New 3D Scanning Techniques for Complex Scenes

Tongbo Chen

Max-Planck-Institut für Informatik
Saarbrücken, Germany

Dekan — Dean
Prof. Dr. Joachim Weickert,
Universität des Saarlandes, Saarbrücken, Germany

Vorsitzender des Prüfungsausschusses — Chairman of the Examination Board
Prof. Dr. Philipp Slusallek,
Universität des Saarlandes, Saarbrücken, Germany

Gutachter — Reviewers
Dr. Hendrik P. A. Lensch,
Max-Planck-Institut für Informatik, Saarbrücken, Germany

Prof. Dr. Michael Goesele,
Technische Universität Darmstadt, Darmstadt, Germany

Prof. Dr. Hans-Peter Seidel,
Max-Planck-Institut für Informatik, Saarbrücken, Germany

Promovierter akademischer Mitarbeiter —
Academic Member of the Faculty having a Doctorate
Dr. Robert Strzodka
Max-Planck-Institut für Informatik, Saarbrücken, Germany

Betreuender Hochschullehrer — Supervisors
Dr. Hendrik P. A. Lensch,
Max-Planck-Institut für Informatik, Saarbrücken, Germany

Prof. Dr. Hans-Peter Seidel,
Max-Planck-Institut für Informatik, Saarbrücken, Germany

Datum des Kolloquiums — Date of Defense
2. Dezember 2008 – December 2nd, 2008

Tongbo Chen
Max-Planck-Institut für Informatik
Campus E1.4
66123 Saarbrücken, Germany
tongbo@mpi-inf.mpg.de

Abstract

This thesis presents new 3D scanning methods for complex scenes, such as surfaces with fine-scale geometric details, translucent objects, low-albedo objects, glossy objects, scenes with interreflection, and discontinuous scenes.

Starting from the observation that specular reflection is a reliable visual cue for surface mesostructure perception, we propose a progressive acquisition system that captures a dense specularity field as the only information for mesostructure reconstruction. Our method can efficiently recover surfaces with fine-scale geometric details from complex real-world objects.

Translucent objects pose a difficult problem for traditional optical-based 3D scanning techniques. We analyze and compare two descattering methods, phase-shifting and polarization, and further present several phase-shifting and polarization based methods for high quality 3D scanning of translucent objects.

We introduce the concept of modulation based separation, where a high frequency signal is multiplied on top of another signal. The modulated signal inherits the separation properties of the high frequency signal and allows us to remove artifacts due to global illumination. This method can be used for efficient 3D scanning of scenes with significant subsurface scattering and interreflections.

Kurzfassung

Diese Dissertation präsentiert neuartige Verfahren für die 3D-Digitalisierung komplexer Szenen, wie z.B. Oberflächen mit sehr feinen Strukturen, durchscheinende Objekte, Gegenstände mit geringem Albedo, glänzende Objekte, Szenen mit Lichtinterreflektionen und unzusammenhängende Szenen.

Ausgehend von der Beobachtung, daß die spekulare Reflektion ein zuverlässiger, visueller Hinweis für die Mesostruktur einer Oberfläche ist, stellen wir ein progressives Meßsystem vor, um Spekularitätsfelder zu messen. Aus diesen Feldern kann anschließend die Mesostruktur rekonstruiert werden. Mit unserer Methode können Oberflächen mit sehr feinen Strukturen von komplexen, realen Objekten effizient aufgenommen werden.

Durchscheinende Objekte stellen ein großes Problem für traditionelle, optischbasierte 3D-Rekonstruktionsmethoden dar. Wir analysieren und vergleichen zwei verschiedene Methoden zum Eliminieren von Lichtstreuung (Descattering): Phasenverschiebung und Polarisation. Weiterhin präsentieren wir mehrere hochqualitative 3D-Rekonstruktionsmethoden für durchscheinende Objekte, die auf Phasenverschiebung und Polarisation basieren.

Außerdem führen wir das Konzept der modulationsbasierten Signaltrennung ein. Hierzu wird ein hochfrequentes Signal zu einem anderes Signal multi-

pliziert. Das so modulierte Signal erhält damit die separierenden Eigenschaften des hochfrequenten Signals. Dies erlaubt uns Meßartefakte aufgrund von globalen Beleuchtungseffekten zu vermeiden. Dieses Verfahren kann zum effizienten 3D-Scannen von Szenen mit durchscheinden Objekten und Interreflektionen benutzt werden.

Summary

This thesis presents new 3D scanning methods for complex scenes, such as surfaces with fine-scale geometric details, translucent objects, low-albedo objects, glossy objects, scenes with interreflection, and discontinuous scenes.

Starting from the observation that specular reflection is a reliable visual cue for surface mesostructure perception, we present a simple and robust method for surface mesostructure acquisition. In contrast to most photometric stereo methods, which take specularities as outliers and discard them, we propose a progressive acquisition system that captures a dense specularity field as the only information for mesostructure reconstruction. Our method can efficiently recover surfaces with fine-scale geometric details from complex real-world objects with a wide variety of reflection properties, including translucent, low albedo, and highly glossy objects. We show results for a variety of objects including human skin, dried apricot, orange, jelly candy, black leather and dark brown chocolate.

We present efficient and practical methods for 3D scanning of translucent objects. Translucent objects pose a difficult problem for traditional structured light 3D scanning techniques. Subsurface scattering corrupts the range estimation in two ways: by drastically reducing the signal-to-noise ratio and by shifting the intensity peak beneath the surface to a point which does not coincide with the point of incidence. In this thesis we analyze and compare two descattering methods in order to obtain reliable 3D coordinates for translucent objects. By using polarization difference imaging, subsurface scattering can be filtered out because multiple scattering randomizes the polarization direction of light while the surface reflectance partially keeps the polarization direction of the illumination. The descattered reflectance can be used for reliable 3D reconstruction using traditional optical 3D scanning techniques, such as structured light. Phase-shifting is another effective descattering technique if the frequency of the projected pattern is sufficiently high. We demonstrate the performance of these two techniques and the combination of them on scanning real-world translucent objects.

We introduce the concept of modulation based separation where a high frequency signal is multiplied on top of another signal. The modulated signal inherits the separation properties of the high frequency signal and allows for removing artifacts due to global illumination. This technique can be used to clean up arbitrary projected signals, e.g. photographs as well as the sinusoid patterns used for phase-shifting. For the modulated phase-shifting, we propose a two-pass separation method exploiting high frequency patterns in two-dimensions that can filter out the global components much more completely than traditional one-pass separation methods. We demonstrate the effectiveness of our approach on a couple of scenes with significant subsurface scattering and interreflections.

Acknowledgments

First of all, I would like to thank my advisers Dr. Hendrik P. A. Lensch, Prof. Dr. Michael Goesele, and Prof. Dr. Hans-Peter Seidel for their interest in this work, their valuable comments, and their continuous support. Prof. Dr. Hans-Peter Seidel is a great leader in computer graphics and more generally in the whole domain of computer science. He created and leads a very creative and productive group, where I fortunately have an opportunity to start my academic career and to pursue my Ph.D. degree.

Prof. Dr. Michael Goesele was my academic adviser during the first two years of my Ph.D. study. He introduced me into Max-Planck-Institut Informatik. He supervised my first academic project and paper at MPII on car paint acquisition. From him I learned a lot of skills about practical acquisition. I'm deeply in debt to Prof. Dr. Michael Goesele both academically and personally.

Dr. Hendrik P. A. Lensch was my academic adviser during the last two years of my Ph.D. study. I learned a lot of skills about how to successfully manage a project and how to nicely write an academic paper. He always encourages me to improve myself in communication skill and social life. He organized very dynamic regular Computer Graphics Lunch for the whole group and highly brain stimulating subgroup meetings. Dr. Hendrik P. A. Lensch himself is very active, positive, and creative and is always a good model for me to follow.

Special thanks to Christian Fuchs for his kindness, his collaboration, and his generosity of sharing knowledge. He is my first collaborator in Max-Planck-Institut Informatik, from whom I start to recognize the *German Style*. He taught me a lot of valuable lessons on how to do experiments, how to talk to people, how to give a talk, how to prepare a presentation etc. with great patience. He is always ready to become my *informal adviser*.

I would like to thanks all the members in the subgroup *General Appearance Acquisition* led by Dr. Hendrik P. A. Lensch. They are Martin Fuchs, Boris Ajdin, Miguel Granados, Matthias Hullin, and Andrei Linţu. This group is a wonderful environment for doing research.

Special thanks to Prof. Dr. Philipp Slusallek who chaired the Ph.D. defense. His humor and kindness will stay in my memory. Thanks to Dr. Robert Strzodka who was the Promovierter akademischer Mitarbeiter (Academic Member of the Faculty having a Doctorate) of the defense.

Special thanks to Martin Fuchs and Carsten Stoll, who are the group administrators and are always ready to help me solve system or hardware problem. Thanks to Sabine Budde, Conny Liegl and Sonja Lienard, who are the secretaries of our group for their generous, kind and patient helps. Thanks to the Helpdesk for their patient and skillful helps. Thanks to Prof. Dr. Holger Theisel and Prof. Dr. Volker Blanz for their kindness and help during my first years of the Ph.D. study.

Thanks to Michael Heinz for his help and his lessons on doing acquisition experiments. Thanks to Andreas Steinel for his valuable discussion and great work in the reflection security project from whom I learned lots of knowledge about astronomy imaging. Thanks to Markus Dürmuth and Prof. Dr. Michael Backes for their great collaboration in the reflection security project. Thanks to Johannes Güther and Dr. Ingo Wald for their great collaboration in the car paint project, which is my first project at MPII and resulted in my first paper.

I would especially like to thank all my present and former colleagues at the Computer Graphics Group at MPI who help to make it such a great place.

I would like to thank Prof. Dr. Steve Seitz, who together with Prof. Dr. Michael Goesele hosted me for two months at GRAIL at the University of Washington. I'm also very grateful to Prof. Dr. Brian Curless for his encouragements.

I would like to thank all the anonymous reviewers of our papers for their insightful and valuable comments and advices.

This work has been partially funded by the DFG Emmy Noether fellowship (Le 1341/1-1) and the Max Planck Center for Visual Computing and Communication (BMBF-FKZ01IMC01).

I would also like to thank my good friends in Saarbrücken who made my social life wonderful. To name some of them, Kuangyu Shi, Zhao Dong, Hanglin Pan, Hongbo Zhu, Danyi Wang, Wenxiang Ying, Fei Guo, Zhiguo Xing, Hui Jiang and so on.

Special thanks to my parents, who always encourage and support me in my life. I'm very grateful to my wife, Yan Wang, for all the support and help. She is always beside me under any situation.

Contents

List of Figures

List of Tables

Chapter 1
Introduction

1.1   Problem Statement

3D scanning of real-world objects or scenes has been a popular topic in computer graphics and computer vision for a long time and important progress has been made since the early days. Nevertheless, there are still some well known open problems.

Generally, 3D scanning techniques can be categorized in terms of being active or passive, or being contact or non-contact. In this thesis, we will focus on active non-contact 3D scanning methods. One of the most popular active non-contact 3D scanning methods is structured light scanning, which is widely employed and tested in research or industry projects, such as the Digital Michelangelo Project [Levoy00] and the Michelangelo's Florentine Pietà project [Bernardini02]. Some following structured light methods improve the performance by incorporating both spatial and temporal coherence [Curless95, Zhang03, Zhang04, Davis05, Zhang05] or by integrating position and normal information [Nehab05]. Most of these methods work on opaque objects and assume the light emitted from the light source will hit the surface once and directly reflect to the sensor. In real world, however, this assumption will often be broken. For example, most sculptures made from marble or alabaster exhibit strong subsurface scattering, which causes serious problems for laser based scanning where the detected peak will be shifted [Godin01, Curless95]. Translucent objects are very common in the real world, including most fruits, animal or human skin, marble, alabaster, jade, etc. It is a challenging task to acquire high quality and high resolution 3D models of such kind of objects. Subsurface scattering is only one of the problems posed for optical based 3D scanning methods. In addition other global light transport effects inside an object or a scene make 3D scanning more difficult, such as volumetric scattering, refraction and interreflection. Even for objects with only direct

reflection, such as specular or low-albedo objects, accurate 3D scanning becomes a hard problem. Recently the scanning of optically challenging objects/scenes attracts more and more interests from both computer graphics and computer vision communities [Kutulakos05, Hasinoff06, Morris07a, Morris07b, Hullin08, Narasimhan05, Magda01, Miyazaki03a, Miyazaki05, Tarini05, Chen06, Ma07, Chen07, Chen08, Hertzmann05].

1.2 Contributions

In this thesis, we contribute to the 3D scanning problem of complex objects/scenes in three different aspects: mesostructure reconstruction for complex objects, such as translucent, glossy or low-albedo objects; 3D scanning of translucent objects using phase-shifting and polarization; and a novel modulated phase-shifting method.

1.2.1 Mesostructure from Specularity

Surface mesostructure represents geometric details that are relatively small but still individually visible such as bumps or dents on a surface.

By drawing inspiration from photographs of real-world translucent objects and from the literature on human vision and perception [Fleming04, Fleming05, Norman04, Todd04a], we found that specular highlights are an important visual cue for surface mesostructure perception and reliable visual information for surface detail representation. In Figure 1.1, the first row shows four images of a piece of orange skin under changing illumination. The small bumps on the orange skin introduce rich visual effects and can be efficiently revealed by specular highlights. Based on this observation, we developed a simple and progressive system that uses specular highlights in order to solve the dense mesostructure reconstruction problem for a variety of real-world complex objects, which possess a significant specular reflection component. Our method is largely independent of the underlying reflectance model, and can therefore successfully handle objects with complex reflectance that have previously been challenging. Efficient mesostructure reconstruction methods can contribute greatly to high-quality graphics models in terms of fine-scale surface geometric details. An accurate and explicit mesostructure model can also benefit related mesostructure modeling techniques such as BTFs (Bidirectional Texture Functions) [Dana99, Müller05].

Our contributions to the mesostructure reconstruction problem include [Chen06]:

- We simplify the problem of mesostructure reconstruction from complex objects, e.g., objects with translucency, which has up to now been expensive
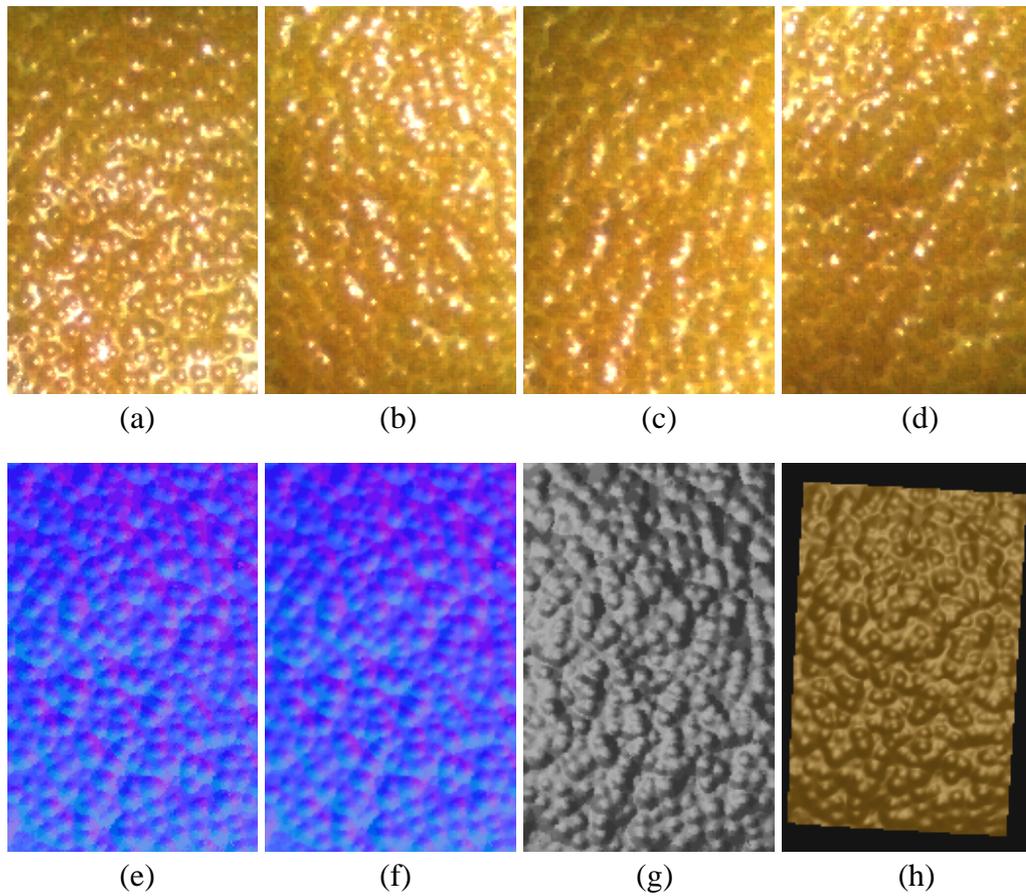
Figure 1.1: *Mesostructure reconstruction of a piece of orange skin. (a-d) Four cropped input images. (e) Recovered normal field (RGB-encoded). (f) Filtered normal field [Tomasi98]. (g) Rendering of the normal field using Ward's isotropic BRDF model [Larson92]. (h) Reconstructed 3D surface rendered at a novel viewpoint.*

or even impossible to solve.

- We use a dense specularity field as the only reliable visual information for mesostructure reconstruction.

- We develop a simple incremental and very flexible acquisition system.

- We acquire high-quality mesostructure, for a variety of real-world objects including human skin, dried apricot, orange skin, jelly candy, black leather, and dark brown chocolate.

| (a) | (b) | (c) |
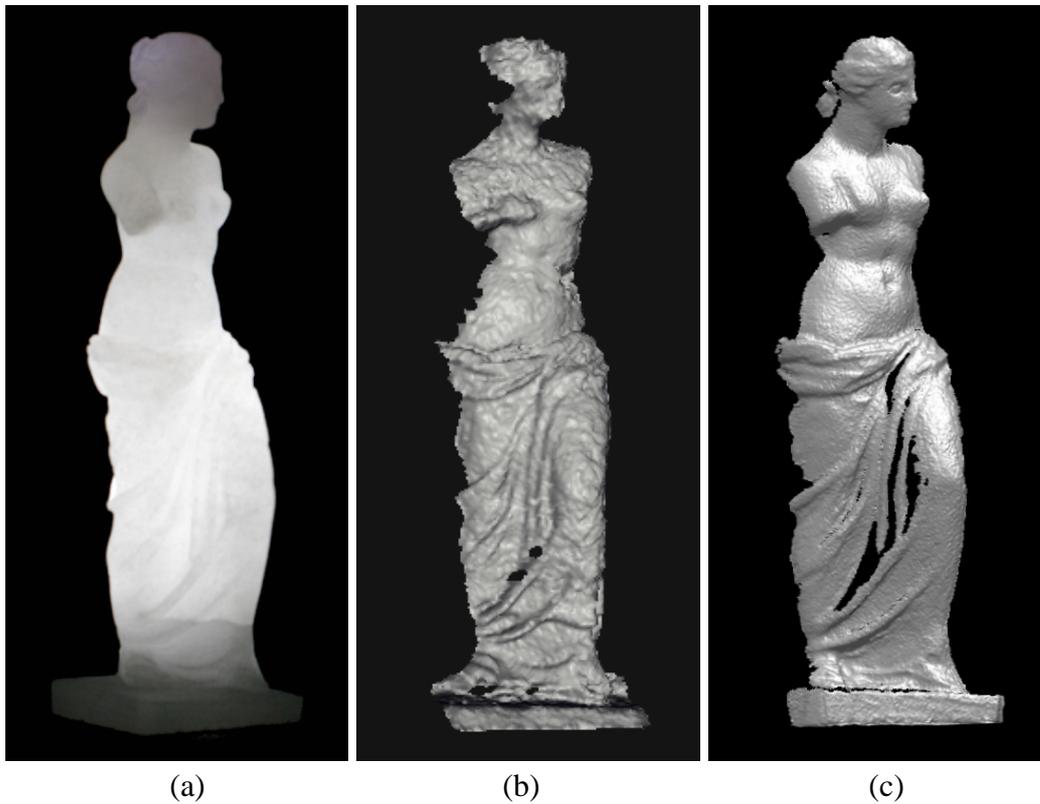
Figure 1.2*: 3D Scanning of a translucent object. (a) is a photograph of an highly translucent alabaster Venus figurine (height ≈ 19cm). (b) is a 3D scan from Minolta VI-910 laser range scanner. (c) By combining phase-shifting and polarization our method faithfully captures the 3D geometry.*

1.2.2   3D Scanning of Translucent Objects

For a number of scenes, structured light 3D scanning techniques run into the problem that the signal observed by the camera for a surface point is actually not only due to direct reflection of the projected pattern but instead contains polluting signals originating from ambient illumination, interreflections from other scene parts, or from subsurface scattering.

These effects are most prominent in translucent objects where the directly reflected signal is furthermore weakened since the incident light is diffused inside the material instead of being fully reflected at the surface.

Subsurface scattering can of course be excluded completely if the object's surface is painted before scanning, as it is done frequently. In this thesis, we propose 3D scanning techniques which are inherently robust against subsurface

scattering.

In order to obtain reliable scans of translucent objects one has to separate the direct reflection from the pollution due to multiple interreflections or scattering. One approach to descattering is to use the fact that light scattered multiple times gets unpolarized. Projecting polarized light and computing the difference of images captured with a polarization filter at two orthogonal orientations thus removes most of the multiple scattering contribution [Wolff94, Rowe95, Schechner05, Treibitz06]. Another method for separating direct from global reflections based on high frequency illumination patterns has recently been proposed by Nayar et al. [Nayar06]. In the same paper Nayar et al. also mention that phase-shifting [Srinivasan85, Zhang06e] can perform the separation and 3D scanning at the same time.

Inspired by the previous work, we propose a new 3D scanning method by combining phase-shifting and polarization. Figure 1.2 show a highly translucent alabaster Venus figurine, which is very difficult to be scanned with traditional optical scanning methods, can be successfully scanned using our method.

Our contributions to the 3D scanning of translucent objects include [Chen07]:

- We analyze the descattering properties of phase-shifting and polarization.

- We propose efficient and practical methods for 3D scanning of translucent objects, which were previously very difficult or even impossible. We demonstrate and assess our methods on a variety of translucent objects, such as fruits, alabaster sculptures, etc.

1.2.3 Modulated Phase-Shifting

The feasible frequency of traditional phase-shifting method is limited by the phase unwrapping method and the optics and resolution of both the camera and the projector and is in one dimension. The limitations can hurt the descattering capability, which is mostly related to the frequency of the projected signal, of the traditional phase-shifting method.

