1 = k_2 = H$. This is identical to the spherical surface assumption we made previously. Pentland [62] also used this assumption, which is strong enough to yield a unique shape interpretation from local surface shading.

4.1.2 Classification of Surface Patch using Curvature

Gaussian curvature is an intrinsic property of the surface, maintaining its sign even when the direction of the normal vector is flipped. A point on a surface can be locally classified into one of the following three types according to the sign of the Gaussian curvature:

- an elliptic surface point, when $K > 0$;
- a hyperbolic surface point, when $K < 0$;
- a parabolic surface point, when $K = 0$.

On the other hand, the sign of mean curvature depends on the direction of the normal. If surface normals were defined to be pointing away from the surface, then surface points can be classified according to their mean and Gaussian curvatures as in table 4.1.

Similarly, surface points can be classified according to their two principal curvatures (with $k_1 \geq k_2$) as in table 4.2. Examples of the principal curvatures calculated during the experiments are shown in figure 4.1, where their values are indicated by the color bars.

Table 4.1: Surface classification based on Gaussian and mean curvatures

$H \backslash K$	+	0	-
-	convex ellipsoid	convex parabolic	hyperbolic saddle-like
0	impossible	planar - minimal surface	hyperbolic - minimal surface
+	concave ellipsoid	concave parabolic	hyperbolic saddle-like

Table 4.2: Surface classification based on principal curvatures

$k_2 \backslash k_1$	-	0	+
-	convex ellipsoid	convex parabolic	hyperbolic saddle-like
0	impossible	plane	concave parabolic
+	impossible	impossible	concave ellipsoid

4.1.3 Relation of Curvature and Synthesis Errors

As principal curvature signs are normal direction dependent, Koenderink [39] introduced a shape index S and a curvedness C function. These depend on the two principal curvatures as follows:

$$S(k_1, k_2) = -\frac{2}{\pi} \tan^{-1} \left(\frac{k_1 + k_2}{k_1 - k_2} \right) \in [-1, 1], \quad (4.12)$$

(a) Principal curvature of male model

(b) Principal curvature of female model

Figure 4.1: The principal curvatures of the 3d models. The values are indicated by the color bars.

$$C(k_1, k_2) = \sqrt{\frac{k_1^2 + k_2^2}{2}} \in [0, +\infty]. \quad (4.13)$$

Intuitively, the shape index (SI) quantifies the shape type of surface patches as a continuous variation from concave to saddle to convex shapes, while curvedness

captures the profile amplitudes of the surface patches.

As shown in figure 4.2, the shape index S defined by Koenderink increases from -1 to 1 by turning counterclockwise from bottom left to top right, while the curvedness C varies between 0 to $+\infty$ from the origin toward the periphery. Since we show the proof that the method in Chapter 3 is invariant to the radius of spherical patch, it implies the curvedness is not related to relighting errors. The comparison in figure 3.7 also shows that the results is only slight improved with multi-scale example images, indicating that the approach has low sensitivity to the curvedness which describes the scale sensitivity.

Figure 4.2: Koenderink's shape index and curvedness, figure from [58]

Experiments show that, statistically, there is an approximate linear relation between a shape index and the synthesis errors as shown in figure 4.3. These models are publicly available from cyberware, whose website lists them as samples <http://cyberware.com/products/scanners/psSamples.html>. We have chose two complex models to be presented in the thesis, as they can already provide all kinds of surface shapes.

For convenience, the shape index we use in our experiment is defined as:

$$SI(k_1, k_2) = \frac{1}{\pi} \text{atan2}(k_2, k_1) + \frac{3}{4} \in [0, 1], \quad (4.14)$$

where $k_1 \geq k_2$, so the convex spherical patches have $SI = 0$, and they are expected to result in minimal errors.

In figure 4.3, each diagram labeled “avg error” plots the average errors for surface patches with different k_1 and k_2 . The image labeled “SI” shows the shape index, while the “error” image indicates the synthesis error. The approximate linear relation is also plotted, with the regression coefficients (slope and intercept) shown on the top. In this plot, the horizontal axis is the shape index, and the vertical axis is the mean error of the data points in each SI interval.

4.2 Local Shape Analysis with Image Measurements

As shown in the previous chapter, we can match spherical surface with equal principal curvature by using zeroth and first order derivatives of image intensity. In this section, we will extend beyond spherical surfaces to locally quadratic ones with the help of second order derivatives of image intensity.

4.2.1 Local Surface Shape

For every point on a surface we can define a coordinate system in which the point of interest P is the origin, the x and y axes lie in the tangent plane of the surface, and the z axis is aligned with the surface normal. Assuming the surface around P is smooth,

(a) Error analysis of male model with convex sphere example

Figure 4.3

(b) Error analysis of female model with convex sphere example

Figure 4.3

(c) Error analysis of male model with concave sphere example

Figure 4.3

(d) Error analysis of female model with concave sphere example

Figure 4.3: Approximate linear relation relation of curvature and synthetic error. In “avg error”, k_1 is the vertical axis, $[-0.5, 0.5]$ from top to bottom, and k_2 is the horizontal axis, $[-0.5, 0.5]$ from left to right.

the local patch can be modeled by a smooth function:

$$z = f(x, y). \quad (4.15)$$

Expanding the function using Taylor series at P , we get

$$\begin{aligned} z = & f(0, 0) + \frac{\partial f}{\partial x}(0, 0)x + \frac{\partial f}{\partial y}(0, 0)y \\ & + \frac{1}{2} \left[\frac{\partial^2 f}{\partial x^2}(0, 0)x^2 + 2\frac{\partial^2 f}{\partial x \partial y}(0, 0)xy + \frac{\partial^2 f}{\partial y^2}(0, 0)y^2 \right] \\ & + O(x, y)^3. \end{aligned} \quad (4.16)$$

The tangent plane is expressed as

$$z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0), \quad (4.17)$$

and since P is the origin, then

$$f(0, 0) + \frac{\partial f}{\partial x}(0, 0)x + \frac{\partial f}{\partial y}(0, 0)y = 0. \quad (4.18)$$

So the local surface function becomes:

$$z = \frac{1}{2} \left[\frac{\partial^2 f}{\partial x^2}(0, 0)x^2 + 2\frac{\partial^2 f}{\partial x \partial y}(0, 0)xy + \frac{\partial^2 f}{\partial y^2}(0, 0)y^2 \right] + O(x, y)^3. \quad (4.19)$$

Without loss of generality, we can rotate the xy -plane about the z axis to eliminate the xy term. The surface in the new local coordinate system $(\bar{x}, \bar{y}, \bar{z})$ is:

$$\bar{z} = \frac{1}{2} \frac{\partial^2 f}{\partial \bar{x}^2}(0, 0)\bar{x}^2 + \frac{1}{2} \frac{\partial^2 f}{\partial \bar{y}^2}(0, 0)\bar{y}^2 + O(\bar{x}, \bar{y})^3. \quad (4.20)$$

In this object coordinate system, $\bar{x}$ and $\bar{y}$ are the principal axes and $\frac{\partial^2 f}{\partial \bar{x}^2}(0,0)$ and $\frac{\partial^2 f}{\partial \bar{y}^2}(0,0)$ are the principal curvatures, *i.e.* the maximum and minimum values of the curvature, k_1 and k_2 .

If the principal curvatures are equal *i.e.* $k_1 = k_2$, the surface point is referred to as an umbilical point, meaning locally the surface is spherical. Pentland [62] proved that for an umbilical point on a Lambertian surface, we can determine the surface parameters by looking at six measurements of the image of the point, which are the intensity and its first and second derivatives $I, \frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}, \frac{\partial^2 I}{\partial x^2}, \frac{\partial^2 I}{\partial y^2}, \frac{\partial^2 I}{\partial x \partial y}$. Two solutions (surface orientation and radius of curvature) can be found, one corresponding to a convex point, the other to a concave point. If the lighting direction is known, then this ambiguity can be solved, the solution becomes unique. Here we investigate more complex surfaces beyond spherical ones. Oliensis [60] claimed an unique solution for the surface parameters is determined using up to second order image derivatives. However, experimentally we find four types of surfaces that may lead to the same six image measurements. Next, we will show the mathematical setup and our numerical experiments.

By omitting higher order terms, the surface patch can be represented as:

$$\bar{z} = \frac{1}{2}(k_1 \bar{x}^2 + k_2 \bar{y}^2), \quad (4.21)$$

in the local coordinate system, which is called the object coordinate system.

Note axis $\bar{z}$ is in the same direction of surface normal $\bar{N}$, while axes $\bar{x}$ and $\bar{y}$ are in the two principal curvature directions, and k_1 and k_2 are the principal curvatures.

So the surface derivatives with respect to $\bar{x}$ and $\bar{y}$ are:

$$\frac{\partial \bar{z}}{\partial \bar{x}} = k_1 \bar{x}, \quad (4.22)$$

$$\frac{\partial \bar{z}}{\partial \bar{y}} = k_2 \bar{y}. \quad (4.23)$$

The surface normal is:

$$\bar{N} = \frac{1}{\sqrt{\left(\frac{\partial \bar{z}}{\partial \bar{x}}\right)^2 + \left(\frac{\partial \bar{z}}{\partial \bar{y}}\right)^2 + 1}} \begin{bmatrix} -\frac{\partial \bar{z}}{\partial \bar{x}} \\ -\frac{\partial \bar{z}}{\partial \bar{y}} \\ 1 \end{bmatrix} = \frac{1}{\gamma} \begin{bmatrix} -k_1 \bar{x} \\ -k_2 \bar{y} \\ 1 \end{bmatrix}, \quad (4.24)$$

where $\gamma = \sqrt{(k_1 \bar{x})^2 + (k_2 \bar{y})^2 + 1}$.

The image intensity equation is:

$$I = \bar{N} \cdot \bar{L} = \frac{1}{\gamma} \begin{bmatrix} -k_1 \bar{x} \\ -k_2 \bar{y} \\ 1 \end{bmatrix} \cdot \bar{L}, \quad (4.25)$$

where $\bar{N}$ and $\bar{L}$ are in the object coordinate system.

By differentiating N with respect to $\bar{x}$, we get the first order derivative:

$$\begin{aligned} \frac{\partial \bar{N}}{\partial \bar{x}} &= \begin{bmatrix} -\frac{k_1}{\gamma} + \frac{k_1^3 \bar{x}^2}{\gamma^3} \\ \frac{k_1^2 k_2 \bar{x} \bar{y}}{\gamma^3} \\ -\frac{k_1^2 \bar{x}}{\gamma^3} \end{bmatrix} \\ &= \frac{1}{\gamma} \begin{bmatrix} -k_1 \\ 0 \\ 0 \end{bmatrix} + \frac{1}{\gamma^3} \begin{bmatrix} k_1^3 \bar{x}^2 \\ k_1^2 k_2 \bar{x} \bar{y} \\ -k_1^2 \bar{x} \end{bmatrix} \\ &= \frac{1}{\gamma} \begin{bmatrix} -k_1 \\ 0 \\ 0 \end{bmatrix} + \frac{k_1^2 \bar{x}}{\gamma^3} \begin{bmatrix} k_1 \bar{x} \\ k_2 \bar{y} \\ -1 \end{bmatrix}. \end{aligned} \quad (4.26)$$

Similarly the first order derivative with respect to $\bar{y}$ is:

$$\frac{\partial \bar{N}}{\partial \bar{y}} = \frac{1}{\gamma} \begin{bmatrix} 0 \\ -k_2 \\ 0 \end{bmatrix} + \frac{k_2^2 \bar{y}}{\gamma^3} \begin{bmatrix} k_1 \bar{x} \\ k_2 \bar{y} \\ -1 \end{bmatrix}. \quad (4.27)$$

At the point of interest, the first order derivatives reduce to:

$$\frac{\partial N}{\partial \bar{x}} = \begin{bmatrix} -k_1 \\ 0 \\ 0 \end{bmatrix}, \quad (4.28)$$

$$\frac{\partial N}{\partial \bar{y}} = \begin{bmatrix} 0 \\ -k_2 \\ 0 \end{bmatrix}. \quad (4.29)$$