In order to alleviate this problem, we introduce the modulated phase-shifting. Figure 1.3 demonstrates the improvements of modulated phase-shifting over traditional phase-shifting and polarization difference imaging. Our contributions include [Chen08]:

- We introduced the modulation based phase-shifting method, which greatly improves the descattering power of traditional phase-shifting method and achieves higher performance in real 3D scanning of scenes with strong global light transport, such as subsurface scattering and interreflection, than previous methods.
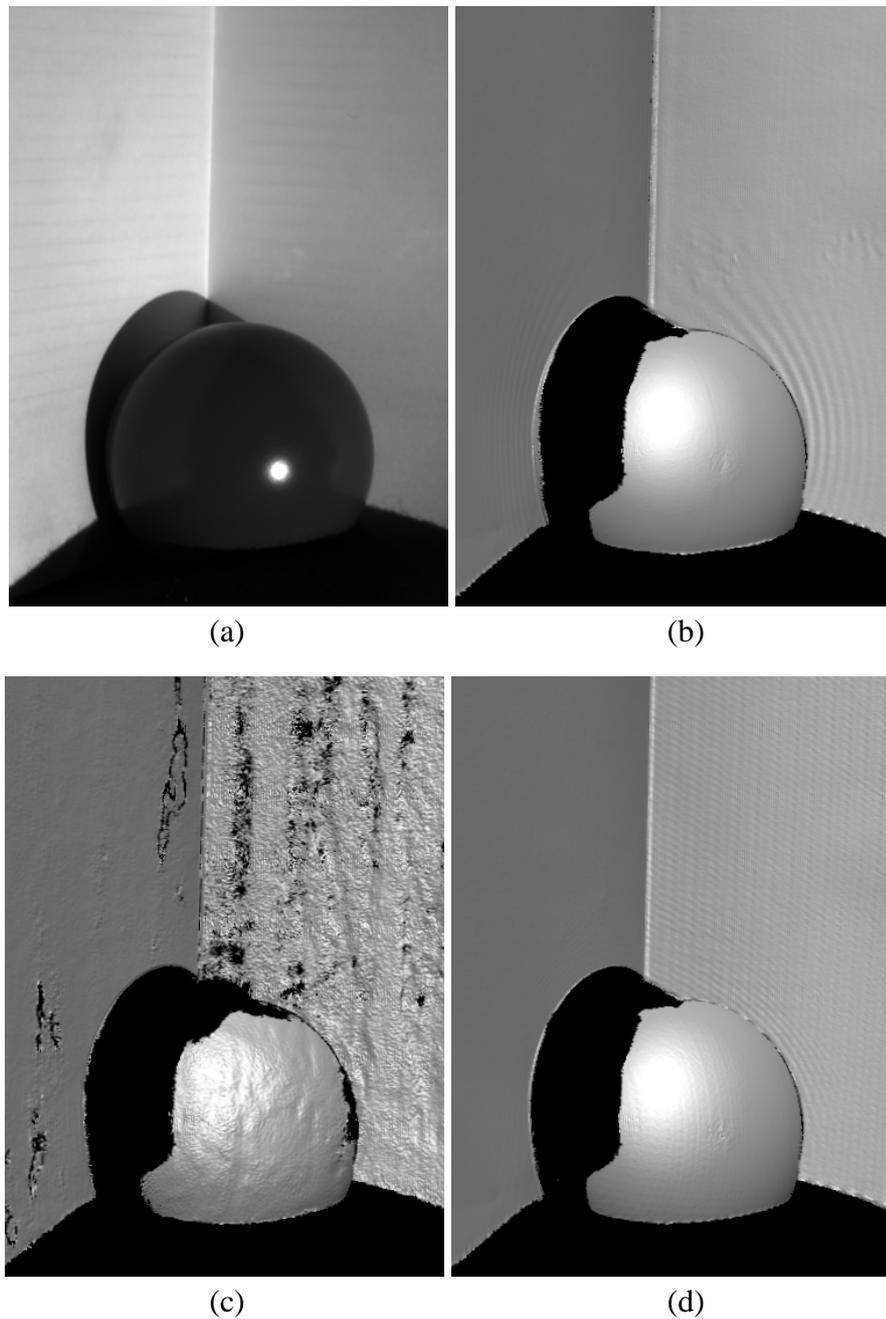
Figure 1.3: *A scene with strong interreflections caused by a glossy sphere between the two pages of a book. (a) Photograph of the scene. (b) Phase-shifting produces ripples in the final geometry due to the reflections of the sphere. (c) The result by polarization difference imaging is very noisy due to the modified polarization after interreflections. (d) The most accurate reconstruction is achieved using modulated phase-shifting.*

- We further show that due to the non-linearity in the separation analysis, a sequential analysis of a 2D pattern once horizontally and once vertically produces an even better separation.

- Moreover, the modulation and separation scheme can be exploited for general low frequency patterns, such as an ordinary photograph.

1.3 Thesis Overview

The remainder of this dissertation is structured as follows: In Chapter 2 we will briefly review the related work. Chapter 3 summarizes the fundamentals of light transport. Most of the optically challenging objects for 3D scanning can be analyzed under the light transport framework. In the following chapters, we will solve some of the open problems. In Chapter 4, we present a novel mesostructure acquisition method by exploiting specularities. In Chapter 5, we propose efficient and practical methods for 3D scanning of translucent objects. We further improve the performance of phase-shifting method by introducing the modulated phase-shifting method in Chapter 6. We conclude the thesis in Chapter 7 and provide an outlook on future work.

Chapter 2
Related Work

In this section, we provide a short introduction to 3D scanning technologies and briefly review the related work, including 3D shape perception, surface mesostructure acquisition and rendering, 3D reconstruction methods, photometric stereo, phase-shifting for 3D scanning, separation of reflection components, and polarization based imaging techniques.

2.1   3D Scanning Technologies

3D models are either created by artists or captured from real world. 3D scanning technologies are the enabling technologies to help people get digital 3D models from the real world. They may find applications in dissemination of museum artifacts, special visual effects production, games, virtual worlds, reverse engineering, design, product prototyping, medicine, web commerce and entertainment industry.

3D scanning technologies divide roughly into to contact and non-contact. The contact methods includes CMM (Coordinate Measurement Machine), jointed arms, slicing etc. The non-contact methods includes optical-based, microwave-based, radar-based, and sonar-based methods. In this thesis we will focus on optical-based methods, which can be roughly classified as active or passive. Examples of passive methods include stereo, multi-view stereo, shape from shading, photometric stereo, shape from silhouettes, depth from focus/defocus etc. The active methods include laser scanner, structured light, time-of-flight, interferometry, tomography, active depth from defocus etc.

An optimal 3D scanning technique usually includes the following characteristics:

- accurate

- fast

- high resolution

- robust for different kind of materials

- stable under various environments

- easy to use and to move

- safe and non-destructive

- capable of capturing object appearance

- low price

However, in practice, there are no single 3D scanning technology has all the above characteristics. Usually an individual scanning method is specialized for specific kind of applications. For example, a high resolution method usually works slower than a low resolution method. The best choice of scanning technology is usually made with compromise and depending on application.

2.2   3D Shape Perception via Specular Reflections

Research on the nature of 3D shape perception has been an important topic for more than two millennia, and remains an active area involving many different disciplines, including psychology, physics, neuroscience, computer science, and mathematics. There are many different types of optical stimulation that provide perceptually salient information about 3D shape, including variations of shading, texture, contours, occlusion, binocular disparity, motion parallax, and optical deformations of a dynamic scene. To keep simplicity, we only review the work on shape perception relying on specular reflections. Many materials, including skin, leather, jade, marble, glass, water, leaves, plastic, steel, ceramic products, oil painting etc., exhibit specular reflections. Previous studies have shown that specular reflections aid shape estimation or even provide reliable and accurate constraints on 3D shape [Blake90, Blake91, Todd04a, Todd04b, Fleming04, Norman04]. In [Weidenbacher06], Weidenbacher et al. proposed a biologically motivated recurrent model for the extraction of visual features relevant for the perception of 3D shape information from images of a mirrored objects and utilized the model output to create a rough nonphotorealistic sketch representation.

Inspired by the work on human vision and perception, some research formalize the problem by a computational framework. In contrast to most of the photometric stereo methods, where specular highlights are detected and separated as outliers, shape-from-specularity methods try to efficiently use the specular reflectance component.

Ikeuchi [Ikeuchi81] analyzed photometric stereo for specular surfaces. Sanderson et al. [Sanderson88] developed a structured light system, SHINY to recover surface depth and orientation using both single and multiple cameras. Zheng and Murata [Zheng00] presented a system in which a rotating specular object was illuminated by an extended circular light sources and reconstructed 3D shape by tracing the specularity trace or using motion stereo. Zisserman et al. [Zisserman89] provided a quantitative analysis of the information available to a camera undergoing known motion. One key result was that the convex/concave ambiguity can be resolved under unknown illumination. Oren and Nayar [Oren96] introduced a comprehensive theoretical framework for the perception of specular surface geometry and developed an algorithm that uniquely recovers 3D surface profiles using a single virtual feature tracked from the occluding boundary of the object. Savarese and Perona [Savarese01, Savarese02] provided a general solution for recovering shape from mirror reflections in a single static image. Solem et al. [Solem04] introduced variational analysis into shape-from-specularity and demonstrated the robustness of shape reconstruction from sparse specularity data. Lu et al. [Lu00] used specular reflections on surfaces to help modeling tangential hairs or grooves.

Francken et al. [Francken08] proposed a mesostructure from specularities method using LCD monitor as the illumination. Ma et al. [Ma07] developed a rapid acquisition method of specular and diffuse normal maps from polarized spherical gradient illumination, where specular reflection is used for the reconstruction of fine-scale geometry of translucent objects. Adato et al. [Adato07] developed a framework for recovering general shape from distorted specular reflection under unknown environment. Nehab et al. [Nehab08] presented a dense 3D reconstruction method for glossy objects using specularity consistency.

2.3   Surface Mesostructure

Surface mesostructure is one of the key components of 3D texture [Koenderink96]. It contributes strongly to the complex surface appearance of real-world objects. One method for modeling and rendering mesostructure is through BTFs (Bidirectional Texture Functions) [Dana99], which can be regarded as a mapping from the 4D space of lighting and viewing directions to the space of 2D images. Most previous work on BTFs aims at capturing appearance data

from natural materials and at efficient representation. Müller et al. [Müller05] gives a comprehensive report on the state of the art of BTFs techniques. Liu et al. [Liu01] used a shape-from-shading method to recover approximate 3D geometry of surface details from a BTF dataset. In [Neubeck05], Neubeck et al. proposed a method for 3D texture reconstruction from extensive BTF data, with only a few and rather weak assumptions about reflectance and geometry. The reconstructed mesostructure can be used for the simplification of the BTF-based texture description and efficient compression of a BTF dataset. Even for the most advanced and expensive laser scanning systems, mesostructure reconstruction of highly specular or translucent objects is still a difficult problem. Most of the scanning technologies based on structured lighting will also fail in reconstructing fine-scale details for very low albedo, translucent, or highly-specular surfaces [Nehab05, Davis05]. To deal with highly-specular surfaces, Wang and Dana [Wang03, Wang06] presented a method that can simultaneously capture fine-scale surface shape and spatially varying BTFs by using a BTF measurement system. Similar to that work, our method will also depend on specular reflection. But we extend the idea to include not only highly-specular surfaces, but also very low albedo glossy or translucent glossy materials. Instead of using a complicated BTF measurement system, we developed a simple, flexible and progressive acquisition system. In [Yu05], Yu and Chang introduced shadow graphs for 3D texture reconstruction. They show that the shadow graph alone is sufficient to solve the shape-from-shadow problem from a dense set of images. They also solved the problem of recovering height fields from a sparse set of images by integrating shadow and shading constraints. However, this method cannot work effectively for objects where shadow is no longer an accurate information, such as skin or fruit.

2.4   3D Reconstruction

Numerous 3D scanning techniques [Besl88, Poussart88, Curless97] have been developed during the last decades. A long processing pipeline is necessary to obtain a complete 3D model from a collection of range scans [Levoy00, Bernardini02]. In this thesis we concentrate just on capturing reliable range maps and do not cover further processing such as registration, merging, or smoothing. Structured light methods (see Salvi et al. [Salvi04] for a survey) analyze a set of images captured under well defined patterns in order to determine the correspondence between camera and projector pixels for each surface point, from which the point's depth can be computed. They range from line sweeping algorithms [Curless95] to optimized stripe boundary codes that allow for real-time scanning [Hall-Holt01]. Rusinkiewicz et al. [Rusinkiewicz02] proposed a real-time 3D model acquisition

system that permits the user to rotate an object by hand and see a continuously-updated model as the object is scanned. Davis et al. [Davis05] presented a common framework, space-time stereo to unify stereo, structured light, and laser scanning. Nehab et al. [Nehab05] introduced an efficient algorithm for combining positions, usually from stereo triangulation and normals, usually from photometric stereo, to obtain highly precise 3D geometry.

Recently, people saw more and more specialized methods for 3D reconstruction of complex objects, including transparent, translucent, specular, inhomogeneous, or geometrically intricate objects. Kutulakos and Steger [Kutulakos05] proposed an effective method for 3D shape reconstruction of refractive and specular objects by light-path triangulation. Miyazaki et al. [Miyazaki03a, Miyazaki05] reconstructed the surface shape of transparent objects using polarization. In [Hasinoff06], Hasinoff and Kutulakos presented confocal stereo for computing 3D shape by controlling the focus and aperture of a lens. The method is specially designed for reconstructing scenes with high geometric complexity or fine-scale texture. To achieve this, they introduced the confocal constancy property. Wei et al. [Wei05] proposed an image-based approach to model hair geometry from images taken from multiple viewpoints. Yamazaki et al. [Yamazaki07] presented a practical shape-from-silhouettes method using coplanar shadowgram imaging to acquire 3D models of intricate objects, such as tree branches, bicycle and insects. In [Morris07b] Morris and Kutulakos presented a method for reconstructing the exterior surface of a complex transparent scene with inhomogeneous interior (e.g. multiple interfaces, reflective or painted interiors, etc.). Trifnov et al. [Trifonov06] provided a visible light tomographic reconstruction method for transparent objects.

2.5 Photometric Stereo

Photometric stereo methods [Woodham80, Rushmeier97] are known to be able to capture fine-scale surface details and to provide an efficient alternative to BTF-based methods. In [Hertzmann05], Hertzmann and Seitz presented an example-based photometric stereo for shape reconstruction with general spatially varying BRDFs. They assumed that there are no cast shadows, no interreflections, and no subsurface scattering. Goldman et al. [Goldman05] proposed a photometric stereo method for iteratively recovering shape and BRDFs. They employed a local reflectance model, which cannot properly account for shadows, interreflections and subsurface scattering. In [Paterson05], Paterson et al. developed a simple system for BRDF and geometry capturing. Their system can handle a variety of real-world objects except highly specular or translucent materials. Wu and Tang [Wu05] presented a simple dense photometric stereo method, using only a

mirror sphere, a spotlight and a DV camera. They achieved surprisingly good results even with the presence of moderate shadows and specular highlights. To our knowledge, photometric stereo methods can rarely recover dense fine-scale surface details from translucent, highly specular, or low albedo glossy materials.

In [Magda01], Magda and Zickler take advantage of Helmholtz reciprocity and light fields to reconstruct surfaces with arbitrary BRDFs. That method makes no assumption of the surface BRDF and works effectively for a variety of non-Lambertian surfaces (e.g. glossy surface), but not for highly translucent objects, where subsurface scattering dominates. Alldrin et al. [Alldrin08] presented a photometric stereo method for simultaneously recovering shape and spatially varying reflectance of a surface. In their method, novel bi-variate approximations of isotropic reflectance functions were employed.

2.6   Phase-Shifting

While most structured light approaches simply assume to observe the undistorted measurements from direct surface reflections the phase-shifting algorithm is more robust against noise caused by global illumination effects such as subsurface scattering. Phase-shifting [Srinivasan85, Wust91] based on sinusoid patterns has been used extensively in optical 3D scanning. The source, usually a digital projector, projects a series of phase-shifted sinusoid patterns into the scene and a camera records the resulting image. From a sequence of shifted patterns one can detect for every camera pixel the phase within one period of the illuminating sinusoid pattern. The collective phase information is called the modulo 2π phase map. In order to determine the absolute position within the illumination pattern the period has to be localized. This process is called phase-unwrapping, for which lower frequency sinusoid patterns are typically used [Ghiglia98]. An unwrapped phase map directly encodes the correspondence between the image field and the projection field. Once this correspondence is determined, the 3D coordinate information of the object can be computed by triangulation.

Most of the advanced phased-shifting methods focus on the different aspects of the above procedure. With the increasingly ubiquitous availability of digital projectors and digital cameras, a typical setup of a phase-shifting system is a projector-camera system. However, the inaccuracy inherent in the commercial projectors introduces new problems. Zhang and Yau [Zhang06d] proposed a new look-up-table method for phase error compensation. Instead of trying to improve scanning accuracy, some researchers are continuously working on the scanning speed. Zhang and Yau [Zhang06e] presented a high-resolution, real-time phase-shifting method with customized hardware. Weise et al. [Weise07] developed a fast 3D scanning system combining stereo and phase-shifting. Most of

the high quality or high speed phase-shifting methods require an efficient phase-unwrapping method [Ghiglia98]. Zhang et al. [Zhang06c] presented a multi-level quality guided phase-unwrapping for a real-time phase-shifting system. Huntley and Saldner [Huntley93] introduced the temporal phase-unwrapping method, which is a multi-scale method and very robust, especially when scanning highly discontinuous surfaces.

2.7 Separation of Reflection Components

Using images captured with a polarization filter at different orientations one can for example separate diffuse from specular reflections [Nayar97, Rowe95, Tyo96, Umeyama04] or attempt to remove depolarized global effects such as multiple scattering due to participating media [Schechner04, Schechner05, Schechner03, Treibitz06]. Making use of structured, high frequency illumination, most global effects can be removed since only direct reflection will propagate high frequencies while global effects drastically damp them [Narasimhan05, Nayar06]. Wu and Tang [Wu04] obtained a full separation into specular, diffuse, and subsurface scattering reflection components by additionally analyzing directional dependence.

The direct reflection component is due to a single reflection at the surface and is therefore directly correlated to the surface geometry. The global component is caused by multiple scattering events, such as interreflection or subsurface scattering. The global components can seriously deteriorate the measurements in 3D scanning [Godin01]. Accurate 3D scanning methods for scenes with complex global light transport usually have a key component of removing or suppressing the global components. In [Seitz05], Seitz et al. presented a method for computing and removing interreflection in photographs of real scenes. Their method is based on the Lambertian assumption and requires a very large number of images to estimate the photometric coupling between all pairs of scene points. Ma et al. [Ma07] presented a rapid acquisition method of specular and diffuse normal maps from polarized spherical gradient illumination. Their method can produce high resolution 3D scans for moderately translucent objects, such as human faces. Based on the insight that direct and indirect scatter traces have different characteristics, Morris and Kutulakos [Morris07b] introduced a new 3D photography method, called scatter-trace photography. Their method provides new possibilities for 3D scanning of inhomogeneous transparent scenes. In [Nayar06] Nayar et al. introduced an efficient method for separating direct and global components by using high frequency illumination. Their approach is based on the insight that global light transport significantly dampens high frequencies in the incident illumination patterns while the direct reflections does not. In the same paper, they proposed several useful high frequency patterns, such as checkerboard patterns,

sine patterns, etc. They also pointed out that phase-shifting can be used for 3D scanning. In [Talvala07], Talvala et al. adapted Nayar et al.'s separation method to remove veiling glare in high dynamic range imaging. They use a high frequency mask to selectively block the light that contributes to the veiling glare. Veeraraghavan et al. [Veeraraghavan07] presented a novel modulation-based method for capturing light field by using an attenuating mask.

2.8   Polarization

Polarization imaging has been widely used in computer vision, metrology, and optics. We refer to the book [Born99] for comprehensive understanding of polarization and its applications. We summarize some typical applications of polarization in computer vision, including reflectance separation, material classification, visibility improvements, 3D reconstruction etc.

Imaging in scattering media such as water and fog is usually suffering from poor visibility due to backscattering and signal attenuation. Polarization can be used to improve visibility under these situations. Some methods assumed a negligible degree of polarization of the objects [Schechner04, Schechner05, Schechner03, Treibitz06]. Other methods assumed the contrary, i.e., that object reflection is significantly polarized, rather than the backscatter [Tyo96]. In [Treibitz08], Treibitz presented a unified method for active polarization descattering which allow both the backscatter and the object reflection to be partially polarized. Based on imaging through a polarizer at two or more orientations, Schechner et al. [Schechner00] presented an approach to recover scenes deteriorated by reflections off a semi-reflecting medium (e.g., a glass window).

Shape-from-polarization is known to be effective to estimate the shape of specular objects such as metals or transparent objects. The degree of polarization of the light reflected from the object surface depends on the reflection angle, which, in turn, depends on the object's surface normal. Wolff et al. [Wolff90, Wolff91] measured the orientation of glass and metal plane by combining polarization analysis and binocular stereo. Clark et al. [Clark97] and Wallace et al. [Wallace99] improved the laser range finder by polarization analysis to estimate the shape of opaque objects. Saito et al. [Saito99] attempted to measure the surface shape of transparent objects; unfortunately, there was an ambiguity problem for determining the true surface normal. Miyazaki et al. [Miyazaki02] extended Saito's method and solved the ambiguity problem from a polarization analysis of thermal radiation. Rahmann and Canterakis [Rahmann01] applied polarization imaging to the reconstruction problem of specular surface. The projection of surface normals is directly provided the polarization image. Morel et al. [Morel06] extended the shape from polarization method to specular metallic surfaces. Recently,

Miyazaki et al. [Miyazaki03b] proposed a method to estimate the surface shape of an opaque specular object from a polarization image from a single view. Miyazaki et al. [Miyazaki03a] further presented a method to recover surface shape of transparent objects from two views. Miyazaki and Ikeuchi [Miyazaki05] introduced inverse polarization raytracing, which incorporates both the path of light and its polarization state, for estimating surface shapes of transparent objects.

2.9   Discussion

Inspired by the previous work on 3D perception, 3D reconstruction, reflectance separation, and polarization, we introduce several novel methods for 3D scanning of complex scenes, including scenes with fine geometric details, subsurface scattering and interreflection.

For human vision, specularity is one of the key visual cues for 3D shape perception of fine geometric details, namely mesostructures. For example, wrinkles on glossy skin are much more obvious on diffuse skin. Wrinkles under directional illumination are more visible than under uniform environment illumination. Based on this perceptual observation, we introduce an interactive photometric method *mesostructure from specularity*, which can do successful mesostructure reconstruction even for low-albedo, or translucent objects, such as leather, skin etc. The system is easy to implement, including one camera and one point light source. The quality of scanning can be interactively and progressively improved by denser sampling of the light source position. A densely spaced light source environment, such as a light stage, could be used for fast and high quality scanning.

Accurate 3D scanning of scenes with global light transportation such as subsurface scattering and interreflection is a challenging task. For most of the optical-based 3D scanning methods, subsurface scattering inside the translucent objects, such as marble, alabaster, jade etc. will severely spoil the scanning accuracy. Nayar et al. [Nayar06] pointed out that high frequency illumination can be used to separate direct and global reflectance components. Their method provides new possibility of efficiently separating the direct surface reflectance from the global reflectance including subsurface scattering, interreflection etc. On the other hand phase-shifting is a very robust and efficient 3D scanning method in the optical engineering community. We combine these two powerful methods and introduce the phase-shifting based 3D scanning method for scenes with global light transportation, such as subsurface scattering and interreflection. With the increasing of frequency, the phase-shifting patterns become more and more effective in descattering performance. Combining with temporal phase unwrapping method, it is possible to scan a highly discontinuous surface. However the frequency of the phase-shifting pattern, the sine pattern, is bounded in traditional phase-shifting

method, which limits the descattering power and further limits the scanning accuracy. The key of successful scanning of scenes with global light transportation is to do successful descattering, i.e. separate the most interesting direct reflection from the rest. In computer vision, polarization phenomena is quite well understood and exploited to do descattering and dehazing for clear vision. We further introduced two methods to improve the performance of pure phase-shifting. First we incorporate polarization to enhance the descattering performance of high frequency illumination. The improvement is most visible when dealing with surfaces with moderate variation of normals. Second we find out the phase-shifting pattern can be further modulated and be in much higher frequency, which leads to higher performance in descattering. The modulation scheme can also be applied to general low frequency signal.

Chapter 3
Light Transport

In this chapter we will review the reflectance models and light transport. We will discuss the relevant effects of direct reflection, multiple scattering or inter-reflections on projected polarized or non-polarized structured light patterns. The signal we are most interested in for 3D scanning is the directly reflected light, which is closely correlated to the surface geometry. All the scanning methods presented in this thesis will focus on how to separate efficiently the direct reflection and how to retrieve accurate 3D information from the direct reflection. For example, the mesostructure from specularity method exploits the specular reflection, which is one component of direct reflections, to reconstruct fine-scale geometry, the mesostructure from complex objects even with global light transport. The phase-shifting and polarization based methods combine both the descattering properties of high frequency illumination and polarization to select the direct reflection for 3D reconstruction.

3.1   Reflectance Models

Light reflectance models have always been of great interest to the computer graphics and computer vision communities. Reflection of light from surfaces can be classified into two broad categories: diffuse and specular. The diffuse component is distributed in a wide range of directions around the surface normal, giving the surface a matte appearance. In the ideal case of Lambertian surfaces, the diffuse reflection does not change with direction at all. The specular reflection, on the other hand is strictly related to the light incident and exitant angles and is concentrated in a compact lobe around the specular direction. The bidirectional reflectance distribution function (BRDF) [Nicodemus77] is the general model to describe surface reflectance. It is a function of the wavelength, surface roughness properties, and the incoming and outgoing directions and is able to correctly

(a) BRDF   (b) BSSRDF
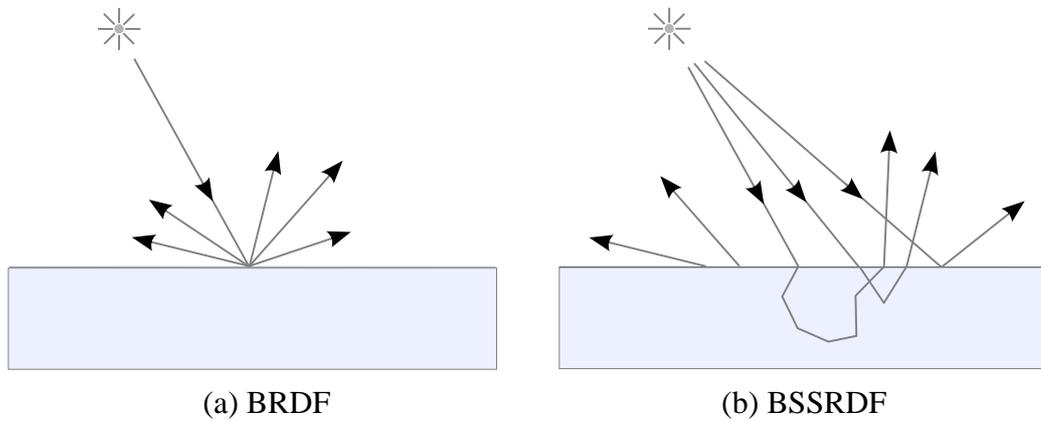
Figure 3.1*: BRDF vs. BSSRDF. A BSSRDF describes light transport between any two rays that enter or leave a surface. It can successfully modeling single/multiple scattering of translucent material. A BRDF is an approximation of the BSSRDF for which it is assumed that light enters and leaves the surface at the same point. It is a function of the wavelength, surface roughness properties, and the incoming and outgoing directions and is able to correctly predict the diffuse and specular components of the reflected light.*

predict the diffuse and specular components of the reflected light.

The BRDF, however, assumes that light entering a surface leaves the surface at the same position (see Figure 3.1(a)). This approximation is valid for most of the the opaque objects, such as metals, but it fails for translucent materials, which exhibit significant subsurface scattering. Translucent materials are ubiquitous in the real world, such as leaves, fruits, marble, jade, milk, cheese, skin, alabaster and so on. A more general reflectance model, called bidirectional surface scattering reflectance distribution function (BSSRDF) [Nicodemus77, Ishimaru78, Jensen01] can describe light transport between any two rays that enter or leave a surface (see Figure 3.1(b)). A BSSRDF, S, relates the outgoing radiance, $L_o(x_o, \overrightarrow{\omega_o})$ at the point x_o in direction $\overrightarrow{\omega_o}$, to the incident flux, $\Phi_i(x_i, \overrightarrow{\omega_i})$ at the point x_i from direction $\overrightarrow{\omega_i}$:

$$dL_o(x_o, \overrightarrow{\omega_o}) = S(x_i, \overrightarrow{\omega_i}, x_o, \overrightarrow{\omega_o})d\Phi_i(x_i, \overrightarrow{\omega_i}). \tag{3.1}$$

The BRDF is an approximation of the BSSRDF for which it is assumed that light enters and leaves the surface at the same point, i.e. $x_o = x_i$.

3.2   Light Transport

Nayar et al. [Nayar06] presented a fast separation method for direct and global components of a scene using high frequency illumination. We will follow the definitions of direct and global components in [Nayar06]. When a scene is lit by a single light source and viewed by a single camera, the radiance of each point in the scene can be broadly classified as two components, namely, direct and global. The direct component is due to the illumination of the point by the source. The global component is due to the illumination of the point by other points in the scene. We call the incident radiance upon the camera from direct component the *direct reflection*, such as specular reflection. Any other reflection different from direct reflection is called *global reflection*, such as subsurface scattering and interreflection.

Figure 3.2 shows schematically the behavior of direct and global illumination components. The scene is lit by a single light source and viewed by a single camera. Consider the scene point P. The light ray A represents its direct illumination by the source. The global component is due to the illumination from other points in the scene. It can be caused by different physical phenomena that are common in the real world. For example, ray F is caused by the interreflection of light between scene points. Ray C and D results from single and multiple subsurface scattering within the medium beneath the surface respectively. Ray E is due to volumetric scattering by a participating medium in the scene. And ray B represents refraction and diffusion of light through translucent or transparent medium.

We will further discuss the relevant effects of direct reflection, multiple scattering or interreflections upon structured light based 3D scanning.

3.2.1   Direct Reflection

The signal we are most interested in for 3D scanning is the directly reflected light (Figure 3.3(a)). The amount of directly reflected light depends on the surface properties such as color, roughness, etc. which can be summarized in the BRDF [Nicodemus77]. In addition, light that is reflected from a smooth surface of a dielectric (or insulating) material is partially polarized. The amount of polarization of the reflected light can be computed according to Fresnel's formulae [Born99] and depends on the material properties and the orientation of surface with regard to the incident and reflected ray directions. A comprehensive polarization reflectance model can be found in Wolff and Boult [Wolff91].
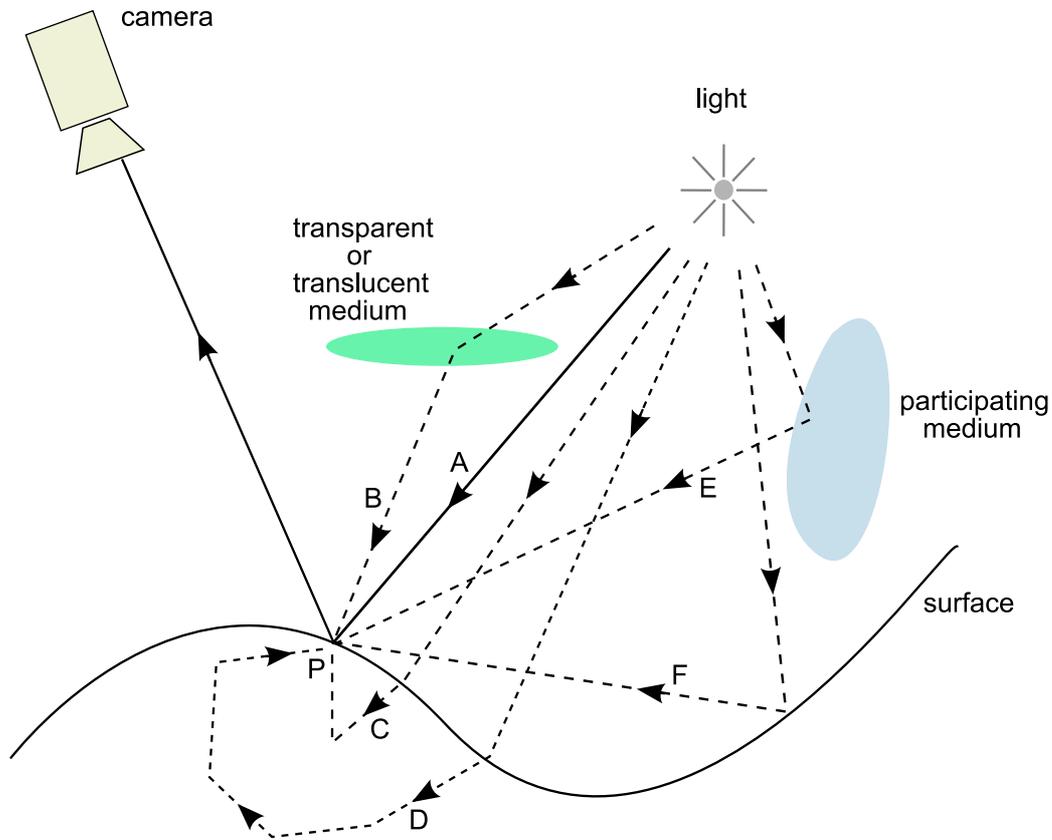
Figure 3.2: *The radiance of a scene point P is due to direct illumination of the point by the source (A) and global illumination due to other points in the scene (B, C, D, E, and F). The global illumination can arise from interreflections (F), single subsurface scattering (C), multiple subsurface scattering (D), volumetric scattering (E), and refraction (B). Only the direct reflection from (A) is directly related to the surface geometry. Therefore, a clean separation and an accurate measurement of direct reflection is very important for 3D scanning of surface geometry.*

3.2.2 Multiple Scattering

The prominent effect that distinguishes translucent from opaque materials is that some light penetrates the surface and is scattered multiple times inside the object before it finally leaves the surface at some other location. Determining the 3D shape of a translucent object requires detecting the first surface intersection of the incoming light ray, i.e. to observe the pure direct reflection (Figure 3.3(a)). Unfortunately, the signal of the direct reflection will be rather weak since some fraction of the incident light will penetrate the surface instead of being reflected. The
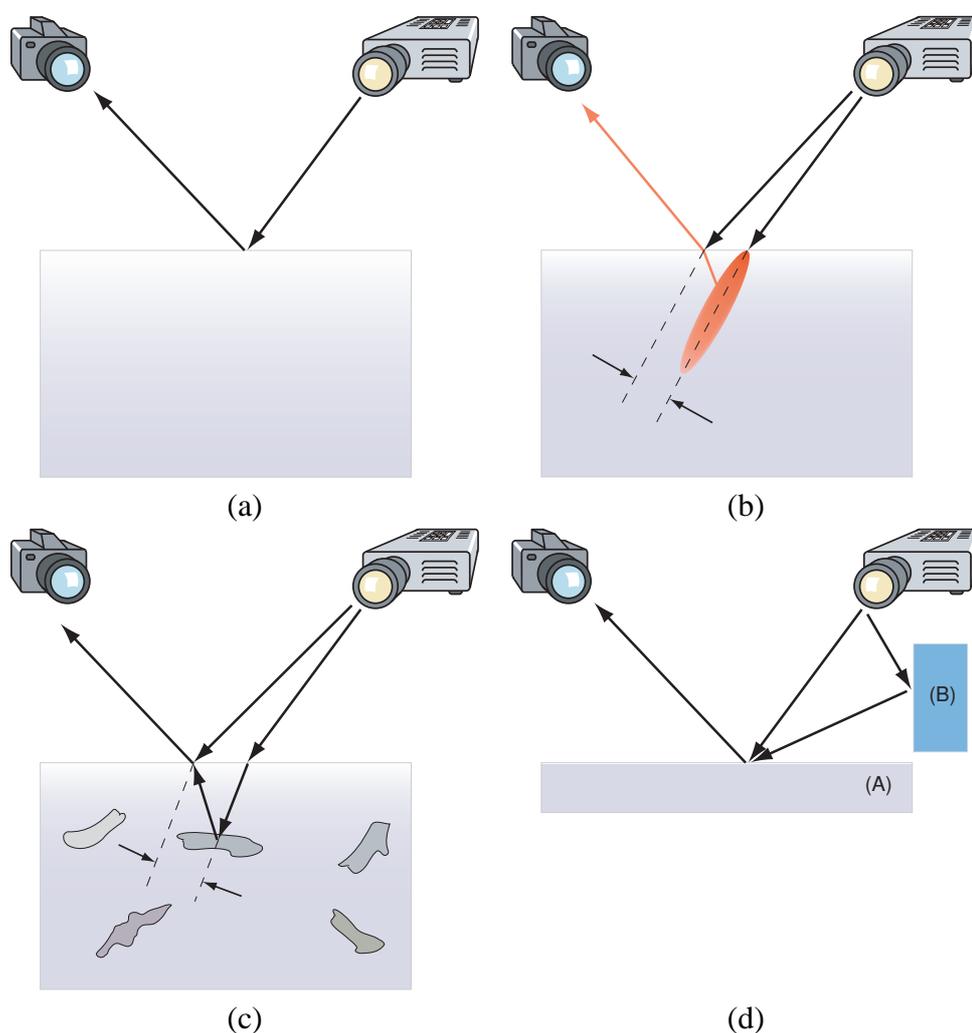
Figure 3.3: *Rays to consider in 3D scanning. (a) 3D geometry can be estimated reliably only from the direct reflection off the surface. (b) The subsurface scattering in translucent objects can shift the observed intensity peak away from the point of incidence. (c) Opaque structures beneath the surface pollute the range estimate. (d) The signals of different projector rays are overlaid due to interreflection from another surface.*

reflected signal will furthermore be heavily polluted by single or multiple scattering created by light incident on other scene points. As pointed out by Godin et al. [Godin01] multiple scattering results in a measurable bias in the depth estimate since the location of the observed intensity peak is shifted away from the point of incidence (Figure 3.3(b)). Multiple scattering can be approximated by a diffusion process [Jensen01] and leads to a significant damping of the high fre-

quencies in the incident illumination. Projecting shifted high frequency patterns, the global component will remain the same while changes can be observed in the direct reflection only. This can be used to remove the global effect algorithmically [Nayar06].

Multiple scattering further influences the state of polarization. While single scattering polarizes light according to the size and shape of a particle and the reference plane spanned by the direction of the incoming light and the scattered direction, multiple scattering due to the random orientation of particles to some degree depolarizes the incident light [van de Hulst81, Schechner03, Schechner05]. We make use of the depolarization properties to remove multiple scattering effects from the measurements.

Another important source of error is depicted in Figure 3.3(c). Here, some structure beneath the surface actually reflects more light than the direct reflection at the surface leading to wrong depth estimates (compare Figures 5.10 and Figure 5.6). While light reflected by those structures keeps the high frequencies of the incident light pattern we show in our experiments that it undergoes some degree of depolarization, which can be utilized.

3.2.3 Interreflections

Similar effects are introduced by interreflections due to nearby surfaces (Figure 3.3(d)). The signal of the direct reflection off an arbitrary surface (not necessarily translucent) is disturbed by the indirect reflection from another surface. The resulting artifacts might range from a small bias added to the depth estimate of the original surface (A) to wrongly detecting the depth of the mirror image of the other surface (B).

Depending on the reflection properties of the other surface (B) the high frequencies of the original pattern will typically be significantly reduced in the indirect reflection; for a glossy or diffuse BRDF, the illumination of a single point on surface (B) will indirectly illuminate a larger region on surface (A), hereby spreading out the signal. For second and higher order interreflections the loss of high frequencies is even more prominent.

Note, however, that interreflections might still result in linearly polarized light depending on the arrangement of surfaces (A) and (B). As a result, polarization is not always suitable for separating the direct component from interreflections.

Chapter 4
Mesostructure from Specularity

The visual appearance of a real-world object is governed by reflectance proper-
ties, illumination condition, and a hierarchy of geometric components. In the ge-
ometric hierarchy, there are basically three different levels of scales, namely, the
macrostructure level, the mesostructure level, and the microstructure level. The
macrostructure level represents the gross surface geometry, typically expressed as
a polygonal mesh or parametric spline surface. The microstructure level involves
surface microfacets that are visually indistinguishable. The mesostructure level
represents geometric details that are relatively small but still individually visible
such as bumps or dents on a surface.

Efficient mesostructure reconstruction methods can contribute greatly to high-
quality graphics models in terms of fine-scale surface geometric details. An ac-
curate and explicit mesostructure model can also benefit related mesostructure
modeling techniques such as BTFs (Bidirectional Texture Functions) [Dana99,
Müller05].

State of the art high-resolution 3D scanning methods include [Nehab05,
Davis05, Bernardini02, Levoy00]. Photometric stereo methods can achieve
high-resolution surface reconstruction with inexpensive setup [Woodham80,
Rushmeier97, Hertzmann05, Goldman05, Paterson05, Wu05]. Existing tech-
niques are, however, rarely able to capture the fine-scale details of real-world
objects with translucency or specular reflection, such as skin, rough fruit skin, etc.

By drawing inspiration from photographs of real-world translucent objects
and from the literature on human vision and perception [Fleming04, Fleming05,
Norman04, Todd04a], we found that specular highlights are important visual cues
for surface mesostructure perception and provide reliable visual information for
surface detail representation. In Figure 4.1, the hand image under point light
illumination conveys much more fine surface geometric details than that under
diffuse illumination. The specular reflection in image Figure 4.1(b) greatly helps
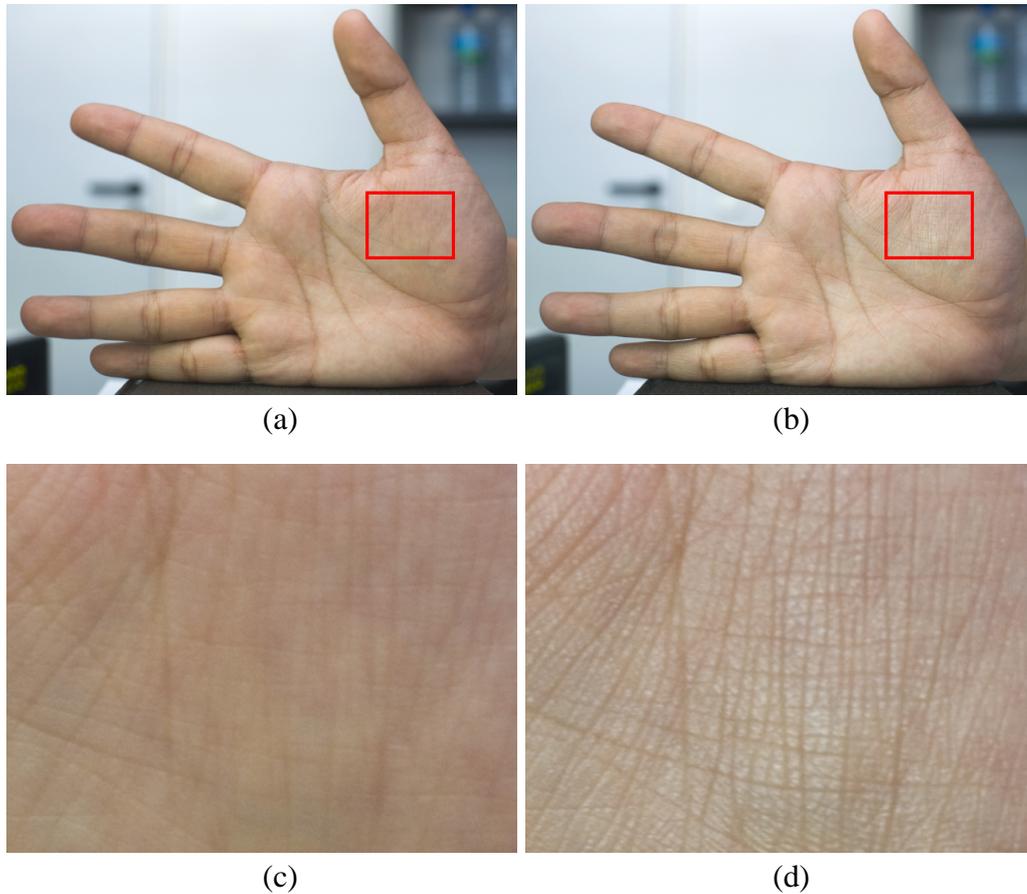
Figure 4.1: *Specularity as a strong visual cue for surface mesostructure. (a) A hand under diffuse illumination. (b) A hand under point source illumination. (c) Zoom-in of (a). (d) Zoom-in of (b). The hand image under point light illumination conveys much more fine surface geometric details than that under diffuse illumination. The specular reflection in image (b) greatly helps the perception of mesostruture of the hand.*

the perception of mesostruture of the palm. The small wrinkles on the hand skin introduce rich visual effects and can be efficiently revealed by specular highlights. Based on this observation, we developed a simple and progressive system that uses specular highlights in order to solve the dense mesostructure reconstruction problem for a variety of real-world complex objects, which possess significant specular reflection. Our method is largely independent of the underlying reflectance model, and can therefore successfully handle objects with complex reflectance that have previously been challenging.