Differentiating N further yields the second order derivatives:

$$\begin{aligned} \frac{\partial^2 N}{\partial \bar{x}^2} &= \partial \left(\frac{1}{\gamma} \begin{bmatrix} -k_1 \\ 0 \\ 0 \end{bmatrix} + \frac{k_1^2 \bar{x}}{\gamma^3} \begin{bmatrix} k_1 \bar{x} \\ k_2 \bar{y} \\ -1 \end{bmatrix} \right) / \partial \bar{x} \\ &= \frac{k_1^2 \bar{x}}{\gamma^3} \begin{bmatrix} k_1 \\ 0 \\ 0 \end{bmatrix} + \frac{k_1^2}{\gamma^5} \begin{bmatrix} k_1 \bar{x} (-k_1^2 \bar{x}^2 + 2k_2^2 \bar{y}^2 + 2) \\ k_2 \bar{y} (-2k_1^2 \bar{x}^2 + k_2^2 \bar{y}^2 + 1) \\ 2k_1^2 \bar{x}^2 - k_2^2 \bar{y}^2 - 1 \end{bmatrix}, \end{aligned} \quad (4.30)$$

$$\begin{aligned}
\frac{\partial^2 N}{\partial \bar{x} \partial \bar{y}} &= \partial \left(\frac{1}{\gamma} \begin{bmatrix} -k_1 \\ 0 \\ 0 \end{bmatrix} + \frac{k_1^2 \bar{x}}{\gamma^3} \begin{bmatrix} k_1 \bar{x} \\ k_2 \bar{y} \\ -1 \end{bmatrix} \right) / \partial \bar{y} \\
&= \frac{k_2^2 \bar{y}}{\gamma^3} \begin{bmatrix} k_1 \\ 0 \\ 0 \end{bmatrix} + \frac{k_1^2 \bar{x}}{\gamma^5} \begin{bmatrix} 3k_2^2 \bar{y} \\ k_1^2 k_2 \bar{x}^2 - 2k_2^3 \bar{y}^2 + k_2 \\ -3k_1 k_2^2 \bar{x} \bar{y} \end{bmatrix},
\end{aligned} \tag{4.31}$$

$$\begin{aligned}
\frac{\partial^2 N}{\partial \bar{y}^2} &= \partial \left(\frac{1}{\gamma} \begin{bmatrix} 0 \\ -k_2 \\ 0 \end{bmatrix} + \frac{k_2^2 \bar{y}}{\gamma^3} \begin{bmatrix} k_1 \bar{x} \\ k_2 \bar{y} \\ -1 \end{bmatrix} \right) / \partial \bar{y} \\
&= \frac{k_2^2 \bar{y}}{\gamma^3} \begin{bmatrix} 0 \\ k_2 \\ 0 \end{bmatrix} + \frac{k_2^2}{\gamma^5} \begin{bmatrix} k_1 \bar{x} (k_1^2 \bar{x}^2 - 2k_2^2 \bar{y}^2 + 1) \\ k_2 \bar{y} (2k_1^2 \bar{x}^2 - k_2^2 \bar{y}^2 + 2) \\ -k_1^2 \bar{x}^2 + 2k_2^2 \bar{y}^2 - 1 \end{bmatrix}.
\end{aligned} \tag{4.32}$$

At the point of interest, the second order derivatives reduce to:

$$\frac{\partial^2 N}{\partial \bar{x}^2} = \begin{bmatrix} 0 \\ 0 \\ -k_1^2 \end{bmatrix}, \tag{4.33}$$

$$\frac{\partial^2 N}{\partial \bar{x} \partial \bar{y}} = 0, \tag{4.34}$$

$$\frac{\partial^2 N}{\partial \bar{y}^2} = \begin{bmatrix} 0 \\ 0 \\ -k_2^2 \end{bmatrix}. \tag{4.35}$$

We have derived the local shape derivatives, in the following section, we will show their relation to image measurements.

4.2.2 Image Measurements

The previous derivations are all conducted in the object coordinate system. Assume the object coordinate system can be transformed into camera coordinate system by a rotation R , *i.e.*

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathfrak{R} \begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{bmatrix}, \quad (4.36)$$

where $\mathfrak{R}$ has three degrees of freedom:

$$\begin{aligned} \mathfrak{R} &\equiv \begin{bmatrix} \mathfrak{R}_x \\ \mathfrak{R}_y \\ \mathfrak{R}_z \end{bmatrix} \equiv \begin{bmatrix} \mathfrak{R}_{\bar{x}} & \mathfrak{R}_{\bar{y}} & \mathfrak{R}_{\bar{z}} \end{bmatrix} \\ &\equiv \begin{bmatrix} \mathfrak{R}_{x\bar{x}} & \mathfrak{R}_{x\bar{y}} & \mathfrak{R}_{x\bar{z}} \\ \mathfrak{R}_{y\bar{x}} & \mathfrak{R}_{y\bar{y}} & \mathfrak{R}_{y\bar{z}} \\ \mathfrak{R}_{z\bar{x}} & \mathfrak{R}_{z\bar{y}} & \mathfrak{R}_{z\bar{z}} \end{bmatrix}. \end{aligned} \quad (4.37)$$

Note, in general, the transformation also involves a translation. However, as we focus on the case with a directional lighting, the image intensity and its derivatives are only determined by the orientation of the surface with respect to the lighting direction. Ignoring the translation will not affect all the subsequent derivation and analysis.

As

$$\begin{aligned}
 \frac{\partial \bar{x}}{\partial x} &= \begin{bmatrix} \mathfrak{R}_{x\bar{x}} & \mathfrak{R}_{y\bar{x}} & \mathfrak{R}_{z\bar{x}} \end{bmatrix} \begin{bmatrix} \frac{\partial x}{\partial x} \\ \frac{\partial y}{\partial x} \\ \frac{\partial z}{\partial x} \end{bmatrix} \\
 &= \begin{bmatrix} \mathfrak{R}_{x\bar{x}} & \mathfrak{R}_{y\bar{x}} & \mathfrak{R}_{z\bar{x}} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\
 &= \mathfrak{R}_{x\bar{x}},
 \end{aligned} \tag{4.38}$$

$$\begin{aligned}
 \frac{\partial \bar{y}}{\partial y} &= \begin{bmatrix} \mathfrak{R}_{x\bar{y}} & \mathfrak{R}_{y\bar{y}} & \mathfrak{R}_{z\bar{y}} \end{bmatrix} \begin{bmatrix} \frac{\partial x}{\partial y} \\ \frac{\partial y}{\partial y} \\ \frac{\partial z}{\partial y} \end{bmatrix} \\
 &= \begin{bmatrix} \mathfrak{R}_{x\bar{y}} & \mathfrak{R}_{y\bar{y}} & \mathfrak{R}_{z\bar{y}} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \\
 &= \mathfrak{R}_{y\bar{y}}.
 \end{aligned} \tag{4.39}$$

Similarly we get a unified notation at the point of interest:

$$\frac{\partial \bar{r}}{\partial r} = \mathfrak{R}_{r\bar{r}}. \tag{4.40}$$

It is obvious that:

$$\frac{\partial^2 \bar{r}_i}{\partial r_i \partial r_j} = 0, \tag{4.41}$$

where i, j denote the x or y .

Under Lambertian model, we can establish following image equations in camera coordinate system.

Zeroth order derivative, image intensity:

$$\begin{aligned}
 I &= N \cdot L \\
 &= \Re \bar{N} \cdot L \\
 &= \Re \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \cdot L \\
 &= \Re_{\bar{z}} \cdot L \\
 &= L^T \begin{bmatrix} \Re_{x\bar{z}} \\ \Re_{y\bar{z}} \\ \Re_{z\bar{z}} \end{bmatrix}.
 \end{aligned} \tag{4.42}$$

First order derivatives, image gradients:

$$\begin{aligned}
 \frac{\partial I}{\partial x} &= \frac{\partial \bar{N}}{\partial x} \cdot \bar{L} \\
 &= \bar{L}^T \left(\frac{\partial \bar{N}}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial x} + \frac{\partial \bar{N}}{\partial \bar{y}} \frac{\partial \bar{y}}{\partial x} \right) \\
 &= L^T \Re \left(\begin{bmatrix} -k_1 \\ 0 \\ 0 \end{bmatrix} \Re_{x\bar{x}} + \begin{bmatrix} 0 \\ -k_2 \\ 0 \end{bmatrix} \Re_{x\bar{y}} \right) \\
 &= L^T \Re \begin{bmatrix} -k_1 \Re_{x\bar{x}} \\ -k_2 \Re_{x\bar{y}} \\ 0 \end{bmatrix},
 \end{aligned} \tag{4.43}$$

$$\begin{aligned}
\frac{\partial I}{\partial y} &= \frac{\partial \bar{N}}{\partial y} \cdot \bar{L} \\
&= \bar{L}^T \left(\frac{\partial \bar{N}}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial y} + \frac{\partial \bar{N}}{\partial \bar{y}} \frac{\partial \bar{y}}{\partial y} \right) \\
&= L^T \Re \left(\begin{bmatrix} -k_1 \\ 0 \\ 0 \end{bmatrix} \Re_{y\bar{x}} + \begin{bmatrix} 0 \\ -k_2 \\ 0 \end{bmatrix} \Re_{y\bar{y}} \right) \\
&= L^T \Re \begin{bmatrix} -k_1 \Re_{y\bar{x}} \\ -k_2 \Re_{y\bar{y}} \\ 0 \end{bmatrix}.
\end{aligned} \tag{4.44}$$

Second order derivatives:

$$\begin{aligned}
\frac{\partial^2 I}{\partial x^2} &= \frac{\partial^2 \bar{N}}{\partial x^2} \cdot \bar{L} \\
&= \bar{L}^T \frac{\partial \left(\frac{\partial \bar{N}}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial x} + \frac{\partial \bar{N}}{\partial \bar{y}} \frac{\partial \bar{y}}{\partial x} \right)}{\partial x} \\
&= \bar{L}^T \left(\frac{\partial^2 \bar{N}}{\partial \bar{x} \partial x} \frac{\partial \bar{x}}{\partial x} + \frac{\partial \bar{N}}{\partial \bar{x}} \frac{\partial^2 \bar{x}}{\partial x^2} + \frac{\partial^2 \bar{N}}{\partial \bar{y} \partial x} \frac{\partial \bar{y}}{\partial x} + \frac{\partial \bar{N}}{\partial \bar{y}} \frac{\partial^2 \bar{y}}{\partial x^2} \right) \\
&= \bar{L}^T \left[\left(\frac{\partial^2 \bar{N}}{\partial \bar{x}^2} \frac{\partial \bar{x}}{\partial x} + \frac{\partial^2 \bar{N}}{\partial \bar{x} \partial \bar{y}} \frac{\partial \bar{y}}{\partial x} \right) \frac{\partial \bar{x}}{\partial x} + \frac{\partial \bar{N}}{\partial \bar{x}} \frac{\partial^2 \bar{x}}{\partial x^2} + \left(\frac{\partial^2 \bar{N}}{\partial \bar{x} \partial \bar{y}} \frac{\partial \bar{x}}{\partial x} + \frac{\partial^2 \bar{N}}{\partial \bar{y}^2} \frac{\partial \bar{y}}{\partial x} \right) \frac{\partial \bar{y}}{\partial x} + \frac{\partial \bar{N}}{\partial \bar{y}} \frac{\partial^2 \bar{y}}{\partial x^2} \right] \\
&= L^T \Re \left(\begin{bmatrix} 0 \\ 0 \\ -k_1^2 \end{bmatrix} \Re_{x\bar{x}} \Re_{x\bar{x}} + \begin{bmatrix} 0 \\ 0 \\ -k_2^2 \end{bmatrix} \Re_{x\bar{y}} \Re_{x\bar{y}} \right) \\
&= L^T \Re \begin{bmatrix} 0 \\ 0 \\ -k_1^2 \Re_{x\bar{x}}^2 - k_2^2 \Re_{x\bar{y}}^2 \end{bmatrix},
\end{aligned} \tag{4.45}$$