Motivated by the simplicity of existing shape-from-specularity methods, we extend them to reconstruct surface mesostructure, fine-scale geometric surface details, from objects with a significant specular component across the surface. A large set of real-world objects possess this property, such as fruit skin, human skin, plant leaves etc. Under the dichromatic reflectance model [Shafer85, Klinker88], the light reflected from a surface comprises two physically different types of reflections, interface or surface reflection and body or subsurface reflection. The body part models conventional matte surfaces. Interface reflection that models highlights, is directly related to the surface (interface between the object and the air). We exploit only the specular reflectance, and developed a very simple system to capture dense mesostructure from complex real-world objects.

4.1 Overview

In our approach, there is no explicit reflectance model assumed. We only exploit the specular reflection, which is directly related to surface geometry.

Figure 4.2 gives an overview of our method. The overall system consists of a camera for capturing images, an LED light as the point light source, four specular spheres for light position estimation and some accessories for calibration and supporting. Before all the images are taken, the system is first both photometrically and geometrically calibrated.

At the acquisition stage, multiple images of the interesting object and the four specular spheres are captured under point illumination from different directions. We locate the regions for the interesting object and four spheres separately. Using histogram thresholding, we build up the specularity map, locate the highlight spots on the four spheres and inversely calculate the light position by intersecting all the four rays reflected from the spheres. The specularity map is interactively displayed on a screen. The user manipulating the LED light is informed and guided by this specularity map with a clear goal in mind, to fill the specularity map as much and fast as possible.

With the known light positions and the geometric calibration of the system, we can calculate the normal for each pixel on the object. Furthermore, the normal map is post-processed, e.g. hole filling and bilateral denoising. Finally, we obtain the height map of the object by integrating the normal map.

4.2 Acquisition System

Most mesostructure acquisition systems are extensions of complicated and expensive BTF acquisition systems [Wang03, Wang05, Wang06, Neubeck05]. In
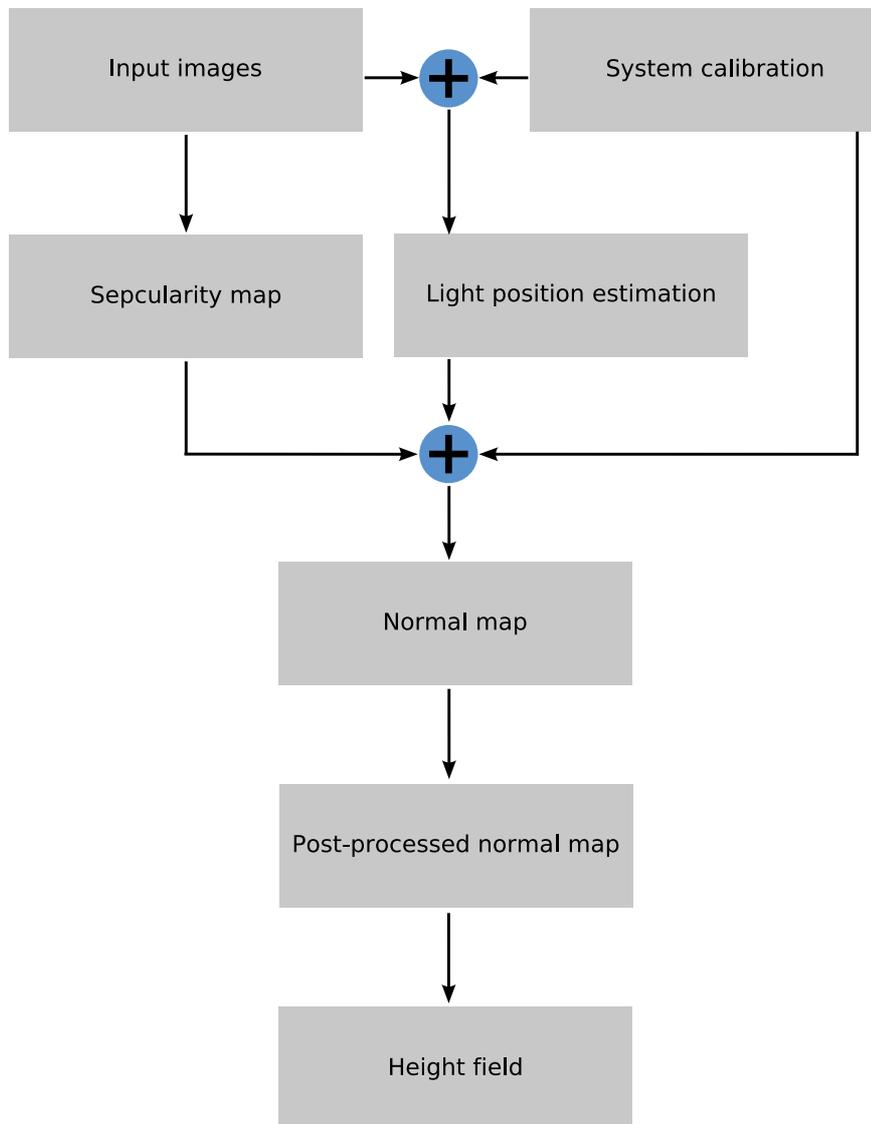
Figure 4.2*: System overview of mesostructure from specularity. The system is first photometrically and geometrically calibrated. Then multiple images are taken under varying point illumination. From the captured images, specularity map can be built using histogram thresholding and light positions can be estimated from the 4 specular spheres. With known light positions, we compute the normal map from the specularity map. The normal map is further post-processed, e.g. hole filling and denoising. The final height field is an integration result of the post-processed normal map.*
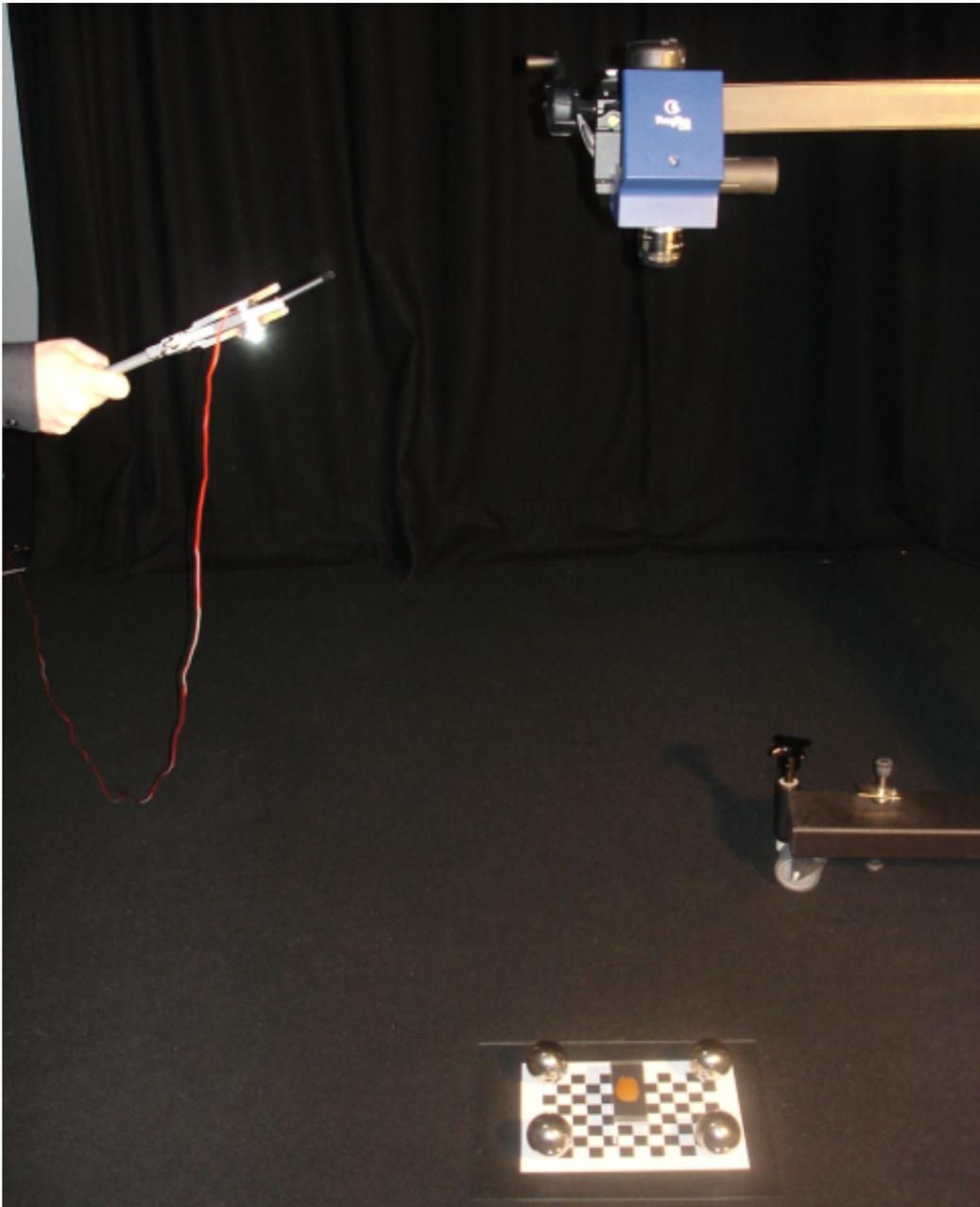
Figure 4.3: *Photograph of the acquisition system. The camera captures images while the LED is moved manually. A checkerboard is used for geometric calibration. Four specular spheres are employed for light source estimation. The sample object is placed at the center of the checkerboard. The camera is about 1.5 meters away from the sample. The light moves roughly on a virtual hemisphere with 1.5-meter radius and points always towards the sample object.*

contrast, we propose a simple, progressive, and flexible acquisition system for high-quality mesostructure recovery. The user can interactively adjust the number and distribution of light source positions without being constrained by the mechanical capabilities of the acquisition system.

The basic setup of the mesostructure acquisition system consists of a digital camera and a point light source (see Figure 4.3). We used a 12-bit 1300×1030-pixel Jenoptik ProgRes C14 digital camera for image acquisition and a 5 Watt Luxeon Star white LED as point light source. A checkerboard is used for camera calibration [Bouguet]. Four specular spheres are positioned at the four corners of the checkerboard to estimate the light source position. The sample object is placed on a small support at the center of the checkerboard. The camera faces downward to the checkerboard with the optical axis perpendicular to the checkerboard plane. We keep the camera 1.5 meters away from the checkerboard. The mesostructure has ignorable magnitude, compared to such a large distance. We also assume the base geometry of the sample object has minute scale, compared to the distance between the camera and the object.

During acquisition, we keep the light source about 1.5 meters away from the object. This allows us to approximate the LED by a point light source. To keep the illumination consistent, we always point the light to the sample object. We capture one image for each position of the point light source. Using histogram thresholding, we can in real-time extract the specular reflection component of the sample object and update the specularity field, which keeps the state of how much specularity data has been captured from the sample object. Pixels for which a specular peak was detected in at least one image are marked red; otherwise they are marked black. During the acquisition, a growing portion of the specularity field will be colored red and the user can use this feedback to move the light source in a way that increases coverage of the specularity field.

This incremental refinement allows flexible control of the quality of the result mesostructure. If the final specularity field is very dense, the resulting mesostructure will be very accurate and highly-detailed. On the other hand, if only a sparse specularity field is captured, the reconstructed mesostructure will be dominated by low frequency features. In practice, less than 200 input images can already lead to a very dense specularity field and consequently highly-detailed mesostructure for complex real-world objects (see Table 4.1). Figure 4.4 shows 64 input images for the orange skin. The specular highlights in these images are strong visual cues for mesostructure perception of the orange skin.
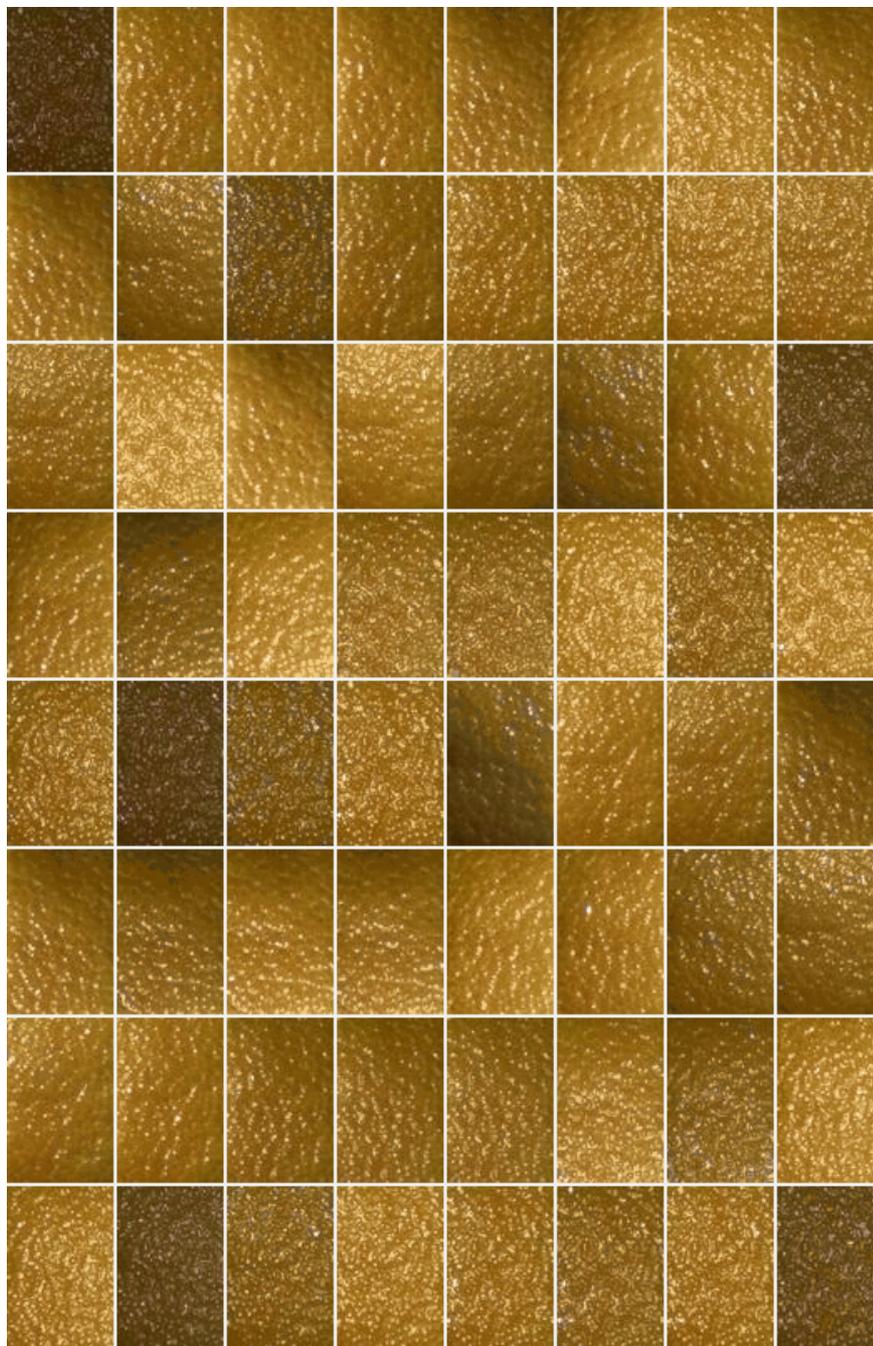
Figure 4.4: *Mosaic of 64 input images for the orange skin. The specularities in these images are strong visual cues for mesostructure perception of the orange skin.*

Sample Objects	Numbers of Input Images
black leather	35
orange skin	65
jelly candy (red)	101
jelly candy	102
human skin	183
dried apricot	188
dark brown chocolate	196

Table 4.1: *Numbers of input images captured for sample objects used in our experiments. Even for complex real-world objects, less than 200 images are sufficient for high quality mesostructure reconstruction.*

4.3 Light Source Estimation

Light source estimation from images is a well-researched problem in computer vision. In our experiments, we use four specular spheres arranged in a rectangle on a common plane to estimate the light source position. Masselus et al. [Masselus02] use diffuse spheres in a similar configuration for this purpose. The law of reflection yields the geometric constraint that the specular normal N is a bisector between the light source direction L and the viewing direction V (see Figure 4.5), i.e.,

$$V = 2(N^T L)N - L. \tag{4.1}$$

With the known geometry of the camera and spheres, and the detected specular peaks on the spheres, we can easily compute the reflected light directions of the spheres according to Equation 4.1. Let the four reflected rays be $R_i = O_i + D_i t_i, i = 0, 1, 2, 3$, where O_i and D_i are the ray origin and ray direction respectively. Since there is only one point light source, we can construct an overdetermined linear system with six equations and four unknowns. Using SVD, we can compute the four unknowns $t_i, i = 0, 1, 2, 3$, and then the light source position.

Figure 4.6 shows a simplified 2D case with only two spheres. O is the optical center of the camera. L is the light. S_1 and S_2 are two specular spheres. Light rays from L are reflected by the spheres S_1 and S_2 and imaged by the camera at points A and B respectively. In the actual computation, we trace reversely light rays OA and OB and intersect the two spheres, S_1 and S_2 at A' and B' respectively. According to the law of reflection and the known geometry of the spheres, we get the light position estimation by intersecting two reflected rays $A'L$ and $B'L$.
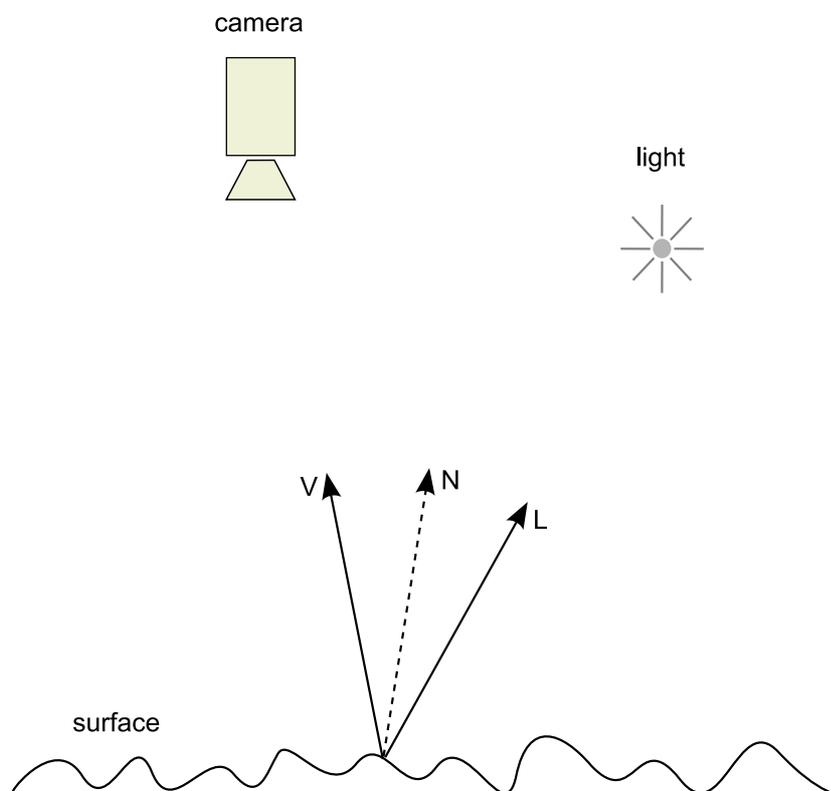
Figure 4.5: *Geometric constraint of specular reflection. At a specular point, the surface normal N is the bisector between the light vector L and viewing vector V. In one captured image, a large set of surface normals can be recovered simultaneously with known light source and camera positions.*

4.4   Specularity Field Construction

Efficient and robust separation of diffuse and specular components of surface reflection for arbitrary materials – especially for translucent or refractive materials – is still an open problem. State of the art techniques exploit both polarization and color appearance [Umeyama04, Tan05, Nayar97, Wolff90].

In experiments, we found that color appearance is no longer a reliable cue for specularity isolation, especially when translucent materials are involved. On the other hand, polarization imaging techniques pose high requirements for progressive and interactive acquisition. With these considerations, we used a simple histogram thresholding method to extract the specular reflection component.

During the progressive acquisition process, for each input image, we construct the intensity histogram as in Figure 4.7. Let $H(i)$ be the histogram, where i
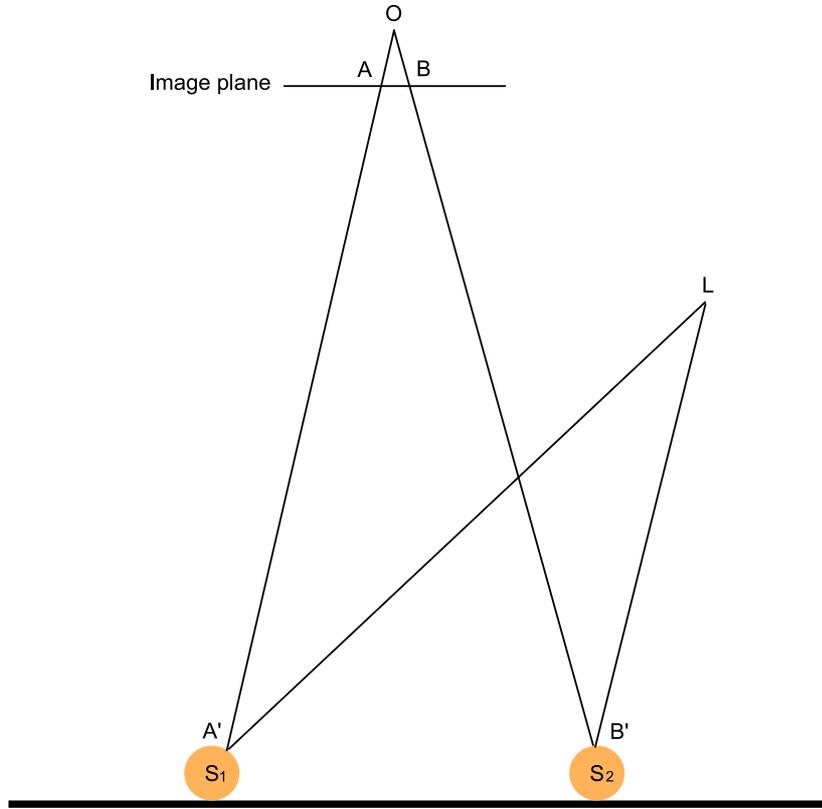
Figure 4.6*: Light source estimation. O is the optical center of the camera. L is the light. S_1 and S_2 are two specular spheres. Light rays from L are reflected by the spheres S_1 and S_2 and imaged by the camera at points A and B respectively.*

represents intensity and $H(i)$ is the number of pixels with intensity i in a specific image. We define the histogram gradient as

$$G_h(i) = |H(i-1) - H(i)|. \tag{4.2}$$

For a specific gradient threshold T_g, we can find the associated intensity threshold T_i from the histogram $H(i)$. Then the specularity field, indicating the position of specular reflection, can be computed by thresholding the image according to T_i, i.e., there is specularity if the pixel value is larger than T_i. Denote $\mathcal{F}_s$ as the specularity field, which is defined as

$$\mathcal{F}_s(x,y) = \begin{cases} 1, & \text{specular highlight was observed at } (x,y); \\ 0, & \text{otherwise.} \end{cases} \tag{4.3}$$

Each pixel (x,y) in the specularity field $\mathcal{F}_s$ will be updated while the acquisition proceeds. The current state of the specularity field is displayed for active
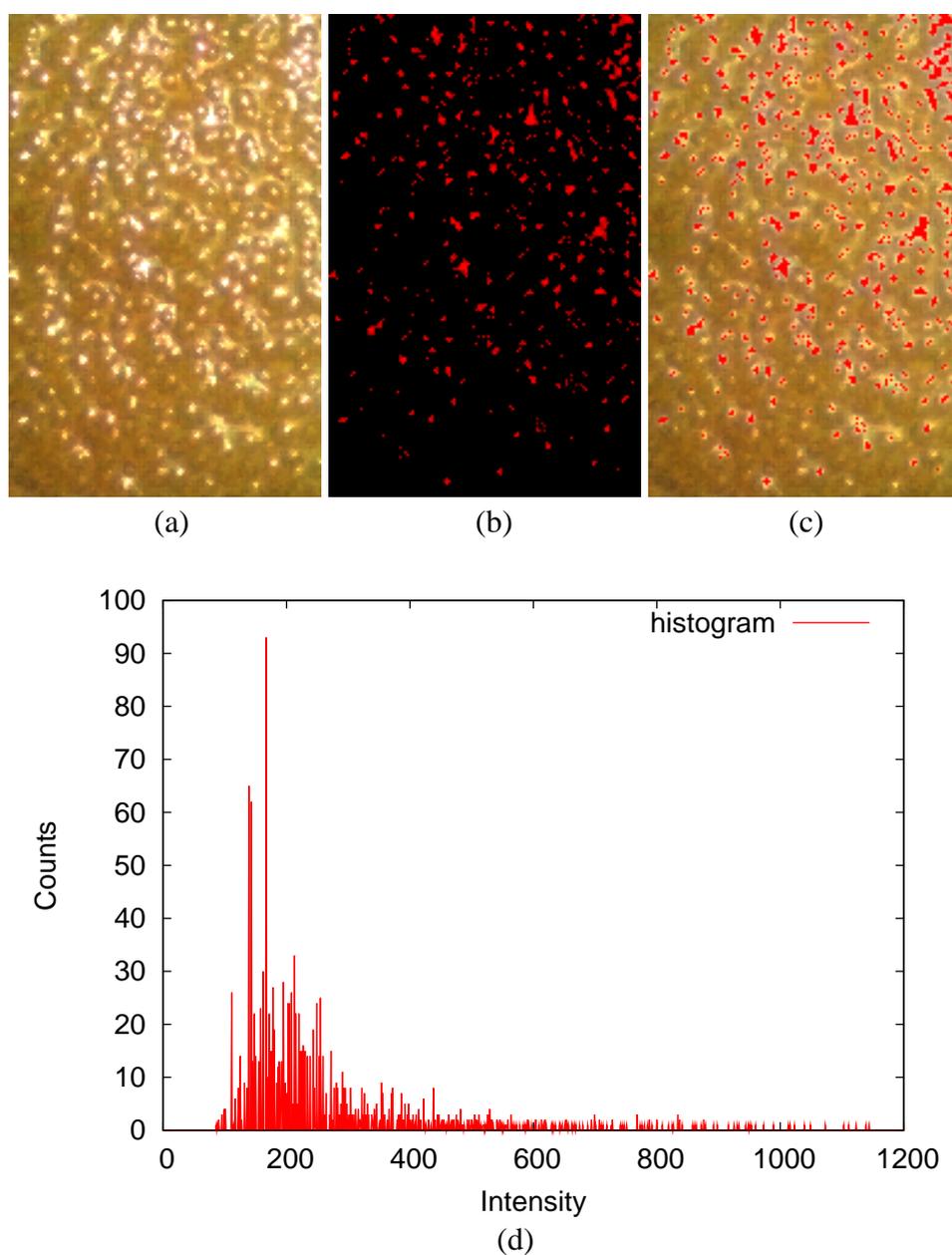
Figure 4.7: *Specularity separation via histogram thresholding. (a) Input image. (b) Extracted specularity field. (c) Input image with specularity marked. (d) Histogram. With a user specified histogram gradient threshold $T_g = 10$, an intensity threshold $T_i = 491$ can be found in the histogram. The specularity field can be computed from the input image by simple thresholding. Note that different input images have the same T_g but different intensity thresholds T_i.*

control of the light source. In our experiments, less than 200 images were required to recover almost complete specularity fields. For each pixel, there are usually multiple images in which it was classified as specular. We select the image where the brightest highlight occurred to calculate the normal. Images with overexposed pixels are avoided using the exposure settings of the camera.

4.5   Inferring Normal Field

With the known specularity field $\mathcal{F}_s$ and the known light source positions for each input image, we can easily compute the surface normal of each pixel (x, y) on $\mathcal{F}_s$ (see Figure 4.5) yielding a normal field $\mathcal{F}_n$. In order to remove noise from $\mathcal{F}_n$, we filter it using the edge-preserving bilateral filter [Tomasi98]. The resulting normal field can be used for direct rendering of mesostructure by bump mapping [Blinn78] or to reconstruct the 3D mesostructure. To simply convey the reconstructed mesostructure, we use Ward's isotropic BRDF model [Larson92] to render the normal field under directional illumination.

4.6   3D Mesostructure Reconstruction

To compute an accurate height field of the surface mesostructure from the estimated normal field we use the method proposed by Frankot and Chellappa [Frankot88] which uses surface smoothness as an integrability constraint.

Let $p = \partial z / \partial x$ and $q = \partial z / \partial y$ be the surface gradients in x and y directions respectively. $N(x, y)$ is the normalized surface normal at position (x, y), which is related to the partial derivatives $\partial z / \partial x$ and $\partial z / \partial y$ through the formula:

$$N(x, y) = \frac{g}{\|g\|} \tag{4.4}$$

where $g = \{\partial z / \partial x, \partial z / \partial y, 1\}$. We will refer to the equality of the second order partial derivatives as the integrability constraint, i.e.,

$$\frac{\partial p}{\partial y} = \frac{\partial q}{\partial x} \tag{4.5}$$

for all positions (x, y) on the support of the normal field. That is, they correspond to a surface with second order partial derivatives independent of the order of differentiation. Consequently, the surface height at any particular position is independent of the path of integration.

We applied this method on all our experimental examples and got stable results even with a noisy input normal field. To increase performance we compute the integration through Fourier expansion.

4.7 Experimental Results

Previous mesostructure reconstruction techniques have difficulties with translucent, highly specular, or low albedo glossy materials. The major advantage of our method is that it can handle successfully these kinds of materials. In experiments, we choose a piece of dark brown chocolate and a piece of black leather, both of them have glossy reflection. To test translucent materials, we use orange skin, dried apricot, human skin, and jelly candy. The results demonstrate that our method can effectively reconstruct surface mesostructure from complex materials.

4.7.1 Very Low Albedo Glossy Materials

Figure 4.8 and Figure 4.9 shows mesostructure reconstructions for a piece of black leather and a piece of dark brown chocolate respectively. The chocolate has larger scale mesostructure than the leather. We successfully recover the large-scale shape of the chocolate and get a relatively smooth surface. With only 35 input images, we obtain accurately the fine-scale details of the leather.

4.7.2 Translucent Glossy Materials

Photometric stereo techniques are capable of measuring fine surface details. Effectively modeling surface details of translucent material is, however, still an open problem. We captured four real-world objects with various degrees of translucency: orange skin (Figure 1.1), dried apricot (Figure 4.10), human skin (Figure 4.11), and jelly candy (Figures 4.12 and 4.13).

Our system can effectively deal with the orange skin and the dried apricot whose fine-scale surface details are accurately recovered. It has, however, some difficulties with the jelly candy's boundary, while the inner part of the jelly candy is successfully reconstructed. The reason for this is twofold. One is that the strong caustics around the jelly candy's boundary makes accurate specularity detection difficult. The other reason is that there are some surface normals on the boundary facing away from the camera, which cannot be captured with the current setup. To overcome this problem, we need additional cameras or capture several views of the object.

The skin case is notoriously known to be difficult in computer vision or graphics. The results in Figure 4.11 demonstrate the effectiveness of our method. From the renderings at different viewpoints, we can clearly see the well-reconstructed fine wrinkles and pores. A promising extension is to measure mesostructure from general 3D skin, which can find important applications in computer graphics.
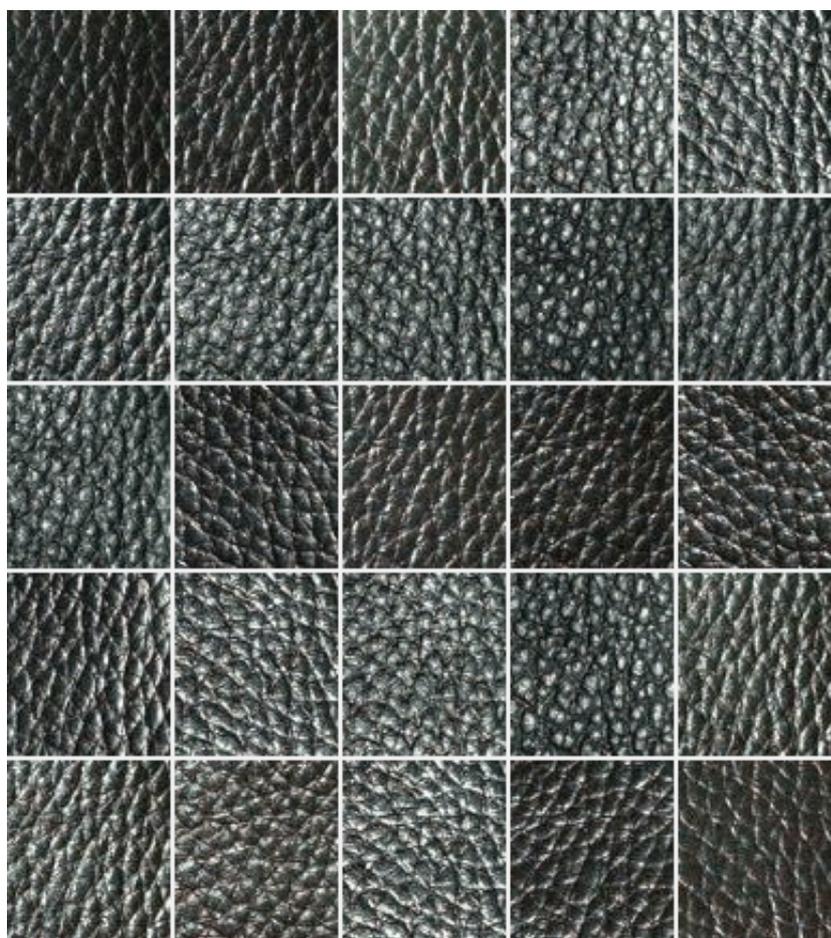
(a)
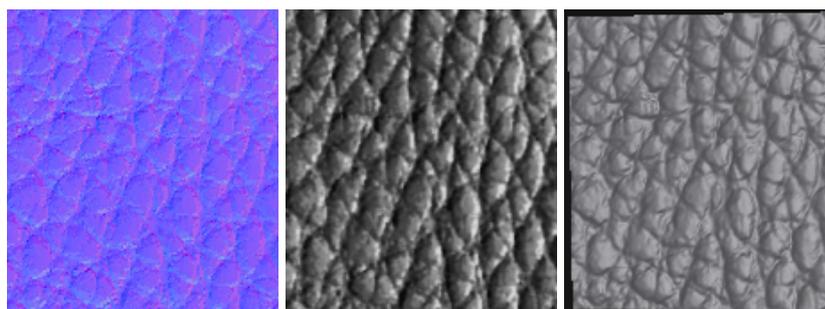
(b) (c) (d)

Figure 4.8: *Mesostructure reconstruction of a piece of black leather. (a) Mosaic of input images. (b) Recovered normal field. (c) Renderings of the normal field using Ward's isotropic BRDF model [Larson92]. (d) Reconstructed 3D model rendered at a novel viewpoint.*
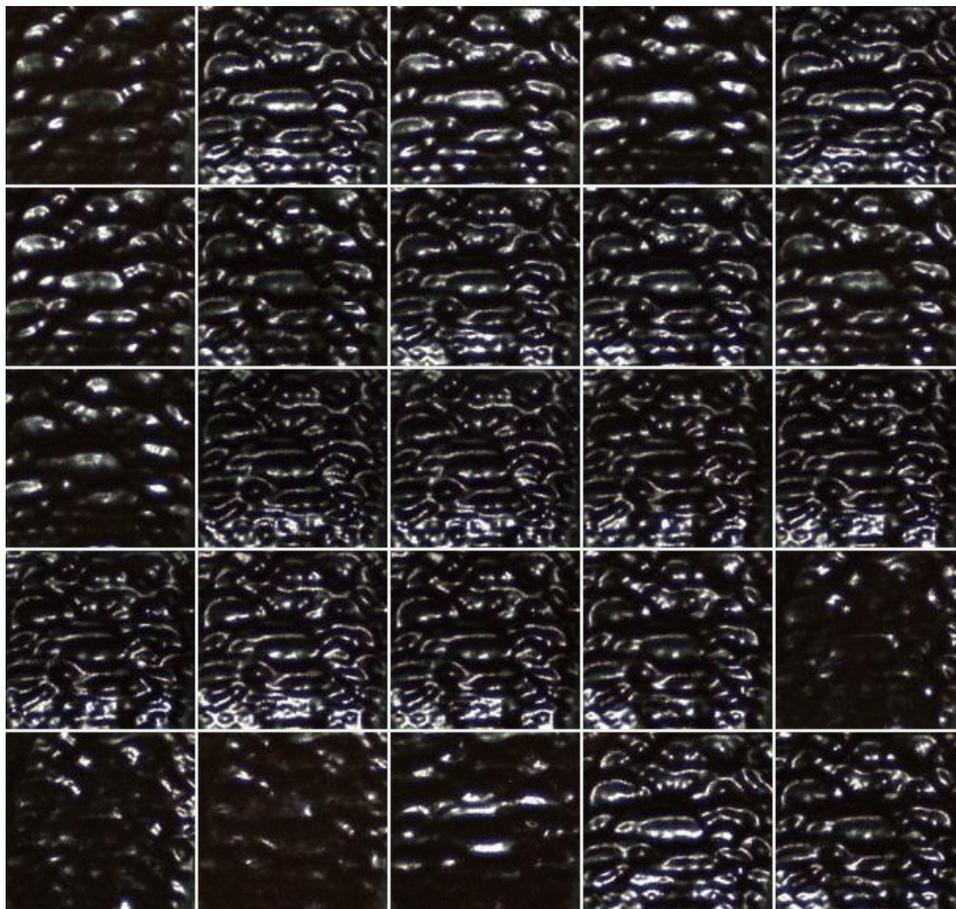
(a)
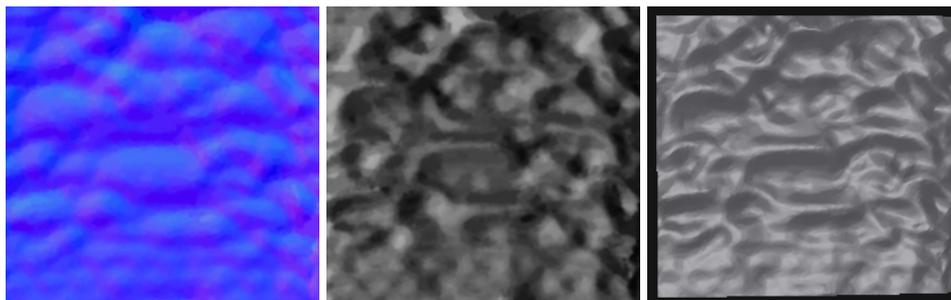
(b)   (c)   (d)

Figure 4.9: *Mesostructure reconstruction of a piece of dark brown chocolate. (a) Mosaic of input images. (b) Recovered normal field. (c) Renderings of the normal field using Ward's isotropic BRDF model [Larson92]. (d) Reconstructed 3D model rendered at a novel viewpoint.*

(a)

(b)   (c)   (d)
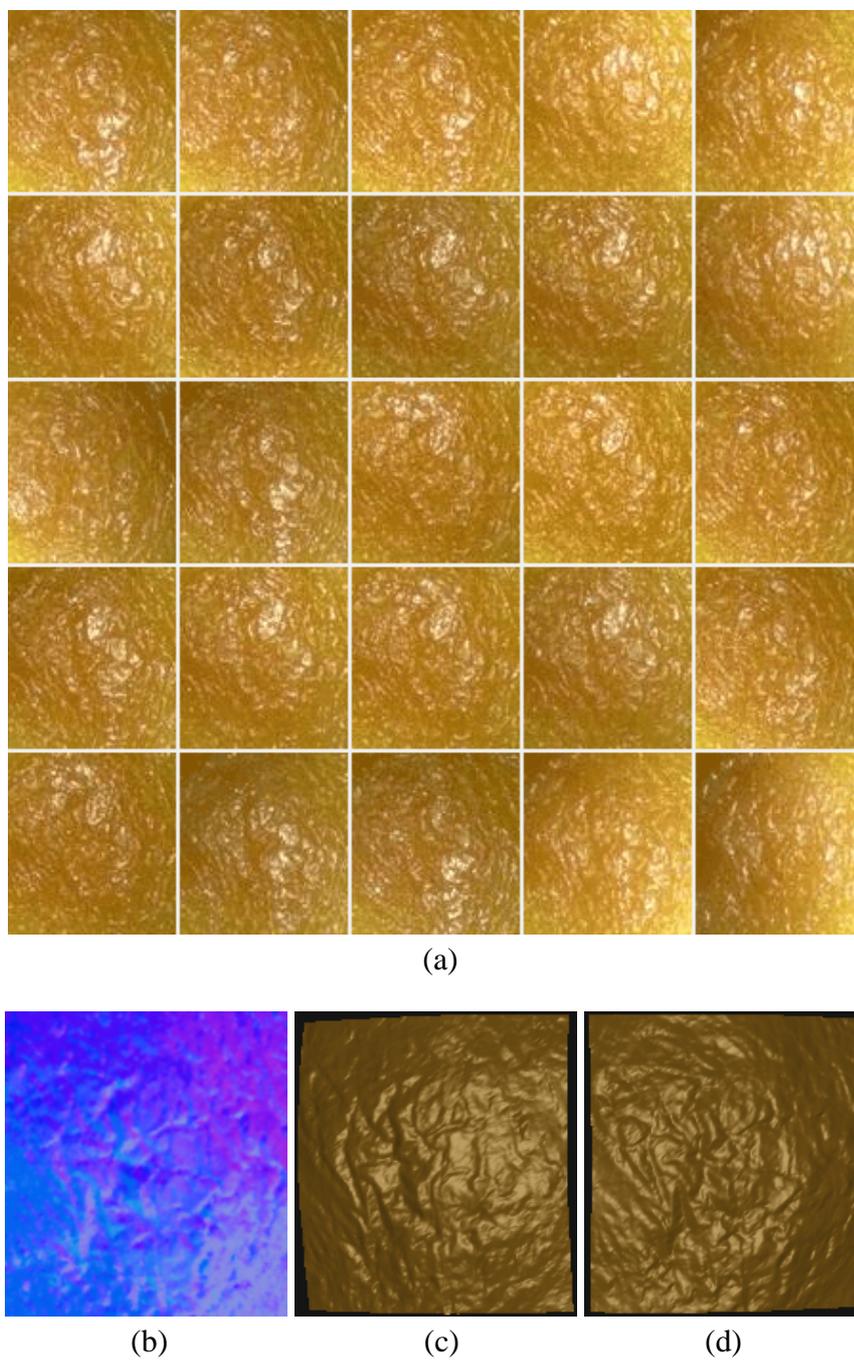
Figure 4.10: *Mesostructure reconstruction of a piece of dried apricot. (a) Mosaic of input images. (b) Recovered normal field. (c-d) Reconstructed 3D model rendered at novel viewpoints.*
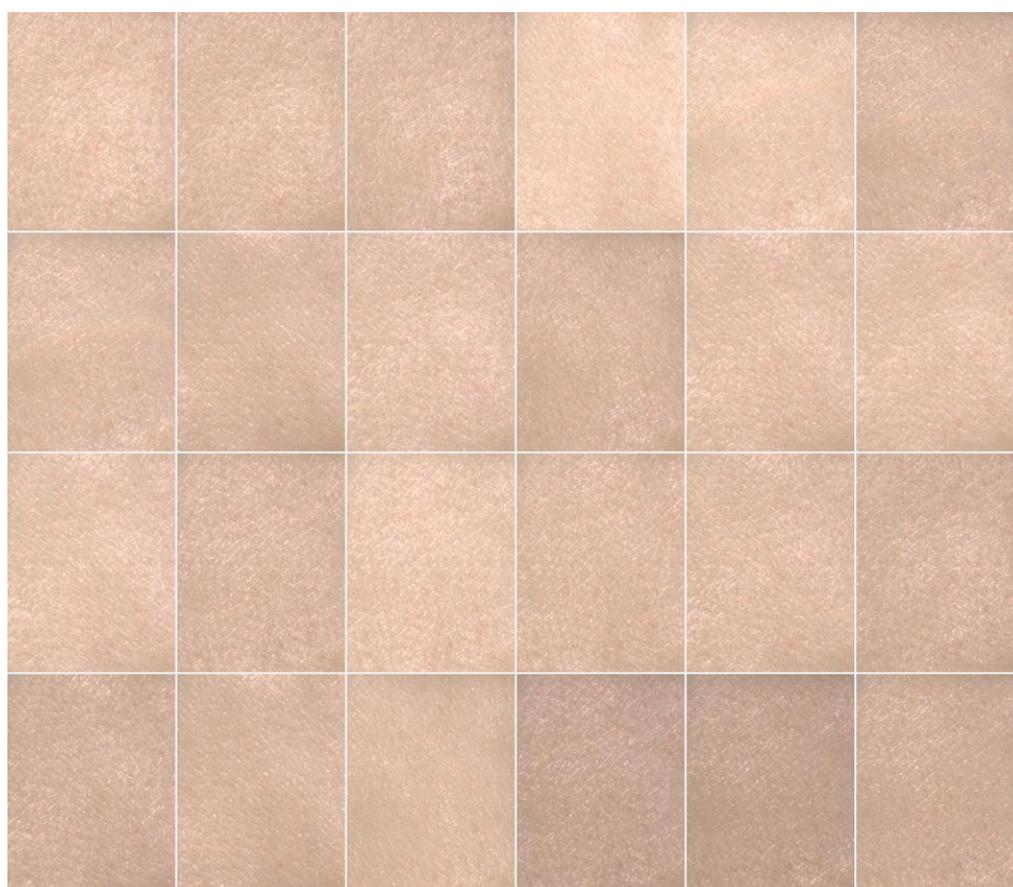
(a)
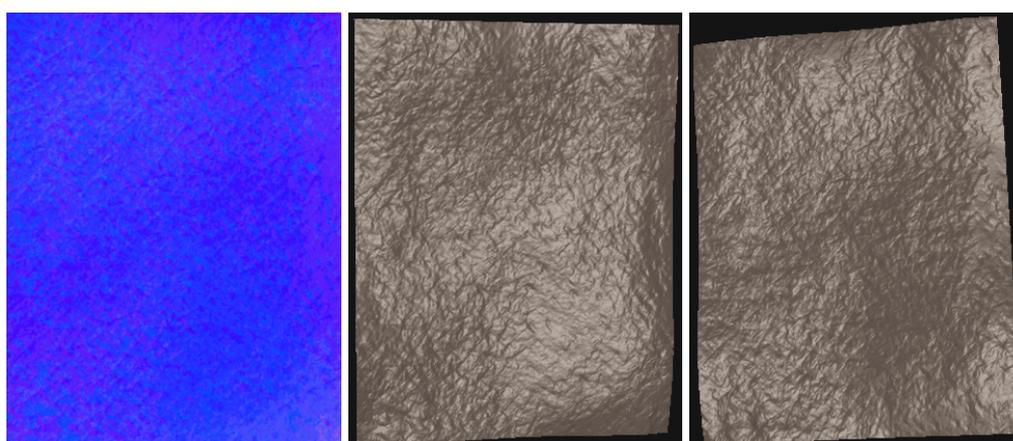
(b)   (c)   (d)