$$\begin{aligned}
\frac{\partial^2 I}{\partial y^2} &= \frac{\partial^2 \bar{N}}{\partial y^2} \cdot \bar{L} = \bar{L}^T \frac{\partial \left(\frac{\partial \bar{N}}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial y} + \frac{\partial \bar{N}}{\partial \bar{y}} \frac{\partial \bar{y}}{\partial y} \right)}{\partial y} \\
&= \bar{L}^T \left(\frac{\partial^2 \bar{N}}{\partial \bar{x} \partial y} \frac{\partial \bar{x}}{\partial y} + \frac{\partial \bar{N}}{\partial \bar{x}} \frac{\partial^2 \bar{x}}{\partial y^2} + \frac{\partial^2 \bar{N}}{\partial \bar{y} \partial y} \frac{\partial \bar{y}}{\partial y} + \frac{\partial \bar{N}}{\partial \bar{y}} \frac{\partial^2 \bar{y}}{\partial y^2} \right) \\
&= \bar{L}^T \left[\left(\frac{\partial^2 \bar{N}}{\partial \bar{x}^2} \frac{\partial \bar{x}}{\partial y} + \frac{\partial^2 \bar{N}}{\partial \bar{x} \partial \bar{y}} \frac{\partial \bar{y}}{\partial y} \right) \frac{\partial \bar{x}}{\partial y} + \frac{\partial \bar{N}}{\partial \bar{x}} \frac{\partial^2 \bar{x}}{\partial y^2} \right. \\
&\quad \left. + \left(\frac{\partial^2 \bar{N}}{\partial \bar{x} \partial \bar{y}} \frac{\partial \bar{x}}{\partial y} + \frac{\partial^2 \bar{N}}{\partial \bar{y}^2} \frac{\partial \bar{y}}{\partial y} \right) \frac{\partial \bar{y}}{\partial y} + \frac{\partial \bar{N}}{\partial \bar{y}} \frac{\partial^2 \bar{y}}{\partial y^2} \right] \\
&= L^T \Re \left(\begin{bmatrix} 0 \\ 0 \\ -k_1^2 \end{bmatrix} \Re_{y\bar{x}} \Re_{y\bar{x}} + \begin{bmatrix} 0 \\ 0 \\ -k_2^2 \end{bmatrix} \Re_{y\bar{y}} \Re_{y\bar{y}} \right) \\
&= L^T \Re \begin{bmatrix} 0 \\ 0 \\ -k_1^2 \Re_{y\bar{x}}^2 - k_2^2 \Re_{y\bar{y}}^2 \end{bmatrix}, \tag{4.46}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 I}{\partial x \partial y} &= \frac{\partial^2 \bar{N}}{\partial x \partial y} \cdot \bar{L} = \bar{L}^T \frac{\partial \left(\frac{\partial \bar{N}}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial x} + \frac{\partial \bar{N}}{\partial \bar{y}} \frac{\partial \bar{y}}{\partial x} \right)}{\partial y} \\
&= \bar{L}^T \left(\frac{\partial^2 \bar{N}}{\partial \bar{x} \partial y} \frac{\partial \bar{x}}{\partial x} + \frac{\partial \bar{N}}{\partial \bar{x}} \frac{\partial^2 \bar{x}}{\partial x \partial y} + \frac{\partial^2 \bar{N}}{\partial \bar{y} \partial y} \frac{\partial \bar{y}}{\partial x} + \frac{\partial \bar{N}}{\partial \bar{y}} \frac{\partial^2 \bar{y}}{\partial x \partial y} \right) \\
&= \bar{L}^T \left[\left(\frac{\partial^2 \bar{N}}{\partial \bar{x}^2} \frac{\partial \bar{x}}{\partial y} + \frac{\partial^2 \bar{N}}{\partial \bar{x} \partial \bar{y}} \frac{\partial \bar{y}}{\partial y} \right) \frac{\partial \bar{x}}{\partial x} + \frac{\partial \bar{N}}{\partial \bar{x}} \frac{\partial^2 \bar{x}}{\partial x \partial y} \right. \\
&\quad \left. + \left(\frac{\partial^2 \bar{N}}{\partial \bar{x} \partial \bar{y}} \frac{\partial \bar{x}}{\partial y} + \frac{\partial^2 \bar{N}}{\partial \bar{y}^2} \frac{\partial \bar{y}}{\partial y} \right) \frac{\partial \bar{y}}{\partial x} + \frac{\partial \bar{N}}{\partial \bar{y}} \frac{\partial^2 \bar{y}}{\partial x \partial y} \right] \\
&= L^T \Re \left(\begin{bmatrix} 0 \\ 0 \\ -k_1^2 \end{bmatrix} \Re_{y\bar{x}} \Re_{x\bar{x}} + \begin{bmatrix} 0 \\ 0 \\ -k_2^2 \end{bmatrix} \Re_{y\bar{y}} \Re_{x\bar{y}} \right) \\
&= L^T \Re \begin{bmatrix} 0 \\ 0 \\ -k_1^2 \Re_{y\bar{x}} \Re_{x\bar{x}} - k_2^2 \Re_{y\bar{y}} \Re_{x\bar{y}} \end{bmatrix}. \tag{4.47}
\end{aligned}$$

To summarize:

$$I = L^T \begin{bmatrix} \Re_{x\bar{z}} \\ \Re_{y\bar{z}} \\ \Re_{z\bar{z}} \end{bmatrix}, \quad (4.48)$$

$$\frac{\partial I}{\partial x} = L^T \Re \begin{bmatrix} -k_1 \Re_{x\bar{x}} \\ -k_2 \Re_{x\bar{y}} \\ 0 \end{bmatrix}, \quad (4.49)$$

$$\frac{\partial I}{\partial y} = L^T \Re \begin{bmatrix} -k_1 \Re_{y\bar{x}} \\ -k_2 \Re_{y\bar{y}} \\ 0 \end{bmatrix}, \quad (4.50)$$

$$\frac{\partial^2 I}{\partial x^2} = L^T \Re \begin{bmatrix} 0 \\ 0 \\ -k_1^2 \Re_{x\bar{x}}^2 - k_2^2 \Re_{x\bar{y}}^2 \end{bmatrix}, \quad (4.51)$$

$$\frac{\partial^2 I}{\partial y^2} = L^T \Re \begin{bmatrix} 0 \\ 0 \\ -k_1^2 \Re_{y\bar{x}}^2 - k_2^2 \Re_{y\bar{y}}^2 \end{bmatrix}, \quad (4.52)$$

$$\frac{\partial^2 I}{\partial x \partial y} = L^T \Re \begin{bmatrix} 0 \\ 0 \\ -k_1^2 \Re_{y\bar{x}} \Re_{x\bar{x}} - k_2^2 \Re_{y\bar{y}} \Re_{x\bar{y}} \end{bmatrix}. \quad (4.53)$$

There are six equations available. The unknowns are the rotation matrix R which transforms the object coordinate system into camera coordinate system (three independent degrees of freedom), two principal curvatures k_1 and k_2 which characterize the

local shape, and the product of light intensity and surface albedo which was omitted in previous derivations. With total six degrees of freedom, the nonlinear system of above six equations is solvable (Here we assume lighting direction L is known, otherwise the system has two more unknowns for the two degrees of freedom of L , the system becomes unsolvable), but the uniqueness of the solution is not determined.

4.3 Experiments

As it is not easy to solve the nonlinear system of image derivatives analytically, we verify the uniqueness of the solution with numerical solver. Under a given lighting condition L , we first chose a set of parameters R , k_1 and k_2 (the product of light intensity and surface albedo is omitted as it is only a linear factor) to calculate the image derivatives using (4.48) (4.49) (4.50) (4.51) (4.52) (4.53). Then we tried to inversely recover the parameters with these resulting derivatives. The numerical solver empirically finds there are 16 solutions (combinations of R , k_1 and k_2), given the same set of six derivatives. The 16 solutions actually only relate to four different surface shapes, each surface shape corresponds to four solutions. Once the orientation (rotation matrix R) and principal curvatures (k_1 and k_2) are fixed, the shape is fixed. However, k_1 and k_2 can be switched with each other, and if the surface rotates about normal by π , R becomes different, these lead to the other three solutions.

We have tried different types of numerical algorithms, including trust-region-dogleg, trust-region-reflective and levenberg-marquardt using different parameters with very small tolerance. Besides the original six equations, additional constraints (or the object functions to be minimised) are listed in table 4.6. Some numerical results are shown in the tables. There are 16 combinations of solved parameters (table 4.3), resulting derivatives (table 4.4) and errors (table 4.5 and table 4.6). The actual four types of

the surface are shown in figure 4.4. In each sub-figure, the corresponding indices of the cases in the tables are listed. We can see even with accurate derivatives up to second order, the solution is not unique.

Table 4.3: Numerical results: solved parameters. There are 16 unique solutions found, one solution each row.

	$\Re_{xx}$	$\Re_{xy}$	$\Re_{xz}$	$\Re_{yx}$	$\Re_{yy}$	$\Re_{yz}$	$\Re_{zx}$	$\Re_{zy}$	$\Re_{zz}$	k_1	k_2
1	0.8979	-0.4128	0.1529	0.2575	0.7743	0.5781	-0.3570	-0.4797	0.8015	1.1986	-0.5409
2	0.9028	-0.3714	0.2169	0.2449	0.8585	0.4505	-0.3536	-0.3536	0.8660	1.2000	0.5000
3	0.3357	-0.6334	0.6972	0.1107	0.7616	0.6386	-0.9354	-0.1372	0.3258	-3.1295	0.5096
4	-0.6785	-0.3313	0.6557	0.7187	-0.1142	0.6859	-0.1524	0.9366	0.3156	-0.5245	-3.1387
5	0.6785	0.3313	0.6557	-0.7187	0.1142	0.6859	0.1524	-0.9366	0.3156	-0.5245	-3.1387
6	-0.9028	0.3714	0.2169	-0.2449	-0.8585	0.4505	0.3536	0.3536	0.8660	1.2000	0.5000
7	-0.8979	0.4128	0.1529	-0.2575	-0.7743	0.5781	0.3570	0.4797	0.8015	1.1986	-0.5409
8	0.3714	0.9028	0.2169	-0.8585	0.2449	0.4505	0.3536	-0.3536	0.8660	0.5000	1.2000
9	-0.3357	0.6334	0.6972	-0.1107	-0.7616	0.6386	0.9354	0.1372	0.3258	-3.1295	0.5096
10	-0.3714	-0.9028	0.2169	0.8585	-0.2449	0.4505	-0.3536	0.3536	0.8660	0.5000	1.2000
11	-0.4128	-0.8979	0.1529	0.7743	-0.2575	0.5781	-0.4797	0.3570	0.8015	-0.5409	1.1986
12	-0.3313	0.6785	0.6557	-0.1142	-0.7187	0.6859	0.9366	0.1524	0.3156	-3.1387	-0.5245
13	0.3313	-0.6785	0.6557	0.1142	0.7187	0.6859	-0.9366	-0.1524	0.3156	-3.1387	-0.5245
14	0.6334	0.3357	0.6972	-0.7616	0.1107	0.6386	0.1372	-0.9354	0.3258	0.5096	-3.1295
15	0.4128	0.8979	0.1529	-0.7743	0.2575	0.5781	0.4797	-0.3570	0.8015	-0.5409	1.1986
16	-0.6334	-0.3357	0.6972	0.7616	-0.1107	0.6386	-0.1372	0.9354	0.3258	0.5096	-3.1295

Table 4.4: Numerical results: resulting derivatives. These solutions all result in the same combination of image measurements.