Figure 4.11*: Mesostructure reconstruction of skin. (a) Mosaic of input images. (b) Recovered normal field. (c-d) Reconstructed 3D model rendered at novel viewpoints.*
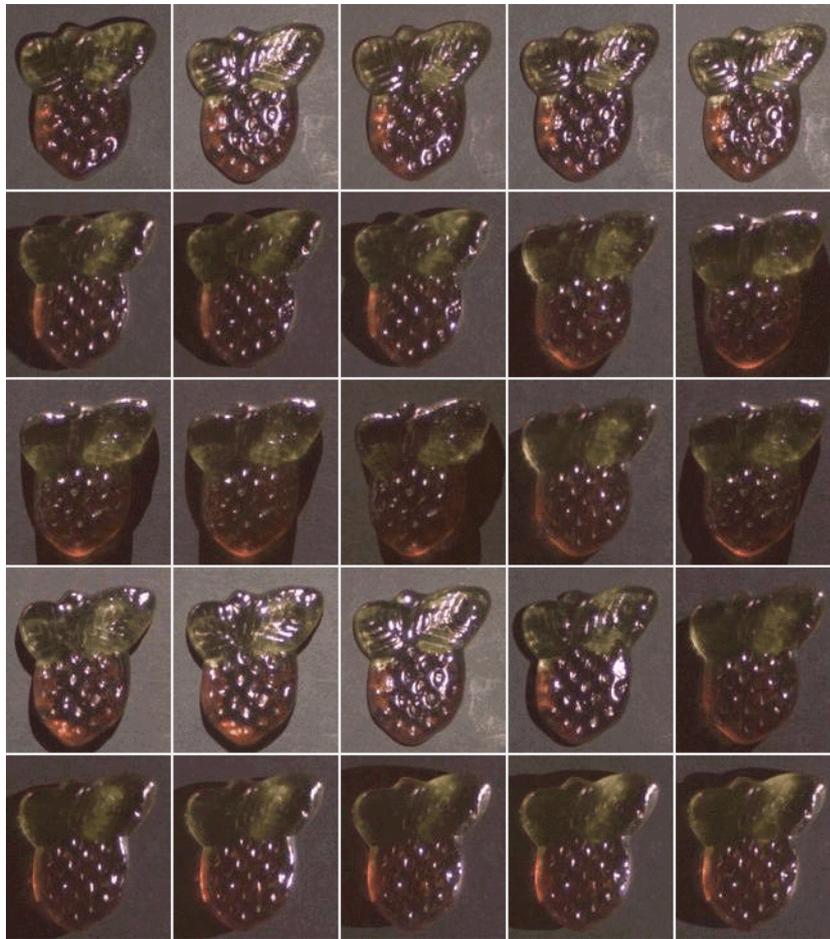
(a)
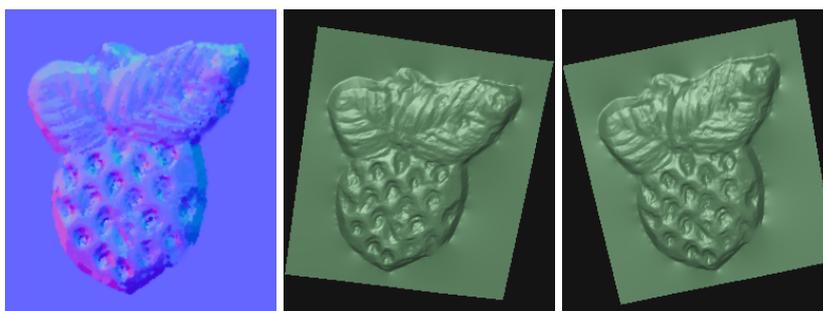
(b)   (c)   (d)

Figure 4.12*: Mesostructure reconstruction of a piece of jelly candy. (a) Mosaic of input images. (b) Recovered normal field. (c-d) Reconstructed 3D model rendered at novel viewpoints.*

(a)   (b)

(c)   (d)
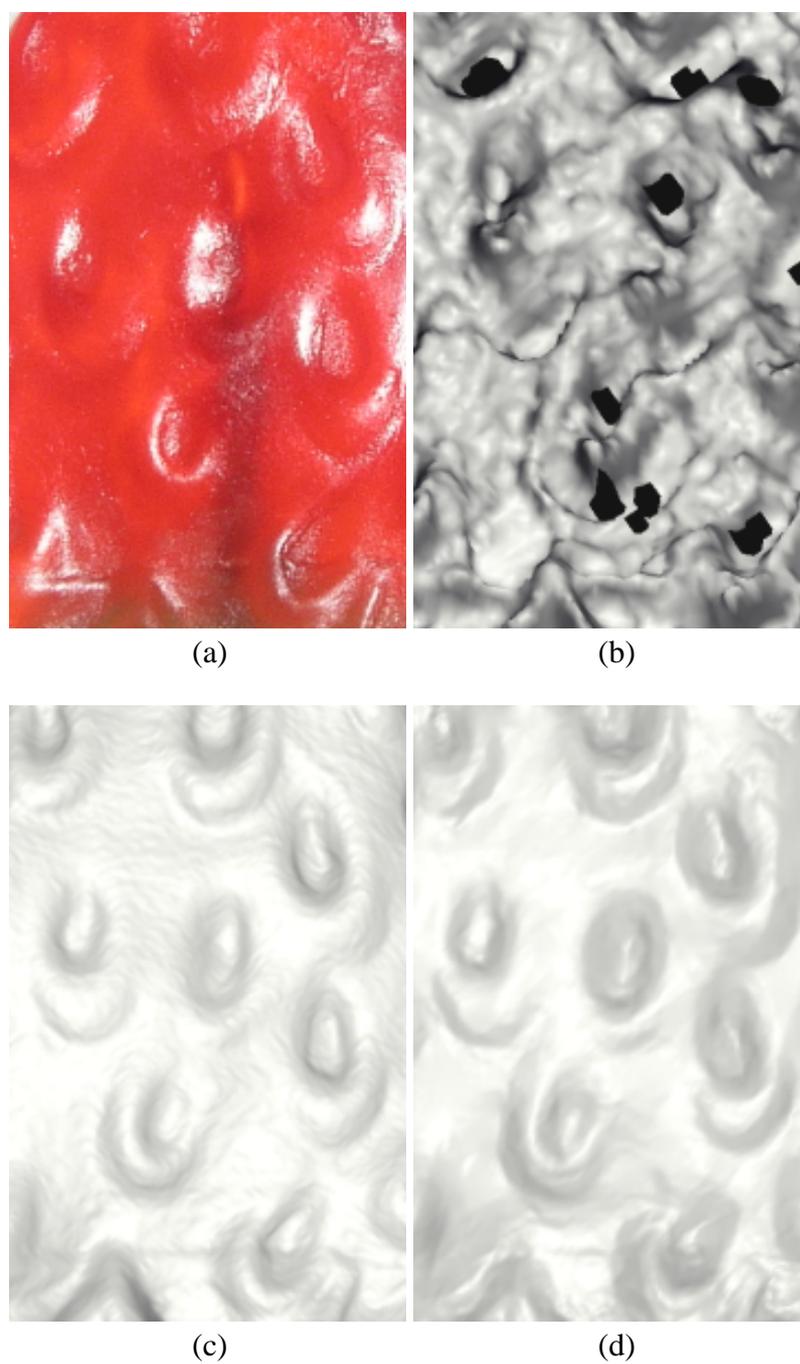
Figure 4.13: *Mesostructure reconstruction of a piece of jelly candy. (a) Input image. (b) Shape by laser scanner. (c) Shape by laser scanner after covering the object with Lambertian powder, which is close to the ground truth. (d) Shape by our method. It is very close to (c).*

We also captured a piece of highly translucent jelly candy with regular geo-metric pattern (see Figure 4.13). In order to compare against ground truth, we captured its surface structure also with a Minolta VI-910 laser range scanner. As expected, the laser scanner has difficulties recovering the shape because the laser light is scattered inside the object. The recovered position information is unreli-able and the final geometry reconstruction shows strong artifacts. We then covered the jelly candy with a layer of very fine and homogeneous white powder in order to make it Lambertian and performed another laser range scan. Our mesostruc-ture reconstruction compares favorably against this last scan, which is the closest approximation of ground truth we were able to capture for this type of object.

4.8   Limitations

Although our method can reconstruct fine-scale geometry details, the mesostruc-ture of complex real-world objects, there are still several limitations existing. To list some of them:

- For high quality reconstruction, we need dozens of input images, which limits our method to be applicable only to still objects. It would be a very interesting extension to decrease the number of input images while still ex-ploiting the fundamental power of shape from specularities. Some recent research went in this direction [Ma07, Francken08].

- The range of recovered normals are limited. The maximum angle between viewing direction and light direction is 90 degrees, and hence the maximum angle between surface normal and viewing direction is 45 degrees.

- Our system only recovers the surface normals of the object. In order to reconstruct the absolute 3D of the surface, we need absolute position infor-mation as reference for solving the ambiguity in normal integration.

- We simply use histogram threshold to separate specular reflection from dif-fuse reflection. It would be interesting to use a more advanced and efficient separation method, such as methods using color, polarization or statistics.

Chapter 5

Phase-Shifting and Polarization-Based 3D Scanning

This chapter focuses on structured light based 3D scanning of translucent objects. The most important problem here is to separate the direct reflection component from any global illumination effect. Current separation approaches are either based on polarization, or on structured, high frequency illumination.

For a number of scenes (see Figure 5.1), structured light based 3D scanning techniques run into the problem that the signal observed by the camera for a surface point is actually not only due to direct reflection of the projected pattern but instead contains polluting signals originating from ambient illumination, interreflections from other scene parts, or from subsurface scattering. These effects are most prominent in translucent objects where the directly reflected signal is furthermore weakened since the incident light is diffused inside the material instead of being fully reflected at the surface. Subsurface scattering can of course be excluded completely if the object's surface is painted before scanning, as it is done frequently. In this thesis, we propose 3D scanning techniques which are inherently robust against subsurface scattering.

In order to obtain reliable scans of translucent objects one has to separate the direct reflection from the pollution due to multiple interreflections or scattering. One approach to descattering is to use the fact that light scattered multiple times gets depolarized. Projecting polarized light and computing the difference of images captured with a polarization filter at two orthogonal orientations thus removes most of the multiple scattering contribution [Wolff94, Rowe95, Schechner05, Treibitz06]. Another method for separating direct from global reflections based on high frequency illumination patterns has recently been proposed by Nayar et al. [Nayar06]. Their approach is based on the insight that global effects significantly damp high frequencies. Illuminating the

(a) Grapes

(b) Alabaster stone

(c) Jelly candy
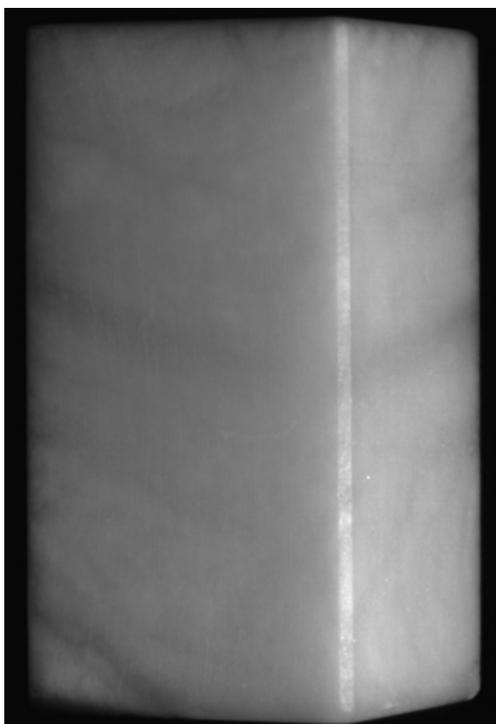
(d) Alabaster slab

Figure 5.1: *Translucent objects. For these objects, traditional optical 3D scanning techniques run into the problem that the signal observed by the camera for a surface point is actually not only due to direct reflection of the projected pattern but instead contains polluting signals originating from ambient illumination, interreflections from other scene parts, or from subsurface scattering. Even for human eyes, it is not easy to recognize the surface geometric details of a highly translucent object because of its strong subsurface scattering.*

scene with shifted high frequency patterns therefore will result in high frequencies observable in the direct reflection part only. Various patterns have been proposed by Nayar et al. [Nayar06] to perform the separation ranging from checkerboard and simple stripe patterns to sinusoids.

In [Nayar06], Nayar et al. also mentioned that phase-shifting [Srinivasan85, Zhang06e] can perform the separation and 3D scanning at the same time. Our 3D scanning approach for translucent objects is also based on phase-shifting. We demonstrate and analyze why descattering based on structured light alone is not sufficient to obtain high quality depth maps of heterogeneous translucent objects. Our proposed method therefore combines phase-shifting with polarization filtering. The increased performance is demonstrated and assessed on a variety of translucent objects.

5.1 Overview

We perform 3D scanning of translucent objects by exploiting both high frequency illumination, which possesses algorithmic descattering properties, and polarization, which performs reflections separation physically/optically. For the high frequency illumination pattern, we choose N-step phase-shifting, which is proven to be robust and accurate and can result in absolution 3D coordinates. Figure 5.2 shows schematically the pipeline of our methods.

The acquisition system consists of a projector, a camera, and two linear polarizers (see Figure 5.7). The projector generates a series of sine patterns while the camera takes images simultaneously. In the simplest setup, i.e. the N-step phase-shifting, no polarizer is used. In parallel polarization imaging, two polarizers are set up with parallel polarization orientations. In the polarization difference imaging, an additional set of images is taken by rotating the two polarizers to be orthogonal. The differences between the parallel and cross image sets, are used for ordinary phase estimation. The image acquisition procedure is repeated for all the frequencies, usually with $f = \{\frac{1}{1024}, \frac{1}{512}, \frac{1}{256}, \frac{1}{128}, \frac{1}{64}, \frac{1}{32}, \frac{1}{16}, \frac{1}{8}\}$. The unit for frequency is $cycles/pixel$.

Before the real acquisition, we first calibrate the projector and the camera photometrically, and geometrically calibrate the projector-camera system, which is crucial for triangulation. For the different scanning schemes, we get different sets of input images. Then we compute the phase map for each frequency, knowing all are N-step phase-shifting. For the recovered phase maps, we perform temporal phase unwrapping and get the absolute phase map for the highest frequency, here $f = \frac{1}{8}$. The projector-camera system is calibrated beforehand, we perform triangulation and obtain the 3D coordinates for all the pixels. Since the pixels are regularly spaced, we easily obtain the 3D mesh.
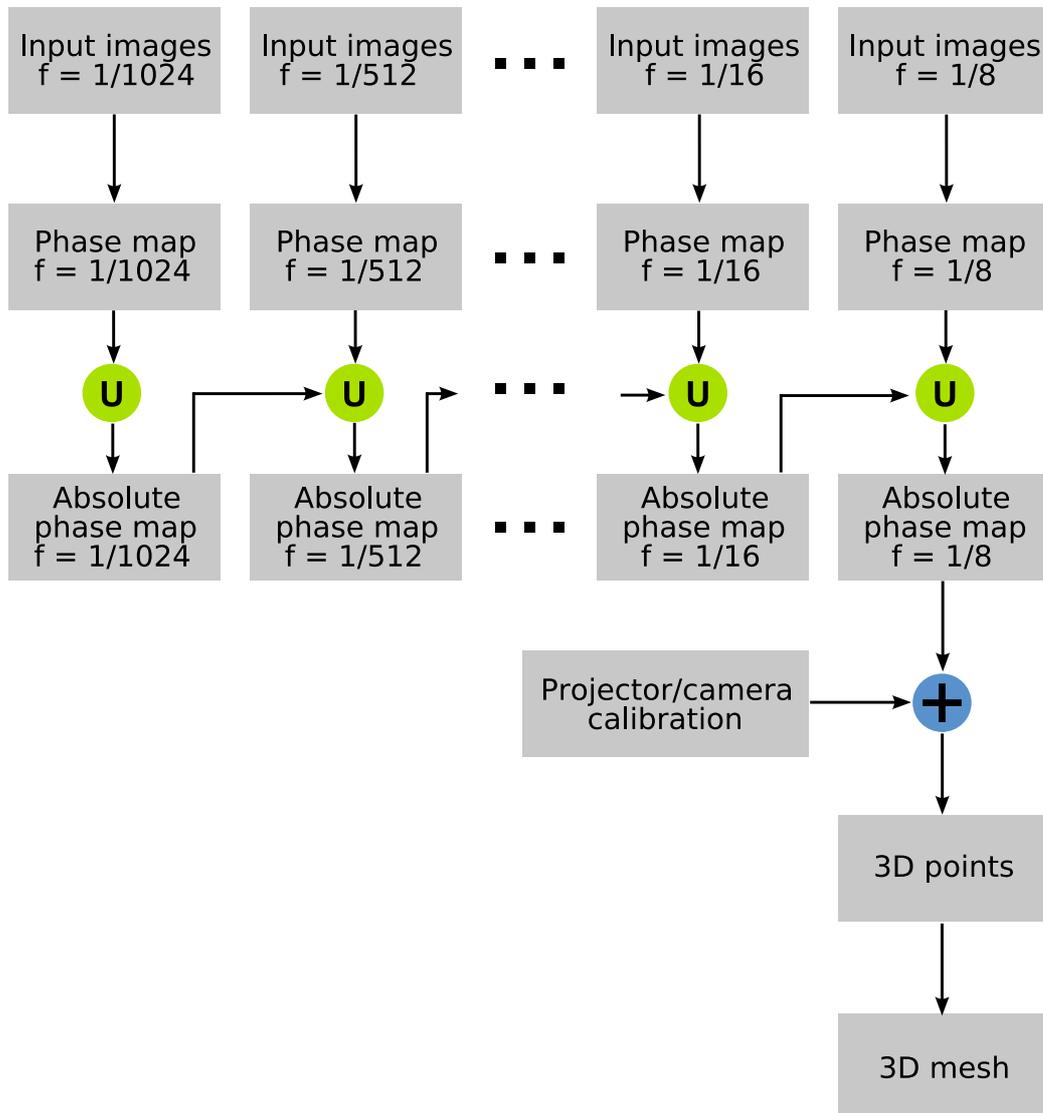
Figure 5.2: *System overview of phase-shifting and polarization based 3D scanning. The input images can be simply the images taken with N-step phase-shifting pattern projected. They can also be the images taken using parallel polarization imaging or polarization difference imaging. The image acquisition procedure is repeated for all the frequencies. Then we compute the phase map for each frequency, knowing all are N-step phase-shifting. For the recovered phase maps, we perform temporal phase unwrapping and get the absolute phase map for the highest frequency. Assuming the projector-camera system is calibrated beforehand, we perform triangulation and obtain the 3D coordinates for all the pixels. Since the pixels are regularly spaced, we can easily get the 3D mesh.*

5.2   Phase-Shifting

In the past years, many different phase-shifting algorithms have been developed [Ghiglia98, Srinivasan85, Wust91, Huang06, Zhang06d, Zhang06e, Weise07, Zhang06c, Zhang06a, Zhang06b]. The measurement accuracy of phase-shifting is usually affected by the noise and inaccuracy of the source and sensor, e.g. imaging noise of the camera, and nonlinearity and light leakage of the projector. One of the most accurate methods is based on the N-step least squares phase-shifting algorithm [Zhang06d] for a projector-camera system. Sinusoid patterns are generated by the projector and shifted by a factor of $\frac{2\pi}{N}$ for N times as

$$L_i(x, y) = 0.5 + 0.5 \cos(2\pi f x + \delta_i), \tag{5.1}$$

where (x, y) is the projector coordinate. The sinusoid is varying only in the x dimension. The frequency is denoted by f. $\delta_i = \frac{2\pi i}{N}, i = 1, 2, ..., N$ denote the N different phase shifts. The unit for frequency is $cycles/pixel$.

  The camera records one image for each phase-shift and N images in total. The resulting intensity at camera pixel (x, y) in step i can be expressed as follows:

$$
\begin{aligned}
I_i(x, y) &= I' + I'' cos(\Phi(x, y) + \delta_i) \\
&= a_0(x, y) + a_1(x, y)cos(\delta_i) + a_2(x, y)sin(\delta_i),
\end{aligned}
\tag{5.2}
$$

where I' is the average intensity, I'' the intensity amplitude, and Φ the phase to be solved. Solving the set of equations given in Eq. 5.2 in a least-squares sense, we obtain,

$$\Phi(x, y) = tan^{-1}(\frac{-a_2(x, y)}{a_1(x, y)}), \tag{5.3}$$

where

$$[a_0(x, y) \ a_1(x, y) \ a_2(x, y)]^T = A^{-1}(\delta_i)B(x, y, \delta_i), \tag{5.4}$$

$$
A(\delta_i) = \begin{bmatrix}
N & \sum cos(\delta_i) & \sum sin(\delta_i) \\
\sum cos(\delta_i) & \sum cos^2(\delta_i) & \sum cos(\delta_i)sin(\delta_i) \\
\sum sin(\delta_i) & \sum cos(\delta_i)sin(\delta_i) & \sum sin^2(\delta_i)
\end{bmatrix},
\tag{5.5}
$$

and

$$B(x, y, \delta_i) = \left[\sum I_i \ \ \sum I_i cos(\delta_i) \ \ \sum I_i sin(\delta_i)\right]^T. \tag{5.6}$$

All the sums are over the N measurements. This result has also been observed in communication theory when detecting noise corrupted signals using synchronous detection [Bruning74]. At the same time we can use the ratio γ of the observed amplitude over the observed bias as a measure for the reliability of the phase estimation:

$$\gamma = \frac{\sqrt{a_1^2 + a_2^2}}{a_0}. \tag{5.7}$$

(a) $\delta_i = 0$ (b) $\delta_i = \frac{2\pi}{3}$ (c) $\delta_i = \frac{4\pi}{3}$
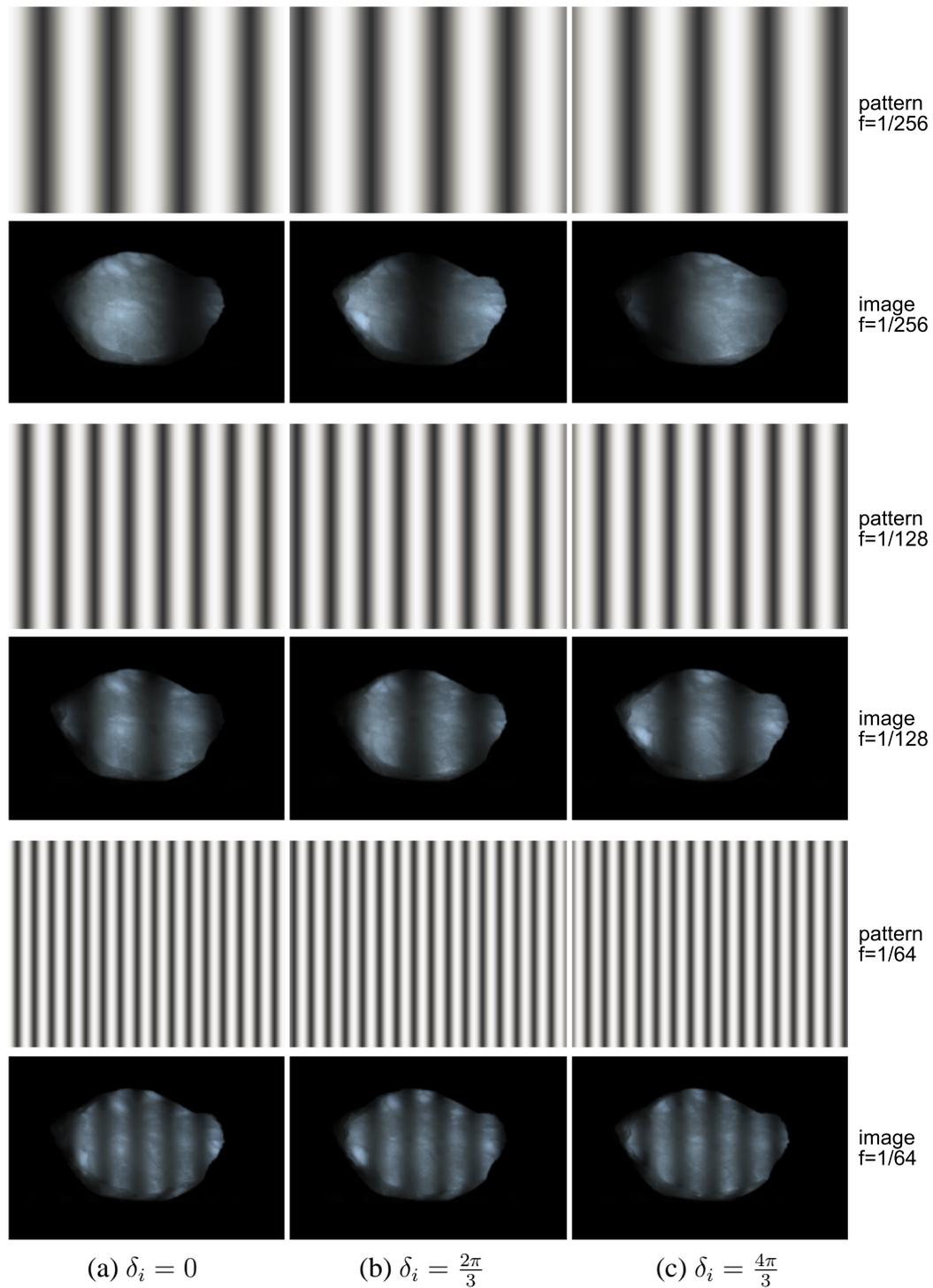
Figure 5.3: *Some of the projected patterns and captured images during phase-shifting. For each frequency, multiple phase-shifted patterns are projected.*

Using phase-shifting for 3D scanning, one is mostly interested in determining the effective phase $\Phi(x, y)$ which is correlated to the disparity and therefore the depth of the pixel. Additional phase-unwrapping employing lower frequency patterns is necessary to locate the period for which the phase is measured.

Figure 5.3 shows some of the projected patterns and captured images during phase-shifting. For each frequency, $f = \frac{1}{256}, \frac{1}{128}, \frac{1}{64}$, multiple phase-shifted patterns, $\delta_i = 0, \frac{2\pi}{3}, \frac{4\pi}{3}$ are projected.

5.3 Separation Properties of Phase-Shifting

Following the idea presented in [Nayar06] one can exploit the different dependencies of the direct and global illumination effects on a high frequency illumination pattern to separate these two components. Only the direct component L_d is correlated to the amplitude term I'' while the global component L_g is not if and only if f is sufficiently high. Both L_d and L_g contribute to the average image I', or a_0. We compute L_d and L_g as:

$$L_d = 2\sqrt{a_1^2 + a_2^2}, \quad \text{and}$$

$$L_g = 2a_0 - 2\sqrt{a_1^2 + a_2^2}. \tag{5.8}$$

Since there are three unknowns, a_0, a_1, and a_2, theoretically, three shifts are the minimum requirement for calculating the phase and the separation. In our experiments we apply either 6 or 8 shifts in order to alleviate distortions due to imaging noise, nonlinearity, vibration, etc.

In the minimum case, there are only three different phase shifts, i.e. $N = 3$, $\delta_i \in \{-2\pi/3, 0, 2\pi/3\}$, one can separate the global and the direct components as:

$$L_d = \frac{2}{3}\sqrt{3(I_0 - I_2)^2 + (2I_1 - I_0 - I_2)^2} \quad \text{and} \tag{5.9}$$

$$L_g = \frac{2}{3}(I_0 + I_1 + I_2) - L_d. \tag{5.10}$$

In [Nayar06], Nayar et al. pointed out that the best separation of direct and global components is achieved with the highest frequency illumination patterns. In other words, the performance of the separation is limited by the frequency of the illumination. This highest frequency is however given by the properties of the projector and might be reduced due to light leakage and built-in image processing. In practice we apply $f = \frac{1}{6}$ as the overall highest frequency, and found $f = \frac{1}{8}$ with $N = 8$ to produce the best phase estimates. The unit for frequency is $cycles/pixel$. While there are practical limits to the highest reliable

frequency the phase-shifting pattern furthermore is only of high frequency in the phase direction, here the x direction. It does not contain any high frequencies in the y direction and therefore has only limited descattering performance.

5.4   Temporal Phase-Unwrapping

Using shifted patterns with a single frequency we can detect the phase within one period of the selected frequency ($\Phi \in [0, 2\pi]$). The period however might be repeated multiple times over the entire scene. The problem is to locate the absolute unwrapped phase Ψ that uniquely identifies the pixel's phase. A number of different methods have been proposed to obtain an unwrapped phase map [Ghiglia98]. If the scene contains depth discontinuities the exact phase and period can be obtained by repeating the phase extraction for multiple (lower) frequencies [Huntley93, Huntley97a, Huntley97b]. Possible approaches are choosing frequencies such that the greatest common divisor of the periods is larger than the number of columns in the projector image [Tarini05]. More robust unwrapping is obtained by creating a series of frequencies $\lambda_j = 0.5\lambda_{j-1}$ until one period spans the projector image width s resulting roughly in $F = \log_2(s)$ frequencies [Huntley97a]. Given the unwrapped phase at one frequency $j+1$, the unwrapping algorithm iteratively locates the phase at step j, the next higher frequency. Starting with $j = F - 1$ and $\Psi_F = \Phi_F$ one computes the unwrapped phase at the next higher frequency by

$$\Psi_j = \Phi_j - 2\pi\text{NINT}\left(\frac{\Phi_j - 2\Psi_{j+1}}{2\pi}\right),\tag{5.11}$$

where NINT rounds to the nearest integer. The unwrapping itself is to some extent similar to decoding binary encoded structured light patterns [Salvi04], but more robust.

Figure 5.4 shows the temporal unwrapping process. In our experiments we use frequencies of $\frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \frac{1}{64}, \frac{1}{128}, \frac{1}{256}, \frac{1}{512}$ and $\frac{1}{1024}$ *cycles/pixel*. At the higher frequencies ($f = \frac{1}{8}$ and $f = \frac{1}{16}$) we use 8 and 16 phase-shifts to obtain the best quality while the lower frequencies are only used for disambiguating the period (making rough and stable binary decisions), therefore six phase shifts turned out to be sufficient. Overall, 60 images are captured for each range scan but the number could be further reduced if necessary. The unwrapping process goes from the lowest frequency, $f = \frac{1}{1024}$ to the highest frequency, $f = \frac{1}{8}$. The phase map at the lowest frequency is the absolute phase map. The phase map at higher frequency is unwrapped using the absolute phase map at lower frequency.

Using multiple frequencies poses the problem that Equation 5.2 only holds for high frequencies. For low frequencies, the global component will also vary with
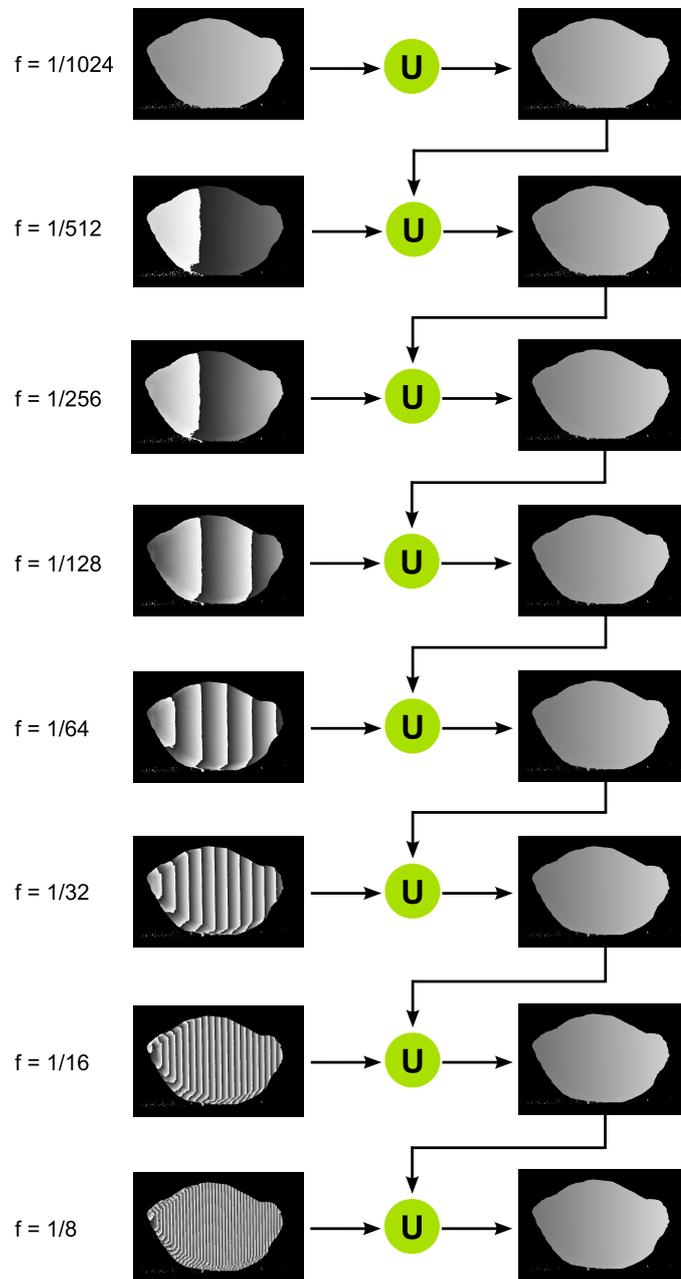
Figure 5.4: *Temporal phase-unwrapping. The phase maps on the left column are those to be unwrapped. The right column shows the unwrapped versions. For the lowest frequency, $f = \frac{1}{1024}$, the unwrapped is the same as the wrapped. For the higher frequencies, the unwrapping is done by using the information of the last unwrapped phase map at lower frequency. The final unwrapped phase map at frequency $f = \frac{1}{8}$ is used for 3D reconstruction.*

the phase-shift and thus the phase and depth estimates will be biased. We demonstrate this effect in Figure 5.5(a) by comparing the phase estimates for different frequencies. The lower the frequency, the larger the deviation of the estimated phase for the individual frequency. A small or moderate drift at a low frequency has typically only very little effect on the combined result since the lower frequencies are just used for estimating the 2π modulo jumps. Figure 5.6(c) shows an example where the deviation on a lower frequency is larger than one period and thus introduces a major offset in the 3D scan.

In the next section we will demonstrate how reliable depth profiles can be computed even for low frequencies when polarization is used in addition to phase-shifting to separate out the global component (Figure 5.5(b) and Figure 5.6(d)).

5.5   Polarization Difference Imaging

Most natural and artificial light sources produce *unpolarized* (or *depolarized, nonpolarized*) light waves whose electric field vectors vibrate in all planes that are perpendicular with respect to the direction of propagation. If the electric field vectors are restricted to a single plane by filtration of the beam with specialized materials, such as long-chain molecules oriented in a single direction, then the light is referred to as *linearly polarized*. When two polarizers are crossed, their transmission axes are oriented perpendicular to each other and light passing through the first polarizer is completely blocked by the second polarizer, which is typically termed as *analyzer*. The *extinction factor* describes how much light is extinguished through a pair of crossed polarizers. Quantitatively, the extinction factor is determined by the ratio of light that is passed by a pair of polarizers when their transmission axes are oriented parallel versus the amount passed when they are positioned perpendicular to each other. A typical extinction factor ranges from 10,000 to 100,000. More generally, the amount of light passing through a pair of polarizers can be quantitatively described by applying Malus' cosine-squared law, as a function of the angles between the polarizer transmission axes, using the equation:

$$L^{'} = L{\cdot}cos^2\theta. \tag{5.12}$$

$L^{'}$ is the amount of light passing through the analyzer and the total amount of light passed through the pair of crossed polarizers. L is the amount of light that is incident upon the polarizer. θ is the angle between the transmission axes of the polarizer and analyzer. By examining the equation, it can be determined that when the two polarizers are crossed ($\theta = 90$), the amount of passed light $L^{'}$ is zero. When the polarizers are partially crossed at $\theta = 30$ and $\theta = 60$ degrees, the light transmitted by the analyzer is reduced by 25 percent and 75 percent respectively.

(a)

(b)
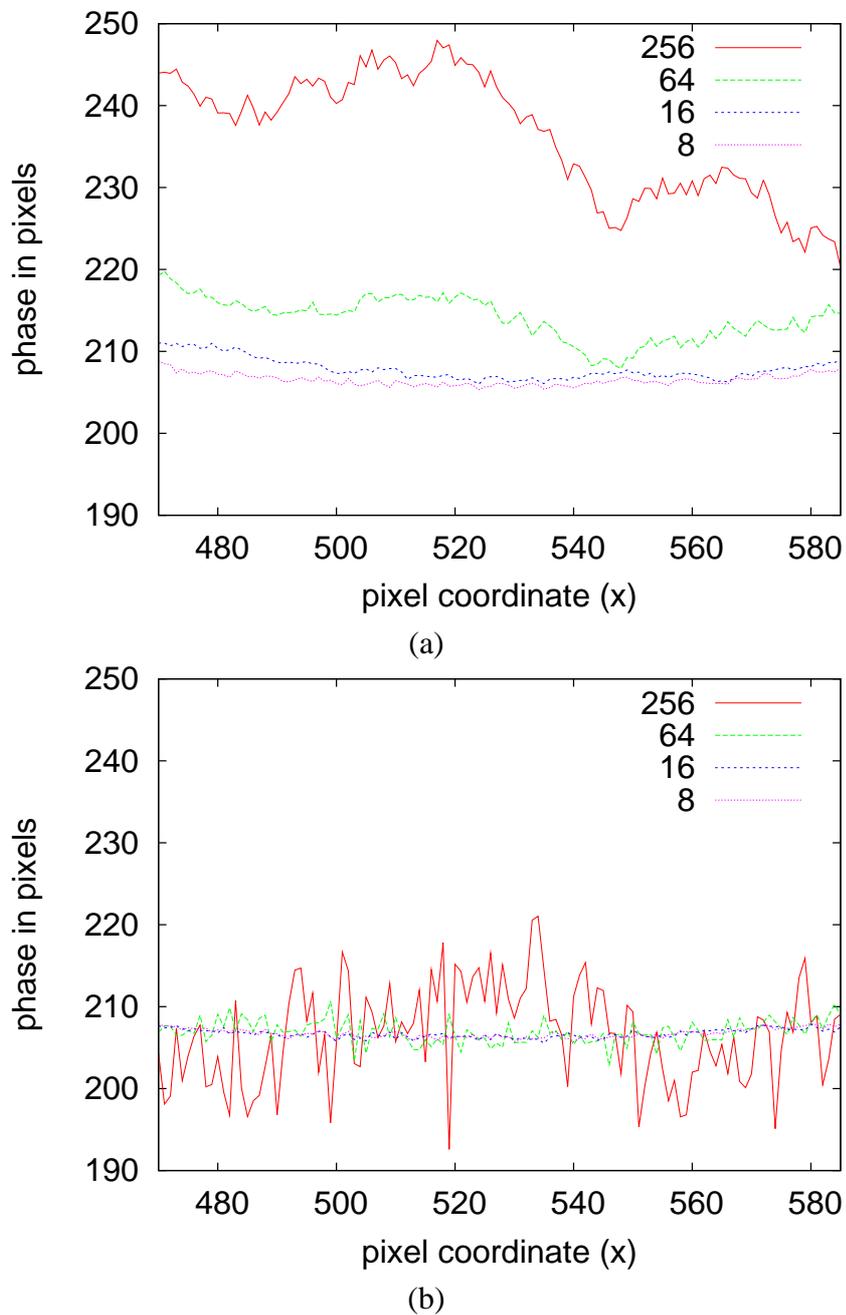
Figure 5.5: *Biased phase reconstruction for low frequency patterns. (a) Phase profiles of individual frequencies for one line on the planar alabaster block. (b) After polarization difference imaging even the lower frequencies result in correct depth estimates. Note that the curves are tilted by the same factor for illustration.*

(a) L_d for $f = \frac{1}{8}$ (b) L_d for $f = \frac{1}{64}$

(c) Phase-shifting   (d) PDI
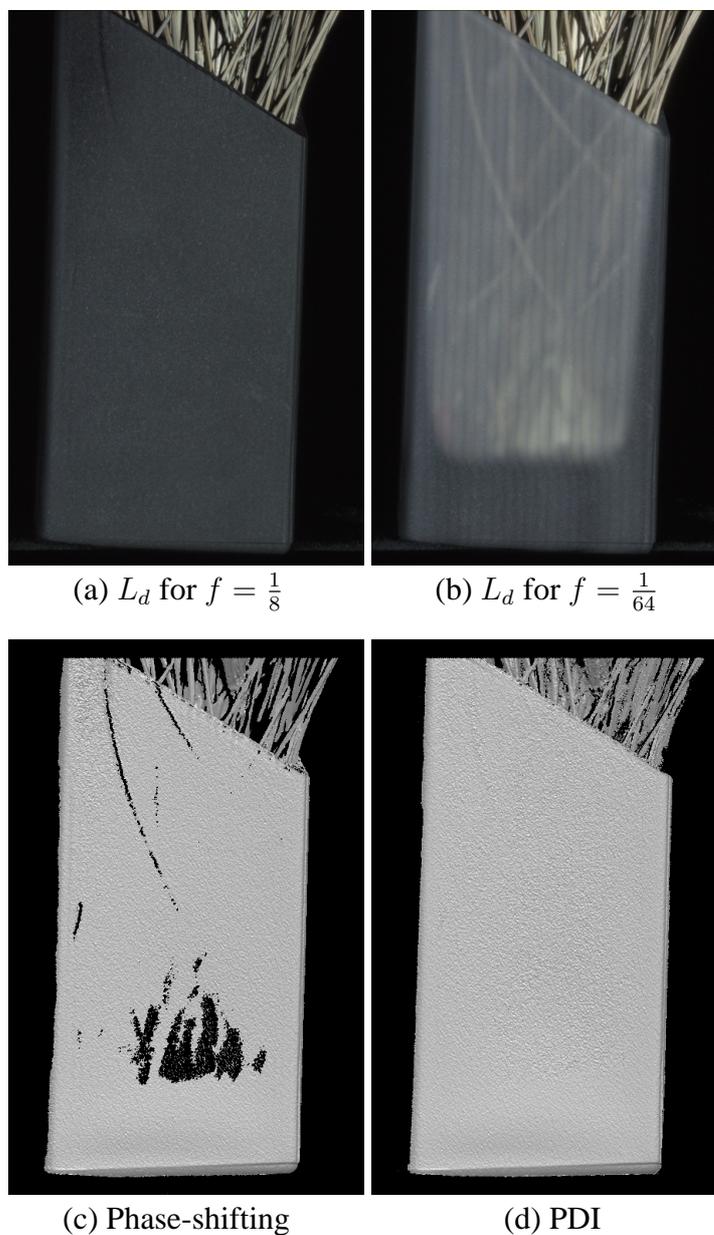
Figure 5.6: *For this translucent vase filled with lavender, the reconstructed direct reflection is dependent on the frequency of the illumination pattern. (a) Most subsurface scattering is removed using the highest frequency. (b) At lower frequencies structures beneath the surface contribute to the direct component polluting the phase-unwrapping results in (c). (d) Using PDI the influence of subsurface structures is largely reduced and the desired shape is captured.*
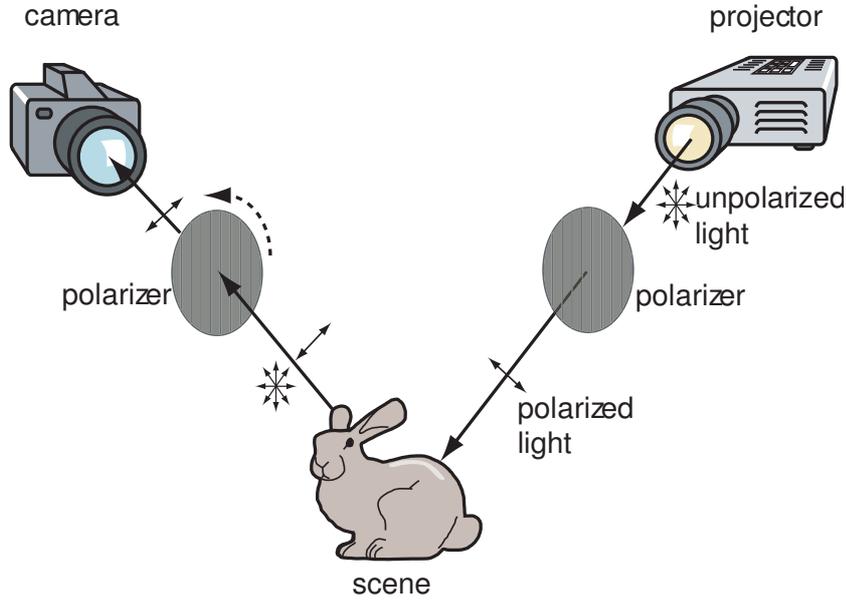
Figure 5.7: *System setup for polarization difference imaging (PDI). The projector is equipped with a linear polarization filter at fixed orientation. The camera captures two image sequences with parallel and with perpendicular orientation of the polarization filters.*

On the one hand polarized light becomes unpolarized after diffuse reflection or multiple scattering. On the other hand, initially unpolarized light reflected off of a smooth surface at an oblique angle will have a transmitted radiance through a polarizer that oscillates sinusoidally as a function of polarizer angular orientation between a maximum I_{max} and a minimum I_{min}. A quantitative measure of the proportion of how much initially unpolarized light becomes linearly polarized on reflection, also called the *degree of polarization (DoP)* is given by

$$\rho = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}. \tag{5.13}$$

ρ varies between 0 and 1 inclusive [Born99]. At $\rho = 0$, reflected light is unpolarized, as is generally true for the diffuse component of reflection. At $\rho = 1$, reflected light is completely linearly polarized as predicted for pure specular reflection at the *Brewster angle* for a dielectric surface. At this angle, only the polarization component perpendicular to the specular plane is reflected.