	I	$\frac{\partial I}{\partial x}$	$\frac{\partial I}{\partial y}$	$\frac{\partial^2 I}{\partial x^2}$	$\frac{\partial^2 I}{\partial y^2}$	$\frac{\partial^2 I}{\partial x \partial y}$
1	0.9276	-0.3798	-0.1423	-1.1206	-0.2511	-0.2214
2	0.9276	-0.3798	-0.1423	-1.1206	-0.2511	-0.2214
3	0.9276	-0.3798	-0.1423	-1.1206	-0.2511	-0.2214
4	0.9276	-0.3798	-0.1423	-1.1206	-0.2511	-0.2214
5	0.9276	-0.3798	-0.1423	-1.1206	-0.2511	-0.2214
6	0.9276	-0.3798	-0.1423	-1.1206	-0.2511	-0.2214
7	0.9276	-0.3798	-0.1423	-1.1206	-0.2511	-0.2214
8	0.9276	-0.3798	-0.1423	-1.1206	-0.2511	-0.2214
9	0.9276	-0.3798	-0.1423	-1.1206	-0.2511	-0.2214
10	0.9276	-0.3798	-0.1423	-1.1206	-0.2511	-0.2214
11	0.9276	-0.3798	-0.1423	-1.1206	-0.2511	-0.2214
12	0.9276	-0.3798	-0.1423	-1.1206	-0.2511	-0.2214
13	0.9276	-0.3798	-0.1423	-1.1206	-0.2511	-0.2214
14	0.9276	-0.3798	-0.1423	-1.1206	-0.2511	-0.2214
15	0.9276	-0.3798	-0.1423	-1.1206	-0.2511	-0.2214
16	0.9276	-0.3798	-0.1423	-1.1206	-0.2511	-0.2214

Table 4.5: Numerical results: errors of resulting derivatives (10^{-015} *).

	$I - GT$	$\frac{\partial I}{\partial x} - GT$	$\frac{\partial I}{\partial y} - GT$	$\frac{\partial^2 I}{\partial x^2} - GT$	$\frac{\partial^2 I}{\partial y^2} - GT$	$\frac{\partial^2 I}{\partial x \partial y} - GT$
1	0	0	-0.0278	0.2220	0.0555	0.0555
2	-0.1110	0.1665	-0.3886	-0.2220	0.1665	0
3	0	-0.0555	0	0.2220	0	0.0555
4	0	0.0555	0.0833	-0.2220	0	-0.0555
5	0.1110	0.2220	0.2498	0	-0.0555	0.0555
6	-0.1110	0.1665	0.0278	0	0.0555	0
7	0.1110	0	0	0.2220	-0.0555	0.0555
8	0	-0.0555	0.0278	0	0.0555	-0.0555
9	0	0.1110	0.0278	0	0.0555	0.1388
10	0	-0.0555	0.0278	0	0.0555	-0.0555
11	0	0	0	0.2220	-0.0555	0
12	0	0.0555	0.0555	0	0.0555	0.0278
13	0.2220	0.1665	0.4718	0	-0.4441	-0.0555
14	-0.1110	0.1665	0.0833	0	0.0555	0.0555
15	0.1110	-0.0555	0.0555	0.2220	0	0.0555
16	0	0	0.0278	0	0.0555	0.0555

Table 4.6: Numerical results: errors of the constraints used in nonlinear solvers (10^{-015*})

	$ \Re_{\bar{x}} - 1$	$ \Re_{\bar{y}} - 1$	$ \Re_{\bar{z}} - 1$	$\Re_{\bar{x}} \cdot \Re_{\bar{y}}$	$\Re_{\bar{x}} \cdot \Re_{\bar{z}}$	$\Re_{\bar{y}} \cdot \Re_{\bar{z}}$	$\det(R)-1$
1	-0.1110	0	0.2220	0.0555	0	0	0
2	0	0.2220	0	-0.0278	0.0555	-0.1665	0.2220
3	0	0	0	0.0278	-0.0555	-0.0069	0
4	0	-0.1110	0	-0.0278	0.0902	0.0555	0
5	0.2220	0	-0.2220	0.0278	0.0278	0.0555	-0.2220
6	0	-0.2220	-0.2220	0.0278	-0.0555	0	-0.1110
7	-0.1110	0	0	0.0833	0.0555	0.0555	0
8	0	-0.1110	0	0	0	-0.0555	0
9	-0.1110	-0.1110	-0.1110	0	0	-0.0278	-0.1110
10	0	-0.1110	0	0	-0.0555	0.0555	0
11	0	0	0	0	-0.0555	0	0.2220
12	0	-0.1110	0	0	0	0.0347	0
13	0.2220	-0.1110	-0.1110	0.0278	0.1110	-0.0694	-0.1110
14	-0.1110	-0.1110	-0.2220	-0.0555	-0.0625	-0.0555	-0.2220
15	0	-0.1110	0.2220	0	0.0555	-0.0555	0
16	0	0	0	0	-0.0139	-0.0555	0.2220

a: 1,7,11,15

b: 2,6,8,10

c: 3,9,14,16

d: 4,5,12,13

Figure 4.4: Four types of surface shapes. Each shape corresponds to four rows in each table, the row numbers are listed in the caption of each shape.

4.4 Conclusion

In this chapter, we analyzed the relation of relighting error and surface shape. We also show that even under Lambertian model, with image derivatives up to second order, we cannot uniquely determine the surface shape. In practice, due to the dynamic range of the images, it is very hard to estimate second order derivatives from intensity measurements. Estimating higher order derivatives from finite differences for use in similarity matching cannot improve relighting results. Therefore, a multiple image approach will be discussed in next chapter to overcome the limitation of the single image approach, it can be applied on more complex objects, including those with non-spherical and non-Lambertian surfaces.

Chapter 5

Multiple Image Relighting

5.1 Introduction

Much previous research on relighting makes strong assumptions on the geometry and reflectance properties. For example, in facial image analysis, there are many approaches that assume a face is convex and Lambertian [30, 8, 82]. In Chapter 3, it was shown that when given a single image of a convex Lambertian object, an image of the object under a different lighting condition can be synthesized via a pair of images of a different reference object, so long as one of the reference images was taken under the same illumination condition as the input image, while the other was taken under the desired illumination condition. However, the convex Lambertian assumption does not hold for most objects, *e.g.* human faces are neither fully convex nor Lambertian. Although various methods have presented visually pleasing results using convex Lambertian models, a more detailed inspection will reveal that non-Lambertian effects, such as cast shadows and specular highlights, are missing.

In order to measure and model more complex BRDF without knowing or explicitly

acquiring 3D geometry, image-based approaches have been explored. Biquadratic polynomials [48] and Cartesian tensors [7] were proposed for compactly modeling ABRDF. These models are fitted using multiple observed images of an object, each taken under a different known lighting condition. Subsequently, an image of the object under a novel illumination can be synthesized. As the images under novel lighting conditions are interpolated or extrapolated from available ones, the illumination directions of the observed images have to be uniformly distributed for better results. The impact of the distribution of observed illumination conditions on tensor-based method was discussed in [41].

While the problem of illumination transfer has attracted a lot of attention recently, many approaches focus on certain classes of objects, such as human faces [61, 17, 18, 71]. The problem of relighting general objects, as what we are addressing in this thesis, have not been as broadly studied.

In this chapter we propose a novel framework for image relighting, which is not only more accurate but also more robust than model fitting methods. Given a few images of an object captured under different illumination, model fitting methods utilize these images to estimate the reflectance functions then predict the images under other illuminations. However, in our proposed framework, these images are used to find the patches with the similar reflectance functions in an reference database. Then new images of the input object can be synthesized under other illumination conditions that are present in the example data. This approach does not require known parameters of the lighting direction, only that matching is carried out with reference images taken under the same illumination. By first modelling the reflectance properties of the example objects with available data, we can further synthesize the images of the input object under lighting conditions that were not captured in the reference database. We also propose methods to alleviate problems caused by different illumination or differ-

ent albedos between example and test data. The framework is demonstrated on real images from standard databases and our own collected data.

The key contributions of this work are as follows:

- To the best of our knowledge, this is the first example-based relighting method which does not assume that test objects to be relit share the same shape or even topology as the reference objects.
- We show how our example-based method can be extended to relight objects under illumination conditions not available in the reference data, which has not been done before in previous example-based methods.
- Our method outperforms state-of-the-art polynomial-based model fitting methods.

5.2 Framework Overview

In this chapter, we describe our framework for relighting by matching and transferring pointwise apparent BRDF exemplars. Our goal is to be able to synthesize novel images of an input object under a wide range of lighting conditions, given only observed images of that object under a few sample lighting conditions. In order to synthesize the images under other illumination conditions rather than the ones observed in these sample images, instead of fitting an ABRDF polynomial-based model based on these images [48, 7], we propose to relight images of objects by transferring ABRDF from images in a reference database. The reference database consists of a number of reference objects taken under a very wide range of known illumination conditions, but the reference objects do not need to be similar to the input object.

Figure 5.1: Framework Overview. Given limited images of a new object as input, for each point on this object, the best match point in reference database is found by searching for the near neighbour in a vector space, where each dimension of the vectors is corresponding to the image intensity level of the points under one lighting condition. Then new images of the object can be synthesized through direct or modeled ABRDF transfer.

For each point on the input object, the assumption is that we know the intensities under a limited set of known lighting conditions. This set of intensities represents an *index* that allows us to find matching points in the reference database that have a similar set of intensities for the same lighting conditions. Since the reference database contains many other lighting conditions, this allows us to extract an accurate data-driven estimate for the ABRDF of that input object point, even if the information comes from a point on a reference object with a totally different overall shape or geometry. See figure 5.1.

This ABRDF is more accurate than one directly modeled by fitting a polynomial function from a limited set of lighting conditions, as established in our experimental comparisons. The experiments are conducted both on face-specific datasets: the original and Extended Yale Face Database B [30, 45], as well as on general object datasets: Birmingham Object Lighting Database (BOLD) [36] and our own collected data.

5.2.1 Similarity Matching

Suppose we have sample images $\mathcal{I} = \{I_1, I_2, \dots, I_M\}$ of the input object, where M is the total number of images available. Each image I_m is taken under certain directional lighting condition L_m , so in total there are $\mathcal{L} = \{L_1, L_2, \dots, L_M\}$ lighting conditions. We want to render the object in arbitrary lighting conditions, but here we focus on cases with one directional lighting source. It is possible with some effort to generalize the method to lighting conditions based on linear superposition or artificial composition [3], but we will not address these in this thesis.

The algorithm will first search for the closest matching patches between the input object and the example objects of the pre-collected reference database. In this reference database, all objects are captured under various directional lighting conditions $\bar{\mathcal{L}}$ where $\bar{\mathcal{L}} \supset \mathcal{L}$, *i.e.* $\bar{\mathcal{L}} = \{L_1, L_2, \dots, L_M, \dots, L_N\}$ where $N \gg M$. So every example object has N images, $\mathcal{I}^k = \{I_1^k, I_2^k, \dots, I_M^k, \dots, I_N^k\}$, $k \in \{1, 2, \dots, K\}$ where K is the number of objects in the database. For every point (u, v) on input object, where (u, v) is the pixel coordinate to which the point is projected in the input images, we find the best-matched point in the database by solving:

$$(k^*, x^*, y^*) = \underset{(k, x, y)}{\operatorname{argmin}} \sum_{m=1}^M |I_m(u, v) - I_m^k(x, y)|, \quad (5.1)$$

where the triplet (k, x, y) indexes the corresponding point in the example data, k is the

id of the object and (x, y) is the pixel coordinate to which the point is projected in the example images. It is solved by nearest neighbor search, and we use an approximate nearest neighbor search method [4, 52] that provides large speedup with minor loss in accuracy. Our hypothesis is that if the intensities of two object points are similar under the sample lighting conditions, their appearances under other lighting conditions are also likely to be similar. In other words, the ABRDFs can be matched based on a limited index comprising pixel intensities obtained under an appropriate set of sample lighting conditions.

5.2.2 Direct ABRDF Transfer

If the desired illumination condition for the novel relighting task is already one of the N illumination conditions in the reference database (but obviously not one of the M observed lighting conditions for the input image), then the pixel intensity transfer is straightforward.

After the best match between the input object point and the reference object point is found, we simply assign the intensity value for the desired lighting condition L_n from example point to the input point as:

$$I_n(u, v) = I_n^{k^*}(x^*, y^*), \quad (5.2)$$

where $n \in \{1, 2, \dots, N\}$.

5.2.3 Modeled ABRDF Transfer

We also consider situations when the lighting conditions in the reference database are not comprehensive enough. In order to synthesize images of input object under

lighting conditions that are not available in the reference database, we propose to parametrically model the ABRDF of example objects in the database using all N known lighting conditions, before carrying out the intensity transfer. This is expected to be much more accurate than directly modeling the ABRDF from just the M known lighting conditions for the input object.

We considered two parametric models: Cartesian tensors and bivariate polynomials. These models are able to represent a multi-lobed reflectance function more appropriate for more realistic rendering. These compare well to simpler models that are able to only handle single-lobed Lambertian reflectances.

ABRDF modeled as Cartesian Tensors

Cartesian tensors of order d can be represented in the form:

$$\mathcal{T}(L) = \sum_{a+b+c=d} T_{abc} L_x^a L_y^b L_z^c, \quad (5.3)$$

where T_{abc} are the tensor coefficients, and $L = [L_x, L_y, L_z]^T$ is the unit vector denoting the direction of illumination.

Using the Lambertian model, $I = \rho N \cdot L$, where ρ is the surface albedo and $N = [N_x, N_y, N_z]^T$ is the unit vector in the direction of surface normal. This is a special case of (5.3) with $n = 1$, *i.e.* $T_{100} = \rho N_x, T_{010} = \rho N_y, T_{001} = \rho N_z$. We can use higher order tensors to model more complex ABRDF model. The number of the coefficients to be fitted, P , is given by the weak composition of partitioning d into $s = 3$ parts, *i.e.* $P = \binom{d+s-1}{s-1} = \binom{d+3-1}{3-1} = \binom{d+2}{2}$.

Given M observed images:

$$I_m = \sum_{p=1}^P T_p \ell_{mp}, \quad (5.4)$$

where $m \in \{1, 2, \dots, M\}$ (*i.e.* M equations with P unknowns) and ℓ is the tensor basis function.

In matrix form it is:

$$\mathbf{I} = \mathbf{L}\mathbf{T}, \quad (5.5)$$

$$\text{where } \mathbf{I} = \begin{pmatrix} I_1 \\ I_2 \\ \vdots \\ I_m \end{pmatrix}, \quad \mathbf{T} = \begin{pmatrix} T_1 \\ T_2 \\ \vdots \\ T_p \end{pmatrix}, \quad \mathbf{L} = \begin{pmatrix} \ell_{11} & \ell_{12} & \cdots & \ell_{1p} \\ \ell_{21} & \ell_{22} & \cdots & \ell_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \ell_{m1} & \ell_{m2} & \cdots & \ell_{mp} \end{pmatrix}.$$

The coefficients are estimated via least squares fitting:

$$\hat{\mathbf{T}} = \mathbf{L}^+\mathbf{I}, \quad (5.6)$$

where $\mathbf{L}^+$ is the pseudoinverse of $\mathbf{L}$.

ABRDF modeled as Bivariate Polynomials

Bivariate polynomials of degree d can be represented in the form that is similar to Cartesian tensors:

$$\mathcal{P}(L) = \sum_{a+b+c=d} P_{abc} L_x^a L_y^b 1^c. \quad (5.7)$$

It can be fitted in a similar manner to that for Cartesian tensors. First establish the equations in matrix form:

$$\mathbf{I} = \mathbb{L}\mathbf{P}. \quad (5.8)$$

Then solve it as:

$$\hat{\mathbf{P}} = \mathbb{L}^+ \mathbf{I}, \quad (5.9)$$

where $\mathbb{L}^+$ is the pseudoinverse of $\mathbb{L}$, which is constructed with the polynomial basis functions.

The linear least squares fitting with nonlinear basis functions is quite efficient because singular value decomposition (SVD) only needs to be done once to get the pseudoinverse, which can be used for every pixel. This is much faster than repeatedly fitting a different nonlinear analytical model for each pixel.

Transfer Method

After the chosen model of ABRDF has been estimated, we have

$$\text{ABRDF}(u, v) = \text{ABRDF}^{k^*}(x^*, y^*), \quad (5.10)$$

i.e.

$$\alpha_p(u, v) = \hat{\alpha}_p^{k^*}(x^*, y^*), \quad (5.11)$$

where α 's are the parameters of the model, $p \in \{1, 2, \dots, P\}$ is a parameter index, and P is the total number of parameters.

We want to estimate the pixel intensity in a new desired lighting condition l . In both the Cartesian tensor and bivariate polynomial models, the intensities can be represented as linear combinations of nonlinear basis functions β_p with coefficients α_p .

The computed intensity in the novel lighting condition is:

$$I_l(u, v) = \sum_{p=1}^P \alpha_p(u, v) \beta_{lp} = \sum_{p=1}^P \hat{\alpha}_p^{k*}(x^*, y^*) \beta_{lp}. \quad (5.12)$$

Note that (5.10) is a generalization of (5.2), where the actual samples of the ABRDFs are available.

5.2.4 Comparison with Model Fitting Methods

We analytically compare the proposed example-based relighting framework to direct model fitting methods.

Although higher order models can approximate the reflectance function better, more observed images are needed to avoid overfitting. As mentioned earlier, $P = \binom{d+2}{2}$ unique coefficients need to be estimated for d th order tensors or bivariate polynomials of degree d , that is 3 ($d = 1$), 6 ($d = 2$), 10 ($d = 3$), 15 ($d = 4$), 21 ($d = 5$), 28 ($d = 6$), 36 ($d = 7$), 45 ($d = 8$) and 55 ($d = 9$). Because we only expect limited samples to be available for the input object (as opposed to objects in the reference database), a relatively low order model has to be chosen to avoid this over-fitting problem.

It has been shown [41] that anti-symmetric (odd order) tensors introduce less visual artifacts than symmetric (even order) ones. In keeping with this previous literature, we also chose to use third order tensors, and picked images under 9 lighting conditions as observed. This leads to an underconstrained problem, which is regularized by obtaining the solution with the smallest norm.

As we are currently only interested in pointwise image relighting, and do not want to confound our research findings with the assumption of smooth ABRDF variation across the area of the object as made in [41], we therefore avoid the use of spline

regularization.

For polynomial-based methods, we find that biquadratic polynomials [48] give the best results with images under nine lighting conditions.

In the experiments we find that example-based method is quantitatively and visually better than the model fitting method, and it is also much less sensitive to the availability of the illumination samples already observed than these model fitting methods. Besides, in order to use model fitting methods, the lighting parameters have to be known, while with example-based method, to synthesize images under observed lighting conditions in examples, we only need to know the images are under the same illuminations during matching.

5.2.5 Color Consistency and Intensity Normalization

The best results of our method obviously come in scenarios when the reference database contain lighting conditions that are nearly identical to those in the input image. In practice, this can differ significantly due to differences in the intensities and types of light sources, and also inaccuracies in the lighting directions. Additionally, we may have to deal with variations in color of the reference objects and light sources.

To deal with these problems in a number of ways. First, we apply our method only to the luminance channel of the images, while keeping the chrominance channels unmodified. Second, we propose an intensity normalization scheme to correct the image intensity. Here we further express ABRDF as $n * \text{NABRDF}$, where n is the norm of ABRDF, and NABRDF is the normalized ABRDF, treated as a vector. This norm n is linearly related to both the intensity of the light source as well as the albedo of the object point, while the NABRDF changes only with lighting direction. Instead of directly matching the ABRDF (image intensities), the NABRDF is to be matched

instead.

However, a potential ambiguity arises on whether matching a *normalized index* of the NABRDF (a normalized M dimensional vector, comprising just the M lighting conditions) is the same as matching the full NABRDF (a normalized N directional vector). We will show this is true.

Suppose the full NABRDF vector is scaled by a factor p , with this scaling selected to normalize only the partial M dimensional NABRDF vector. Then using subscripts t and e to refer to “test” and “example”:

$$\begin{aligned} p\text{NABRDF}_t &= p\text{NABRDF}_e \\ \Rightarrow \text{NABRDF}_t &= \text{NABRDF}_e. \end{aligned} \tag{5.13}$$

As

$$\frac{pn_t}{pn_e} = \frac{i_t\rho_t}{i_e\rho_e} = \frac{n_t}{n_e}, \tag{5.14}$$

where i ’s are light intensities, ρ ’s are albedos, pn is the norm of the matching part of ABRDFs and n is the norm of the full ABRDFs. Then

$$\begin{aligned} \text{ABRDF}_t &= n_t * \text{NABRDF}_t \\ &= \frac{pn_t}{pn_e} n_e * \text{NABRDF}_e \\ &= \frac{pn_t}{pn_e} \text{ABRDF}_e. \end{aligned} \tag{5.15}$$

It means that after we transfer the intensities from example data, we need to scale the values by the ratio of the norms of partial ABRDFs.

5.3 Experiments

The proposed example-based relighting framework is evaluated by conducting both face-specific experiments on the Yale Face Database B (YalefaceB) and the Extended Yale Face Database B (extended YalefaceB) [30, 45], as well as general object experiments on the Birmingham Object Lighting Database (BOLD) [36] as well as our own collected data.

For both face and general object categories, we conducted experiments in which input images were synthesized to illumination conditions that were already in the reference database (which we will call the “observed” experiments), as well as entirely novel illumination conditions (called the “unobserved” experiments).

YalefaceB and extended YalefaceB databases collected the data in the same format. In these databases there are a total of 38 subjects, each of which is captured under 64 illumination conditions. Because some images are either corrupted during acquisition or contain substantial noise, we limited our experiments to 20 subjects which had good quality data, each with 49 illumination directions from the front hemisphere facing the subjects.

In the “observed” experiments, we took these 49 illumination conditions to be the full set of illumination conditions in the reference database, while 9 out of 49 were the lighting conditions that were available for the input test object (and therefore form the matching index).

In the “unobserved” experiments, the difference is that the desired illumination condition is one out of the 49 conditions but assumed to be unavailable in the lookup procedure, and therefore must be synthesized through fitting a ABRDF model from the remaining 48 observed lighting conditions.

In order to demonstrate the generality of our method, we also conducted experiments on the BOLD database, which contains high quality color images of general objects. In the “objects” category of BOLD, each object is captured under 56 illumination conditions which are coming from the front hemisphere above horizon. In the “observed” experiments, all 56 lighting conditions were treated as observed ones in the reference database, while 9 were available for the input test object. The “unobserved” experiments here are similar to the ones for the face databases, whereby one out of the 56 conditions is assumed to be unavailable and to be synthesized. Note that there are two cameras in the BOLD setup for capturing stereoscopic image pairs. In our experiments, we chose the images that are from the left camera, and the illumination azimuths were adjusted according to the camera’s toe in.

Finally, we also collected some of our own data to evaluate the robustness of the methods to differences in light sources and material properties between example and test data. Our data were captured under 18 selected illumination conditions out of the original 56 conditions in the BOLD database.

5.3.1 Results from the “Observed” Experiments

First, we did a cross validation experiment on the 20 subjects from the YalefaceB databases. Each time, we chose one subject as the test subject, while the other 19 subjects form the reference database.

Note that if a test object is contained in the reference database, we will get perfect relighting results. This is the ideal situation in which we are able to find an exact match for a point on the test object in the reference database.

When the test object is not contained in the reference database, the best matches are the ones that minimize the combined intensity differences under *all* 49 illumination

Figure 5.2: Comparison between Example-based Relighting and Model Fitting Results (Observed in Examples). Database: YalefaceB. “PF” [48], “TF” [41] and “ER” means Polynomial Fitting, Tensor Fitting and Example-based Relighting, respectively. The number before them is the order of the model, and the number after is the number of input images for fitting or matching. Same abbreviations will be used in following figures.

conditions. The average absolute difference per pixel for each subject is shown in blue in figure 5.2. This forms the baseline for our method, as it is the upper bound of our performance given the dataset used – it is essentially determined by the dissimilarity in the shapes of the test objects from the reference objects.

Next, we investigated the outcome when ABRDF matching is only limited to the 9 illumination conditions available for the test object. Note that this number of illumination conditions is in keeping with those used in previous works [45, 41]. To be a fair comparison, all methods use 9 images under these illumination conditions as input, including Polynomial Fitting(PF) [48], Tensor Fitting(TF) [41] and our Example-based Relighting(ER) method.

After finding the best matched ABRDF index per pixel, we transferred the inten-

sities under other illuminations from the matched examples to the test pixel. The average absolute differences between the synthesized and actual images of the test object under all observed lighting conditions are shown as green asterisks. We observe that the results for our example-based relighting method are quite close to the baseline.

Figure 5.3: Comparison between Example-based Relighting and Model Fitting Results (Observed in Examples): One case. Database: YalefaceB. The input images for fitting or matching are marked by crosses, and the test cases are marked by circles. For better visualization, the displayed error images are amplified by a factor of 3 and shown in negative gray-scale. For the abbreviations, please refer to figure 5.2. Large version of the results can be found in Appendix A.

Figure 5.4: Comparison between Example-based Relighting and Model Fitting Results (Observed in Examples): One case, with bad samples. Database: YalefaceB. Large version of the results can be found in Appendix A.

Figure 5.5: Comparison between Example-based Relighting and Model Fitting Results (Observed in Examples): One case. Database: BOLD. For the explanation of abbreviations and the meaning of error images, please refer to figure 5.2 and 5.3. Large version of the results can be found in Appendix A.

We compared our method against those using third order tensor and biquadratic polynomial models. The average differences to the ground truth are shown as green circles and green triangles, and we can see that these do not perform as well compared to our example-based method.

One of the 20 cross validation tests is shown in figure 5.3. We can see that the results from our example-based method looks visually similar to the ground truth, especially under harsher lighting conditions. Note that the 9 lighting conditions used in the matching are denoted as crosses in the left bottom of figure 5.3.

Another advantage of the example-based method is that it is much less sensitive to the choice of illumination conditions in the input images, *e.g.* as shown in figure 5.4, a poor choice of illumination conditions for the test object often led to poor model fitting. However, the example-based relighting method is much more robust to this choice of illumination conditions. We ran hundreds of tests where we randomly selected 9 illumination conditions used in the test images. In these tests, the average errors of example-based method remain very constant, while huge error variations are seen for the model fitting methods.

The experiments with high quality color images are conducted on BOLD. Here we pick a few of objects with most complex shapes or materials as examples. Each color channel is treated independently. Example results of relighting one general object are shown in figure 5.5.

5.3.2 Results from the “Unobserved” Experiments

Similar to the previous experiments on YalefaceB databases, each time we chose one subject as the test subject, while the other 19 subjects were considered as part of the reference database.

Figure 5.6: Comparison between Example-based Relighting and Model Fitting Results (Unobserved in Examples). Database: YalefaceB. Besides the abbreviations used in figure 5.2, “W” is short for “With”.

First we estimated the ABRDF of example subjects with all available lighting conditions. Now because we have sufficient data in reference database, we can approximate the reflectance function with a higher order model with higher degrees of freedom. In each test, one condition out of 40 conditions (9 out of 49 are always observed) is considered as unobserved, while the other 48 conditions are observed.

Our method involved determining the best example object point in the reference dataset based on matching 9 illumination conditions of the input test object, after which a parametric model is fitted to the 48 illumination conditions obtained from the example object point.

It turns out that the best choice of parametric model for estimating the unobserved sample is data-dependent. In our experiments, we chose the bicubic polynomial model which has the best predictive power when fitting images under 48 illumination conditions. As shown in figure 5.6, green points are the average errors of direct model fitting

Method	YalefaceB (O)	YalefaceB (U)	BOLD (O)	BOLD (U)
2PF	11.1	12.0	11.7	13.0
3TF	12.1	14.8	13.3	15.5
ER (Ours)	6.2	11.0	5.4	11.4

Table 5.1: Average errors for the methods mentioned in this chapter. Besides the abbreviations used before, “O”, and “U” are short for Observed and Unobserved.

methods, while magenta points are the average errors of using our method. We can see the proposed method is also better than direct model fitting methods.

Two of the tests on YalefaceB databases and BOLD database are shown in figure 5.7 and figure 5.8 respectively.

In the experiments we find that example-based method is quantitatively better, and in most cases, visually closer to the ground truth than model fitting methods, and it is also much less sensitive to the availability of the illumination samples already observed. The average errors of all test cases are shown in table 5.1. As we are currently only interested in pointwise image relighting, in some cases (especially for the direct ABRDF transfer, in which the ABRDF is not smoothed as in modelled ABRDF transfer), our results may look a bit noisy as the transferred ABRDFs are independent, not necessarily smoothly varying across the area of the object. While in model fitting methods, the ABRDFs locations are fixed, so the smoothness is maintained. To alleviate this problem, the same poisson reconstruction technique mentioned in Chapter 3.3 can be applied. One example result from screened poisson solution can be found in figure A.1.

Figure 5.7: Comparison between Example-based Relighting and Model Fitting Results (Unobserved in Examples): One Case. Database: YalefaceB. For the explanation of abbreviations and the meaning of error images, please refer to figure 5.2 and 5.3. Large version of the results can be found in Appendix A.

Figure 5.8: Comparison between Example-based Relighting and Model Fitting Results (Unobserved in Examples): One Case. Database: BOLD. Large version of the results can be found in Appendix A.

5.3.3 Color Consistency and Intensity Normalization

As shown in figure 5.9, if the all example data are with different color and intensity than the test data, the results of the direct example-based relighting method on each RGB channel looks like either the objects are painted with the same material as the example objects or they are illuminated by the same light source. We can solve the unequal color problem by applying the example-based method only on luminance channel, and unequal intensity problem can be alleviated by the proposed normalization scheme. Results are shown in the fourth and fifth column of figure 5.9 respectively.

Figure 5.9: Color Consistency and Intensity Normalization. Besides the abbreviations used in figure 5.2, “L” means luminance and “N” means normalization. The displayed error images are amplified by a factor of 3 and shown in negative gray-scale.

5.3.4 Analysis

The crosses in figure 5.3, 5.4, etc. are the lighting conditions used for matching. In figure 5.3 they are chosen as in [41], which are well distributed in the lighting direction space, while in figure 5.4, the 9 conditions are randomly chosen. We can see from the figures that model fitting methods can reasonably handle the problem with well distributed inputs, but get much worse otherwise. Because essentially the interpolation and extrapolation of the model fitting methods heavily depends on the samples already known. But our proposed example-based method is much less sensitive to the choice of illumination conditions in the input images. Average results are shown in figure 5.2, 5.6 and table 5.1, where we can see our results (ER) are quantitatively better than the model fitting methods (PF and TF). Referring to the ground truth, we can see that most “failure” cases of the model fitting methods are under harsher lighting conditions, it suggests that extrapolation is difficult. However, with our methods, the errors are not significant in those cases. Because our method with direct transfer (“observed”) only depends on the complexity of ABRDFs, not the complexity of the chosen model and the availability of the inputs, while in our method with indirect transfer (“Unobserved”), the transferred model can be fitted with sufficient training data, so over-fitting can be reasonably avoided. Therefore, overall, the example-based method is more accurate and robust.

The implicit assumption of this work is that the high dimensional ABRDFs of the objects actually lie in a low dimensional manifold in the lighting direction space. PCA analysis produces the expected results shown in figure 5.10, where about 95% of variance are accounted by the the first 5 principal components for YalefaceB databases. That means 5 illumination conditions are sufficient if well-chosen. In the experiments, it is found that reducing from 9 images to 5 is not so crucial for example-based method. Our method performs well when the matched observations can reasonably cover the

complexity of the ABRDFs. However for model fitting methods, the interpolation or extrapolation heavily relies on the availability of samples and the over-fitting problem can hardly be avoided.

Figure 5.11 reveals the strong correlation between the best match distance of 9 illuminations and the error of relighting results under other 40 lighting conditions, which suggests the possibility of a hybrid framework, in which other methods may be used to reduce the relighting errors of the points with high match distances. The outliers mostly correspond to points that have complex ABRDFs. This suggests that only ABRDFs with a relatively small amount of lobes can be matched using limited observations. In figure 5.12, the correlation of the errors from example-based relighting and polynomial fitting are shown in two cases. On the left, there is only example data from one face, while on the right, there is sufficient example data from all other faces. Since a point above $y = x$ means that the example-based method is better, we improve the relighting results by providing more example data.

Figure 5.10: PCA Analysis. Database: YalefaceB.

Figure 5.11: Relation between the Best Match Distance and Example-based Relighting (ER) Error. Warmer colors mean higher density of points. The corresponding points of the outliers are shown on the right. Database: YalefaceB.

Figure 5.12: Relation between Example-based Relighting (ER) Error and Polynomial Fitting (PF) Error, with (right) or without (left) sufficient example data. Warmer colors mean higher density of points. Database: YalefaceB.

5.4 Conclusion

In order to avoid 3D acquisition as required in geometry-based relighting methods or laboratory setup for each object as required in traditional image-based relighting

approaches, an example-based relighting framework is proposed. In this framework, there are a number of images of reference objects captured under different illumination conditions. Given limited images of a new test object captured under a small subset of the reference illumination conditions, new images can be synthesized for the input object under other illumination conditions, which may or may not be present in the reference images. Much previous research on relighting makes strong assumptions on the geometry and reflectance properties of the surface. In contrast, our method is a purely data-driven technique. The input data is similar to that used in traditional image-based relighting methods, that is a set of images of the input object taken from a fixed viewpoint under different lighting conditions. However, the number of the required images can be significantly reduced using the proposed example-based framework. This is achieved through matching of pointwise ABRDF exemplars, and do not require that the reference objects to have similar shape or even topology to the input object. Our hypothesis is that by matching image intensities (essentially the samples of the ABRDFs) of two points under several lighting conditions, the ABRDFs of the two points are also well-matched. We can then either use the ground truth samples of the example ABRDF, or a fitted example ABRDF to render the input object under other lighting conditions. The effectiveness of the framework is demonstrated on standard real image databases and our own data, with comparisons showing that the method outperforms existing image-based model fitting methods.

Chapter 6

Conclusion and Future Work

6.1 Summary of Research

An image captures the appearance of a scene under a given condition. In many cases some of the physical conditions such as camera setup, atmosphere medium, surface properties, and illumination are not desirable at the moment of capture. People are interested in enhancing and extending the capabilities of photography by post-processing the images, *e.g.* deblurring, refocusing, dehazing and recoloring. The central theme of this thesis is to investigate methods for virtually changing the illumination conditions post-capture. In order to avoid 3D acquisition which is required in geometry-based relighting methods, and a laboratory setup for each object which is required in traditional image-based relighting approaches, we focus on exploring approaches using a set of reference data.

In Chapter 3, we proposed an analogy framework that assumed spherical surface patches and the Lambertian reflectance model. Analysis shows it to be feasible, both theoretically and experimentally. This approach takes only one image as input, and

finds the best reference patches from example images. Both image intensities and image gradients have theoretically been shown to be useful for finding the correspondence in the context of analogy-based relighting. Images under other illuminations then can be synthesized via lookup. By applying Poisson-derived methods, problems in relighting real images are alleviated with improved visual quality.

The initial approach is capable of handling predominantly convex spherical surfaces, which means at each local surface patch, the two principal curvatures are equal. In Chapter 4, we establish a relation between the relighting error and the deviation from spherical patch, with analysis extended to non-spherical surfaces and image derivatives up to second order. The nonlinear mathematical framework which relates image and surface under the Lambertian model was derived. Numerical results show that under the same lighting condition, four types of surface patches can result in the same combination of image intensity, first order derivatives and second order derivatives. This ambiguity is the limitation of the single image approach.

In Chapter 5, both strong assumptions on the geometry and reflectance are relaxed by using a multiple image data-driven technique. The input data is similar to traditional image-based relighting methods, that is a set of images of the input object taken from a fixed viewpoint under different lighting conditions. However, the number of the required images can be significantly reduced using the proposed example-based framework. Our hypothesis is that by matching image intensities (essentially the samples of the ABRDFs) of two points under several lighting conditions, the ABRDFs of the two points are also well-matched. We can then either use the ground truth samples of the example ABRDF, or a fitted example ABRDF to render the input object under other lighting conditions. We demonstrate this framework on several databases, which includes grayscale face images, as well as textured and colored images of general objects. We also collected some data under illumination with intensity and color that

differ from standard database, and tested our proposed color consistency and intensity normalization solutions. As shown in the experiments, our approach is more robust. Our results are quantitatively better, and in most cases, visually closer to the ground truth than those from model fitting methods.

The example-based framework discussed in this thesis provides a new way of relighting images. Explicit measurements of the geometry and reflectance can be avoided. Also, unlike previous example-based methods, there is no assumption of same shape or even topology between test and reference objects.

6.2 Future Work

The research described in this thesis includes theoretical analysis and experimental verification of the example-based framework for relighting images. Single-image and multiple-image approaches both have many potential research directions.

6.2.1 Single Image Relighting

In the experiments, we found an approximate linear relation between the shape index of local patches and the relighting errors that occurs as the local shapes deviate from a spherical surface. Given that it is very difficult to estimate the local shapes from traditional shape-from-shading techniques, an alternative approach would be to detect the class and pose of the object. The class and pose would then be able to provide a good prior for the shape index of the local patches themselves. We can explore the effectiveness of simple template-matching techniques to more advanced object recognition methods. Separately, illumination and reflectance invariant shape detectors have been initially studied in [55]. Stark et al. [73, 74] recently investigated on object class de-

tection using shading cues. Such methods can be used to locate the objects or portion of the scene which have elemental shapes (sphere, cylinder, *etc.*).

It may be possible to classify the surface points in the image into two categories (convex and concave) using shading cues by assuming the overall shape of the object is convex. Such binary labels can guide the relighting synthesis to preferentially select convex and concave examples.

While the initial approach is capable of handling curved surfaces but not flat surfaces, many images like typical streetscape photos contain vertical cylinders as well as planes. It would be helpful to integrate other surface orientation estimation methods into the framework for relighting more general scenes. The orientation of the planes can be estimated from multiple approaches, such as plane normals from vertical corners, and shape from texture for highly textured surfaces. Although texture is one of the primary problems for our method due to the variations in albedo, the use of other texture-based features may allow for correctly matching textured surface patches with the same orientation. Furthermore, 3D shape hypothesis verification methods may also be applied based on image evidence.

6.2.2 Multiple Image Relighting

For multiple image relighting approach, there are two remaining questions could be further studied:

- How many images do we need for each input object, and under which lighting conditions? For economic and portability considerations, this requirement should be as few as possible.
- How many images do we need for an example object, and under which lighting

conditions? Though we can collect sufficient data of the example objects under various lighting conditions, it is reasonable to keep the database as small as possible.

These questions are related to the complexity of the object geometry and reflectance.

As we are currently only interested in pointwise image relighting, our results may look a bit noisy as the transferred ABRDFs are independent, not necessarily smoothly varying across the area of the object. However, the quantitative results show our approach still outperforms other methods. Poisson reconstruction can be used to reduce the gradient noise. Besides, in future, we plan to refine pointwise relighting results with other constraints.

As the analysis of the results show there is a strong correlation between the best match distance and relighting errors, this suggests the possibility of a hybrid framework, where other methods may be used to reduce the relighting errors of the points with high match distances.

6.2.3 Overall

Future research may involve relaxing some of these assumptions.

We assume the scenes are illuminated by a dominant directional lighting which approximates to a distant light indoor and the sun light outdoor. Extensions on more complex lighting model are obvious directions of future research.

The present solution is based on a purely local approach. Some global constraints may be introduced to achieve better results. For example, the relighting result of a natural scene should have an image edge distribution that fits a Laplace distribution.

Flash and non-flash images have been combined to enhance digital photographs,

through de-noising, detail transfer and tone mapping. In our context, we may consider these images to be of the same scene under different illumination conditions: one with the original environmental illumination, and the other illuminated by a dominant point light source. This may allow flash images of other scenes to be relit under a different environmental illumination.

Relighting cast shadows is exceedingly difficult, as it requires knowledge of the geometry of the object as well as the layout of the scene. In near future, if inexpensive depth cameras become available, images with rough depth information might become the standard output of digital cameras. The depth information may be very useful for solving the shadow relighting problem.

Besides exploring physically correct solutions, research on human vision [43] reveals that people usually relate lower luminance values to greater surface depth. So we may assume that a perceptually appropriate surface depth map can be directly estimated from the image intensities. The plausibility of such assumption has been demonstrated in a material editing application [38]. If the similarity between two objects can be matched based on perceptual structure, the perceptually appropriate relighting of an object in a different environment may be synthesized from the samples of another object.

Appendix A:

Large Results of Multiple Image Relighting

In all following figures, PF, TF and ER denote Polynomial Fitting, Tensor Fitting and Example-based Relighting, respectively. The number before them is the order of the model, and the number after is the number of input images for fitting or matching. “W” and “PR” are short for “with” and Poisson Reconstruction.

For better visualization, the displayed error images are contrast amplified by a factor of 3 and shown in negative gray-scale.

As we are currently only interested in pointwise image relighting, our results may look a bit noisy as the transferred ABRDFs are independent, not necessarily smoothly varying across the area of the object, unlike model fitting methods whose smoothness is often assumed and enforced. In some cases, especially for the direct ABRDF transfer, where the ABRDF is not smoothed as in modelled ABRDF transfer, even small error in pointwise ABRDFs may cause large gradient noise, to which humans are sensitive. Poisson Reconstruction can be applied to alleviate the problem. The quantitative results show our approach outperforms other methods. In future, we plan to refine pointwise relighting results with other constraints.

(a) Ground Truth

Figure A.1

(b) 2PF9

Figure A.1

(c) Mean Error of 2PF9: 11.5724

Figure A.1

(d) 3TF9

Figure A.1

(e) Mean Error of 3TF9: 13.6704

Figure A.1

(f) ER9

Figure A.1

(g) Mean Error of ER9: 6.2452

Figure A.1

(h) ER9PR

Figure A.1

(i) Mean Error of ER9PR: 5.9371

Figure A.1: Comparison between Example-based Relighting and Model Fitting Results (Observed in Examples): One Case with good samples. Database: YalefaceB. Large version of Figure 5.3. In this high resolution version, the results of our method may appear noisy, as it is a pointwise transfer. Poisson reconstruction methods mentioned in Chapter 3.3 are used to refine the results.

(a) 2PF9

Figure A.2

(b) Mean Error of 2PF9: 16.9584

Figure A.2

(c) 3TF9

Figure A.2

(d) Mean Error of 3TF9: 29.3583

Figure A.2

(e) ER9

Figure A.2

(f) Mean Error of ER9: 5.9195

Figure A.2: Comparison between Example-based Relighting and Model Fitting Results (Observed in Examples): One Case with bad samples. Database: YalefaceB. Large version of Figure 5.4. A poor choice of illumination conditions for the test object often led to poor model fitting. However, the example-based relighting method is much more robust to this choice of illumination conditions.

(a) Ground Truth

Figure A.3

(b) 2PF9

Figure A.3

(c) Mean Error of 2PF9: 13.4763

Figure A.3

(d) 3TF9

Figure A.3

(e) Mean Error of 3TF9: 13.8039

Figure A.3

(f) ER9

Figure A.3

(g) Mean Error of ER9: 5.9492

Figure A.3: Comparison between Example-based Relighting and Model Fitting Results (Observed in Examples): One Case. Database: BOLD. Large version of Figure 5.5.

(a) Ground Truth

Figure A.4

(b) 2PF9

Figure A.4

(c) Mean Error of 2PF9: 11.3002

Figure A.4

(d) 3TF9

Figure A.4

(e) Mean Error of 3TF9: 13.4555

Figure A.4

(f) ER9W3PF48

Figure A.4

(g) Mean Error of ER9W3PF48: 9.9241

Figure A.4: Comparison between Example-based Relighting and Model Fitting Results (Unobserved in Examples): One Case. Database: YalefaceB. Large version of Figure 5.7.

(a) Ground Truth

Figure A.5

(b) 2PF9

Figure A.5

(c) Mean Error of 2PF9: 12.3996

Figure A.5

(d) 3TF9

Figure A.5

(e) Mean Error of 3TF9: 14.3751

Figure A.5

(f) ER9W3PF55

Figure A.5

(g) Mean Error of ER9W3PF55: 9.8393

Figure A.5: Comparison between Example-based Relighting and Model Fitting Results (Unobserved in Examples): One Case. Database: BOLD. Large version of Figure 5.8.

Bibliography

- [1] E. H. Adelson and J. R. Bergen. *The Plenoptic Function and the Elements of Early Vision*. MIT Press, 1991.
- [2] Y. Adini, Y. Moses, and S. Ullman. Face recognition: The problem of compensating for changes in illumination direction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):721–732, 1997.
- [3] D. Akers, F. Losasso, J. Klingner, M. Agrawala, J. Rick, and P. Hanrahan. Conveying shape and features with image-based relighting. In *IEEE Visualization*, pages 349–354, 2003.
- [4] S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. Wu. An optimal algorithm for approximate nearest neighbor searching fixed dimensions. *Journal of the ACM*, 45(6):891–923, 1998.
- [5] N. Ayache and F. Lustman. Fast and reliable passive trinocular stereovision. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 422–427, London, England, 1987.
- [6] S. Bae, S. Paris, and F. Durand. Two-scale tone management for photographic look. *ACM Transactions on Graphics*, 25(3):637–645, 2006.
- [7] A. Barmpoutis, R. Kumar, B. C. Vemuri, and A. Banerjee. Beyond the lambertian assumption: A generative model for apparent brdf fields of faces using anti-symmetric tensor splines. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–6, Anchorage, AK, USA, 2008.
- [8] R. Basri and D. W. Jacobs. Lambertian reflectance and linear subspaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(2):218–233, 2003.
- [9] A. Basso, H. P. Graf, D. C. Gibbon, E. Cosatto, and S. Liu. Virtual light: digitally-generated lighting for video conferencing applications. In *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, pages 1085–1088, Thessaloniki, Greece, 2001.

-
- [10] P. N. Belhumeur and D. J. Kriegman. What is the set of images of an object under all possible illumination conditions? *International Journal of Computer Vision*, 28(3):245–260, 1998.
 - [11] D. Beymer, A. Shashua, and T. Poggio. Example based image analysis and synthesis. AI Memo AIM-1431, MIT AI Lab, 1993.
 - [12] P. Bhat, B. Curless, M. F. Cohen, and C. L. Zitnick. Fourier analysis of the 2d screened poisson equation for gradient domain problems. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 114–128, Marseille, France, 2008.
 - [13] V. Blanz and T. Vetter. A morphable model for the synthesis of 3d faces. In *Proceedings of the International Conference and Exhibition on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pages 187–194, Los Angeles, CA, USA, 1999.
 - [14] J. F. Blinn. Models of light reflection for computer synthesized pictures. In *Proceedings of the International Conference and Exhibition on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pages 192–198, New York, NY, USA, 1977.
 - [15] A. Bloomfield. Lighting. <http://www.cs.virginia.edu/~asb/teaching/cs445-fall106/slides/06-lighting.ppt>, 2010.
 - [16] S. Boivin and A. Gagalowicz. Image-based rendering of diffuse, specular and glossy surfaces from a single image. In *Proceedings of the International Conference and Exhibition on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pages 107–116, Los Angeles, CA, USA, 2001.
 - [17] J. Chen, G. Su, J. He, and S. Ben. Face image relighting using locally constrained global optimization. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 44–57, Crete, Greece, 2010.
 - [18] X. Chen, M. Chen, X. Jin, and Q. Zhao. Face illumination transfer through edge-preserving filters. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 281–287, Colorado Springs, 2011.
 - [19] K. J. Dana, B. van Ginneken, S. K. Nayar, and J. J. Koenderink. Reflectance and texture of real world surfaces. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 151–157, San Juan, Puerto Rico, 1997.
 - [20] P. E. Debevec. Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *Proceedings of the International Conference and Exhibition on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pages 189–198, Orlando, FL, USA, 1998.

-
- [21] P. E. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar. Acquiring the reflectance field of a human face. In *Proceedings of the International Conference and Exhibition on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pages 145–156, New Orleans, LA, USA, 2000.
 - [22] Y. Dobashi, K. Kaneda, H. Nakatani, and H. Yamashita. A quick rendering method using basis functions for interactive lighting design. *Computer Graphics Forum*, 14(3):229–240, 1995.
 - [23] A. A. Efros and W. T. Freeman. Image quilting for texture synthesis and transfer. In *Proceedings of the International Conference and Exhibition on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pages 341–346, Los Angeles, CA, USA, 2001.
 - [24] A. A. Efros and T. K. Leung. Texture synthesis by non-parametric sampling. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 1033–1038, Kerkyra, Greece, 1999.
 - [25] F. P. Ferrie and M. D. Levine. Where and why local shading analysis works. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(2):198–206, 1989.
 - [26] W. T. Freeman, T. R. Jones, and E. C. Pasztor. Example-based super-resolution. *IEEE Computer Graphics and Applications*, 22(2):56–65, 2002.
 - [27] M. Fuchs, V. Blanz, H. P. A. Lensch, and H.-P. Seidel. Reflectance from images: A model-based approach for human faces. *IEEE Transactions on Visualization and Computer Graphics*, 11(3):296–305, 2005.
 - [28] M. Fuchs, V. Blanz, and H.-P. Seidel. Bayesian relighting. In *Proceedings of the Eurographics Symposium on Rendering*, pages 157–164, Konstanz, Germany, 2005.
 - [29] A. S. Georgiades, P. N. Belhumeur, and D. J. Kriegman. Illumination-based image synthesis: creating novel images of humanfaces under differing pose and lighting. In *Proceedings of the IEEE Workshop on Multi-View Modeling and Analysis of Visual Scenes (MVIEW)*, pages 47–54, Fort Collins, CO, USA, 1999.
 - [30] A. S. Georgiades, P. N. Belhumeur, and D. J. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):643–660, 2001.
 - [31] Giackop. Tutorial: Add dynamic lighting to a flat photograph. <http://psd.tutsplus.com/tutorials/photo-effects-tutorials/add-dynamic-lighting-to-a-flat-photograph>, 2008.
 - [32] T. Haber, C. Fuchs, P. Bekaert, H.-P. Seidel, M. Goesele, and H. P. A. Lensch. Relighting objects from image collections. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 627–634, Miami, FL, USA, 2009.

-
- [33] Haeberli. Synthetic lighting for photography. <http://www.graficaobscura.com/synth/index.html>, 1992.
 - [34] A. Hertzmann, C. E. Jacobs, N. Oliver, B. Curless, and D. Salesin. Image analogies. In *Proceedings of the International Conference and Exhibition on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pages 327–340, Los Angeles, CA, USA, 2001.
 - [35] A. Hertzmann and S. M. Seitz. Example-based photometric stereo: Shape reconstruction with general, varying brdfs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1254–1264, 2005.
 - [36] X. Jiang, A. J. Schofield, and J. L. Wyatt. Correlation-based intrinsic image extraction from a single image. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 58–71, Crete, Greece, 2010.
 - [37] M. J. Jones and P. Viola. Fast multi-view face detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Madison, WI, USA, 2003.
 - [38] E. A. Khan, E. Reinhard, R. W. Fleming, and H. H. Bühlhoff. Image-based material editing. *ACM Transactions on Graphics*, 25(3):654–663, 2006.
 - [39] J. J. Koenderink. *Solid Shape*. MIT Press, 1990.
 - [40] R. Kumar. *Relighting, Pose Change and Recognition of Faces*. PhD thesis, University of Florida, 2009.
 - [41] R. Kumar, A. Barmpoutis, A. Banerjee, and B. C. Vemuri. Non-lambertian reflectance modeling and shape recovery of faces using tensor splines. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(3):533–567, 2011.
 - [42] E. H. Land and J. J. McCann. Lightness and retinex theory. *Journal of the Optical Society of America*, 61(1):1–11, 1971.
 - [43] M. S. Langer and H. H. Bühlhoff. Depth discrimination from shading under diffuse lighting. *Perception*, 29(6):649, 2000.
 - [44] C.-H. Lee and A. Rosenfeld. Improved methods of estimating shape from shading using the light source coordinate system. *Artificial Intelligence*, 26(2):125–143, 1985.
 - [45] K.-C. Lee, J. Ho, and D. J. Kriegman. Acquiring linear subspaces for face recognition under variable lighting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(5):684–698, 2005.
 - [46] D. Mahajan, R. Ramamoorthi, and B. Curless. A theory of spherical harmonic identities for brdf/lighting transfer and image consistency. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 41–55, Graz, Austria, 2006.

-
- [47] D. Mahajan, R. Ramamoorthi, and B. Curless. A theory of frequency domain invariants: Spherical harmonic identities for brdf/lighting transfer and image consistency. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2):197–213, 2008.
 - [48] T. Malzbender, D. Gelb, and H. J. Wolters. Polynomial texture maps. In *Proceedings of the International Conference and Exhibition on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pages 519–528, Los Angeles, CA, USA, 2001.
 - [49] S. R. Marschner and D. P. Greenberg. Inverse lighting for photography. In *Proceedings of IS&T/SID Fifth Color Imaging Conference*, pages 262–265, Scottsdale, AZ, USA, 1997.
 - [50] V. Masselus, P. Peers, P. Dutré, and Y. D. Willems. Smooth reconstruction and compact representation of reflectance functions for image-based relighting. In *Proceedings of the Eurographics Symposium on Rendering*, pages 287–298, Norrköping, Swede, 2004.
 - [51] I. Motoyoshi, S. Nishida, L. Sharan, and E. H. Adelson. Image statistics for surface reflectance perception. *Nature*, 447(7141):206–209, 2007.
 - [52] M. Muja and D. G. Lowe. Fast approximate nearest neighbors with automatic algorithm configuration. In *Proceedings of the International Conference on Computer Vision Theory and Applications*, pages 331–340, Lisbon, Portugal, 2009.
 - [53] R. Ng, R. Ramamoorthi, and P. Hanrahan. Triple product wavelet integrals for all-frequency relighting. *ACM Transactions on Graphics*, 23(3):477–487, 2004.
 - [54] A. Ngan, F. Durand, and W. Matusik. Experimental validation of analytical brdf models. In *Proceedings of the International Conference and Exhibition on Computer Graphics and Interactive Techniques (SIGGRAPH)*, page 90, New York, NY, USA, 2004.
 - [55] P. Nillius, J. Sullivan, and A. A. Argyros. Shading models for illumination and reflectance invariant shape detectors. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Anchorage, AK, USA, 2008.
 - [56] J. S. Nimeroff, E. Simoncelli, and J. Dorsey. Efficient re-rendering of naturally illuminated environments. In *Proceedings of the Eurographics Workshop on Rendering*, pages 359–373, Darmstadt, Germany, 1994.
 - [57] K. Nishino and S. K. Nayar. Eyes for relighting. *ACM Transactions on Graphics*, 23(3):704–711, 2004.
 - [58] J. F. Norman, J. T. Todd, H. F. Norman, A. M. Clayton, and T. R. McBride. Visual discrimination of local surface structure: Slant, tilt, and curvedness. *Vision research*, 46(6-7):1057–1069, 2006.

-
- [59] M. Okabe, G. Zeng, Y. Matsushita, T. Igarashi, L. Quan, and H.-Y. Shum. Single-view relighting with normal map painting. In *Proceedings of the Pacific Conference on Computer Graphics and Applications (Pacific Graphics)*, pages 27–34, Taipei, Taiwan, 2006.
 - [60] J. Oliensis. New results in shape from shading. In *Proceedings of the DARPA Image Understanding Workshop*, pages 145–153, Pittsburgh, Pennsylvania, USA, Sept. 1990.
 - [61] P. Peers, N. Tamura, W. Matusik, and P. E. Debevec. Post-production facial performance relighting using reflectance transfer. *ACM Transactions on Graphics*, 26(3):52, 2007.
 - [62] A. P. Pentland. Local shading analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(2):170–187, 1984.
 - [63] B. T. Phong. *Illumination for Computer-Generated Images*. PhD thesis, University of Utah, 1973.
 - [64] F. H. Pighin, R. Szeliski, and D. Salesin. Resynthesizing facial animation through 3d model-based tracking. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 143–150, Kerkyra, Greece, 1999.
 - [65] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, 2nd edition, 1993.
 - [66] R. Ramamoorthi and P. Hanrahan. A signal-processing framework for inverse rendering. In *Proceedings of the International Conference and Exhibition on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pages 117–128, Los Angeles, CA, USA, 2001.
 - [67] R. Ramamoorthi and P. Hanrahan. Frequency space environment map rendering. *ACM Transactions on Graphics*, 21(3):517–526, 2002.
 - [68] T. Riklin-Raviv and A. Shashua. The quotient image: Class based recognition and synthesis under varying illumination conditions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 566–571, Fort Collins, CO, USA, 1999.
 - [69] P. Sen, B. Chen, G. Garg, S. R. Marschner, M. Horowitz, M. Levoy, and H. P. A. Lensch. Dual photography. *ACM Transactions on Graphics*, 24(3):745–755, 2005.
 - [70] A. Shashua and T. Riklin-Raviv. The quotient image: Class-based re-rendering and recognition with varying illuminations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(2):129–139, 2001.
 - [71] H. Shim and T. Chen. An example-based face relighting. In *Proceedings of the SPIE*, volume 8289, page 82890B, 2012.

- [72] C. Silpa-Anan and R. Hartley. Optimised kd-trees for fast image descriptor matching. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, Anchorage, AK, USA, 2008.
- [73] M. Stark, M. Goesele, and B. Schiele. Shading cues for object class detection. In *Proceedings of the IEEE Workshop on 3D Representation for Recognition (3dRR)*, pages 649–656, Kyoto, Japan, 2009.
- [74] M. Stark, M. Goesele, and B. Schiele. A shape-based object class model for knowledge transfer. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 373–380, Kyoto, Japan, 2009.
- [75] C. Tchou, J. Stumpfel, P. Einarsson, M. Fajardo, and P. Debevec. Unlighting the parthenon. In *Proceedings of the International Conference and Exhibition on Computer Graphics and Interactive Techniques (SIGGRAPH)*, page 80, New York, NY, USA, 2004.
- [76] S. Ullman. *The Interpretation of Visual Motion*. MIT press, 1979.
- [77] Z. Wen, Z. Liu, and T. S. Huang. Face relighting with radiance environment maps. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 158–165, Madison, WI, USA, 2003.
- [78] T.-T. Wong, P.-A. Heng, S.-H. Or, and W.-Y. Ng. Image-based rendering with controllable illumination. In *Proceedings of the Eurographics Workshop on Rendering*, pages 13–22, St. Etienne, France, 1997.
- [79] R. J. Woodham. Photometric stereo: A reflectance map technique for determining surface orientation from image intensity. In *Proceedings of the SPIE*, volume 155, pages 136–143, 1978.
- [80] Y. Yu and J. Malik. Recovering photometric properties of architectural scenes from photographs. In *Proceedings of the International Conference and Exhibition on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pages 207–217, Orlando, FL, USA, 1998.
- [81] L. Zhang, S. Wang, and D. Samaras. Face synthesis and recognition from a single image under arbitrary unknown lighting using a spherical harmonic basis morphable model. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 209–216, San Diego, CA, USA, 2005.
- [82] S. K. Zhou, G. Aggarwal, R. Chellappa, and D. W. Jacobs. Appearance characterization of linear lambertian objects, generalized photometric stereo, and illumination-invariant face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(2):230–245, 2007.