As discussed in Chapter 3 multiple scattering depolarizes the incoming light. Schechner et al. [Schechner03, Schechner04, Schechner05, Treibitz06] have made extensive use of this phenomenon to compute clear pictures through haze or murky water by taking several polarized images from which the depolarized part can be

removed afterwards. Based on the estimated signal loss induced by the participating media the authors further compute approximate depth maps of the underlying scene.

The Polarization Difference Imaging (PDI) [Lythgoe79] is a bio-inspired technique originally developed for optical imaging which provides significantly enhancements in target detection and feature extraction over the conventional methods. In our setup depicted in Figure 5.7 linear polarizers are put in front of the camera and the projector. We then capture the phase-shift image sequence twice, once when the camera's polarizer axis is oriented parallel to the projector's polarizer axis, yielding $I_j^{\parallel}$, and a second time using cross-polarization, $I_j^{\perp}$. A polarization difference image [Rowe95, Tyo96] is then computed as

$$I_j^{\triangle} = |I_j^{\parallel} - I_j^{\perp}|. \tag{5.14}$$

The idea is that depolarized light will add exactly the same contribution to both image sequences, independent of the camera's filter orientation, and thus will be completely removed in $I_j^{\triangle}$. Relying on $I_j^{\triangle}$ we then perform the 3D reconstruction.

Figure 5.8 shows visually the effectiveness of polarization upon the separation of direct/global components. While the direct component extracted using ordinary phase-shifting even at high frequency contains some subsurface scattering from underlying structures, they are partially removed by parallel polarization and almost completely eliminated after applying PDI. After using polarization, the contrast of the projected pattern can be greatly improved (see Figure 5.9). The sine patterns become clearer after employing parallel polarization and are further improved after using PDI. Figure 5.10 demonstrates the effect of polarization filtering on the quality of the 3D reconstruction of a quite planar block of alabaster. The most important difference between 3D reconstruction by phase-shifting without polarization filtering and with polarization difference imaging applied is that scattering events beneath the surface are much better removed in the latter case. However, as will be discussed in the result section, there are some scenes where the PDI approach filters out too much of the direct reflection. In these cases using the image sequence with parallel orientation of the polarization filters provides a good trade-off between no polarization and PDI.

5.6 Acquisition System

Figure 5.7 shows schematically the system setup for polarization enhanced phase-shifting. All images have been acquired with a 14-bit 1360×1024-pixel Jenoptik ProgRes CFcool CCD camera and a Mitsubishi XD490U XGA DLP Projector whose native resolution is 1024×768. We performed a photometric calibration

(a) Photograph

(b) L_d from Phase-shifting

(c) L_d from parallel polarization

(d) L_d from PDI
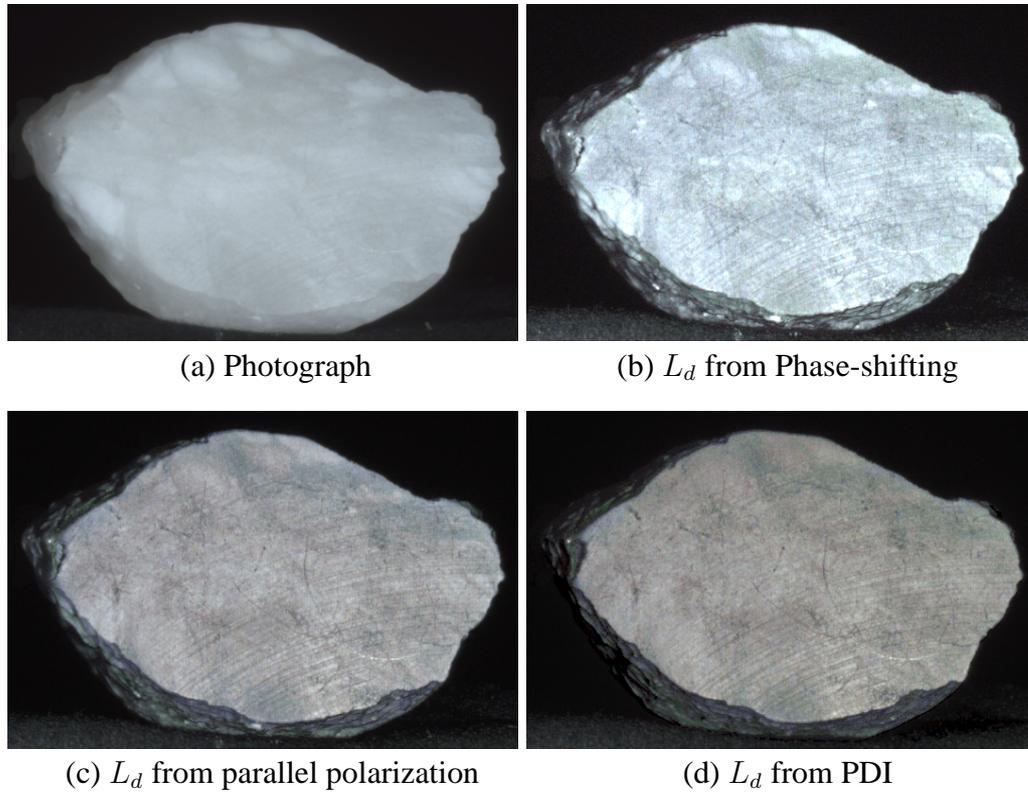
Figure 5.8: *While the direct component L_d (b) extracted using no polarization filters clearly contains some subsurface structures they are partially removed by parallel polarization $L_d^{\parallel}$ (c) and not present after applying PDI, $L_d^{\triangle}$ (d).*

(a) Phase-shifting

(b) Parallel polarization

(c) PDI
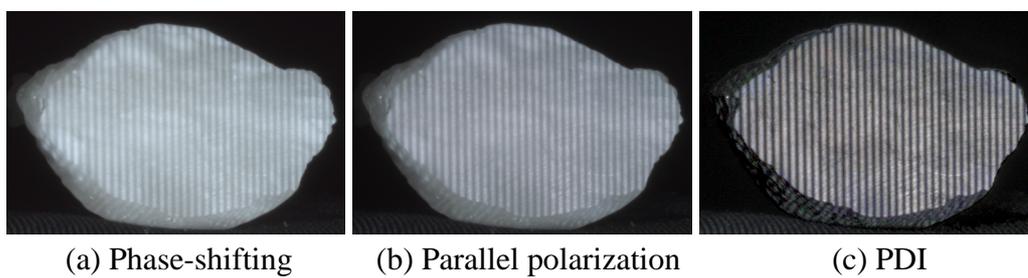
Figure 5.9: *The contrast in the high frequency input images is improved by parallel polarization and further by PDI.*
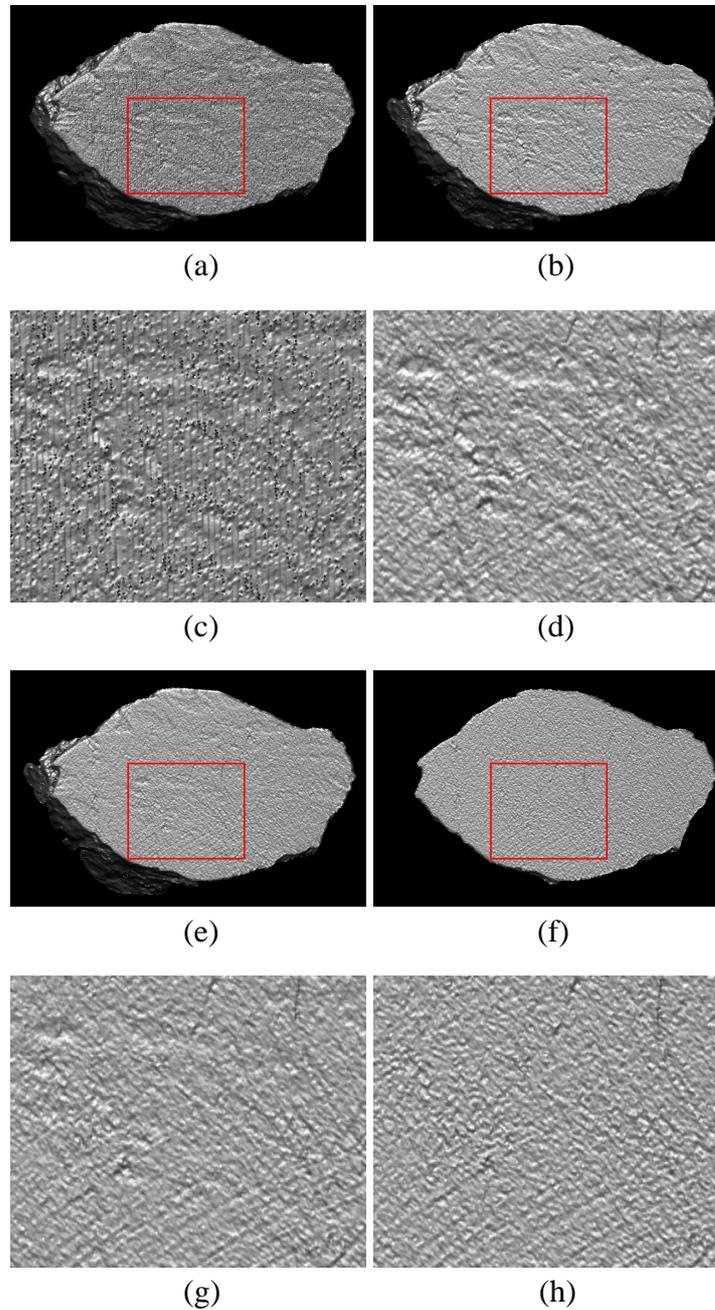
Figure 5.10: *Reconstruction results for a planar surface of heterogeneous alabaster. 3D reconstruction from (a) line sweeping, (b) phase-shifting without polarization, (e) with parallel polarization, and (f) with PDI. The influence of the subsurface structures on the final 3D geometry has been completely removed by PDI. (c), (d), (g) and (h) are the zoom-ins of (a), (b), (e) and (f) respectively.*

for both devices and captured HDR images [Robertson03] using four different exposures. The measured maximum simultaneous contrast of a sine pattern with a period of 8 pixels reflected by a gray card is 180:1 (max/min). We performed geometric calibration between the camera and the projector [Zhang00]. Linear polarization filters have been placed in front of the projector and the camera to acquire the PDI image sequences.

During the acquisition time, the projector generates a sequence of phase-shifting pattern at different frequencies. The camera takes simultaneously images of the interesting scene. For ordinary phase-shifting, no polarizer will be used. For parallel polarization imaging, the two polarizers, in front of the camera and the projector respectively, are set with parallel polarization orientations. Then we turn the polarizer in front of the camera, the analyzer, 90 degrees to make the polarization orientations of the two polarizer crossed. The difference between the images taken under parallel polarization and cross polarization will be the input for PDI based phase-shifting.

5.7   Calibration

In this section, we focus on the geometric calibration of the projector-camera system (see Figure 5.11). A planar checkerboard with blue/white checkers (see Figure 5.11(a)) is used for calibrating both the camera and the projector.

We take two sets of images, one for the camera, and the other for the projector. For each pose of the planar checkerboard with blue and white checkers, we first take an image with a desktop light (see Figure 5.11(c)). Then turn off the desktop light and project a R/BR (Color is Red=255, Green=0, and Blue=0. Color BR is Blue=255, Red=255, and Green=0) checkerboard pattern onto the planar checkerboard (see Figure 5.11(d)). Figure 5.11(e) is the blue channel of Figure 5.11(d). Figure 5.11(f) is the color invert of Figure 5.11(e). After taking these two images, we change the pose of the planar checkerboard, and repeat the procedure for K ($K > 8$) times and get two sets of calibration images. The camera-set consists of K images, shown in Figure 5.11(c). The projector-set consists of K images, see Figure 5.11(f). We perform camera calibration on the camera-set using Zhang's method [Zhang00] and obtain both the intrinsic and extrinsic parameters. Since the projected pattern share the same plane with the printed pattern, we can compute the positions of the checker corners of the projector-set images using homographies. With the known checker corner positions, we can calibrate the projector both intrinsically and extrinsically by using the same calibration algorithm of the camera.

The calibration, including the intrinsic and extrinsic parameters of both the projector and the camera, will be used for triangulation to get the 3D coordinates
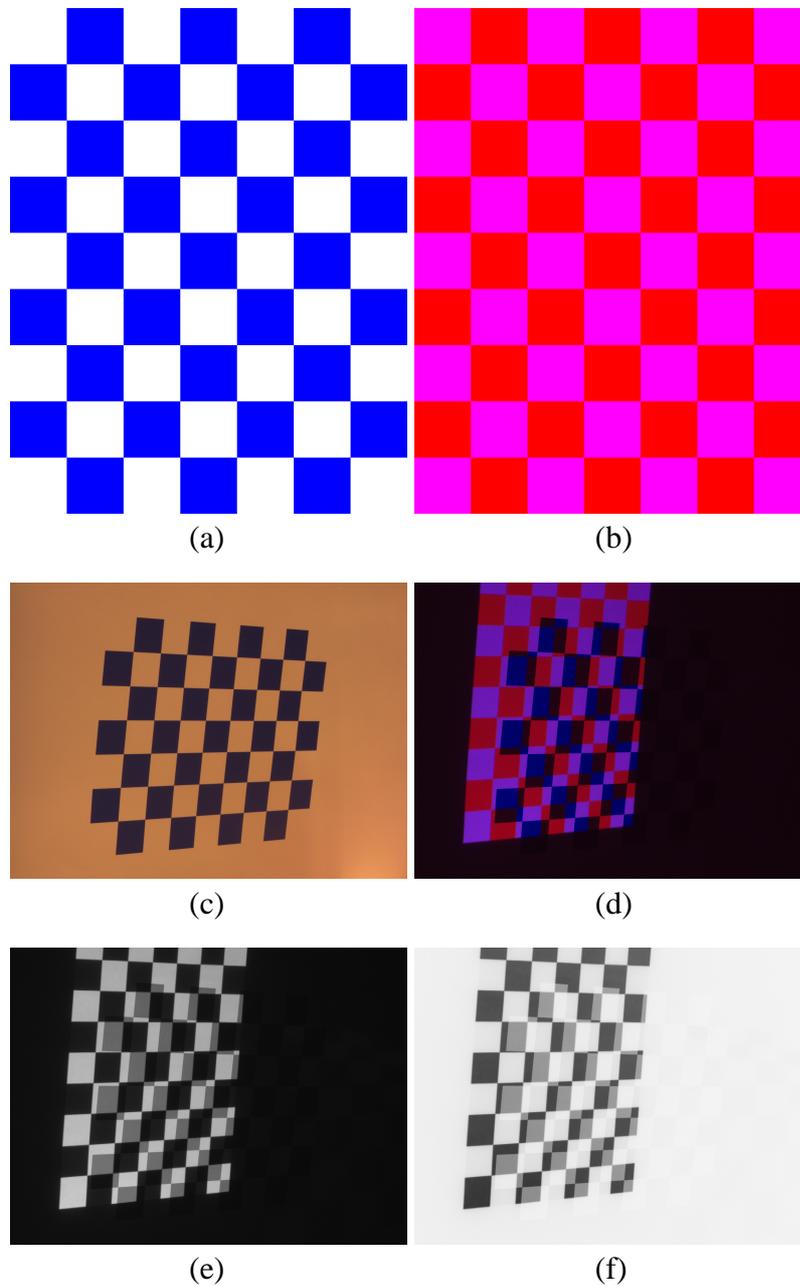
Figure 5.11: *Projector-camera calibration. (a) Blue-white checkerboard pattern, printed on a planar surface. (b) R/BR checkerboard, where one checker is red and one checker is a mixture of red and blue. (c) Image taken for the planar blue-white checkerboard. (d) Image taken under (b) being projected onto the planar checkerboard. (e) Blue channel of (d). (f) Color inversed version of (c). (c) and (f) will be used for corner detection.*

(a) Photograph

(b) 3D from phase-shifting

(c) 3D from parallel polarization

(d) 3D from PDI
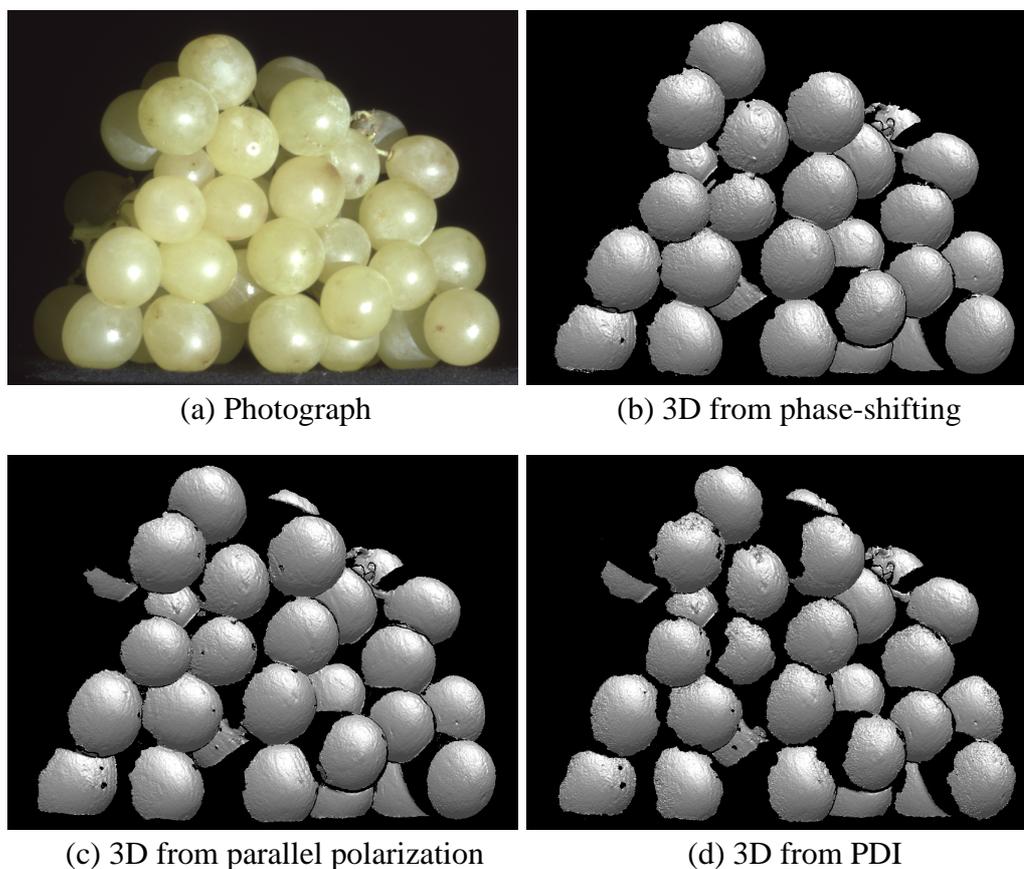
Figure 5.12*: 3D reconstructions of grapes using different methods. The reconstruction results of phase-shifting without polarization (b) and with parallel polarization (c) are of comparable quality since the objects are mostly homogeneous. At grazing angles too much signal from the direct reflection is filtered out by PDI (d) resulting in more holes and noisier depth maps*

as soon as we have the pixel-to-pixel correspondences between the camera and the projector. Building this correspondence is the main part in the phase-shifting based 3D scanning method, where the correspondence is encoded in the absolute phase map.

5.8   Results

In the following section we assess the descattering capabilities of phase-shifting with and without polarization on a set of translucent scenes: a highly translucent, almost homogeneous alabaster figurine (Figure 5.14 and Figure 5.15), a

(a) Photograph

(b) 3D from phase-shifting

(c) 3D from parallel polarization

(d) 3D from PDI
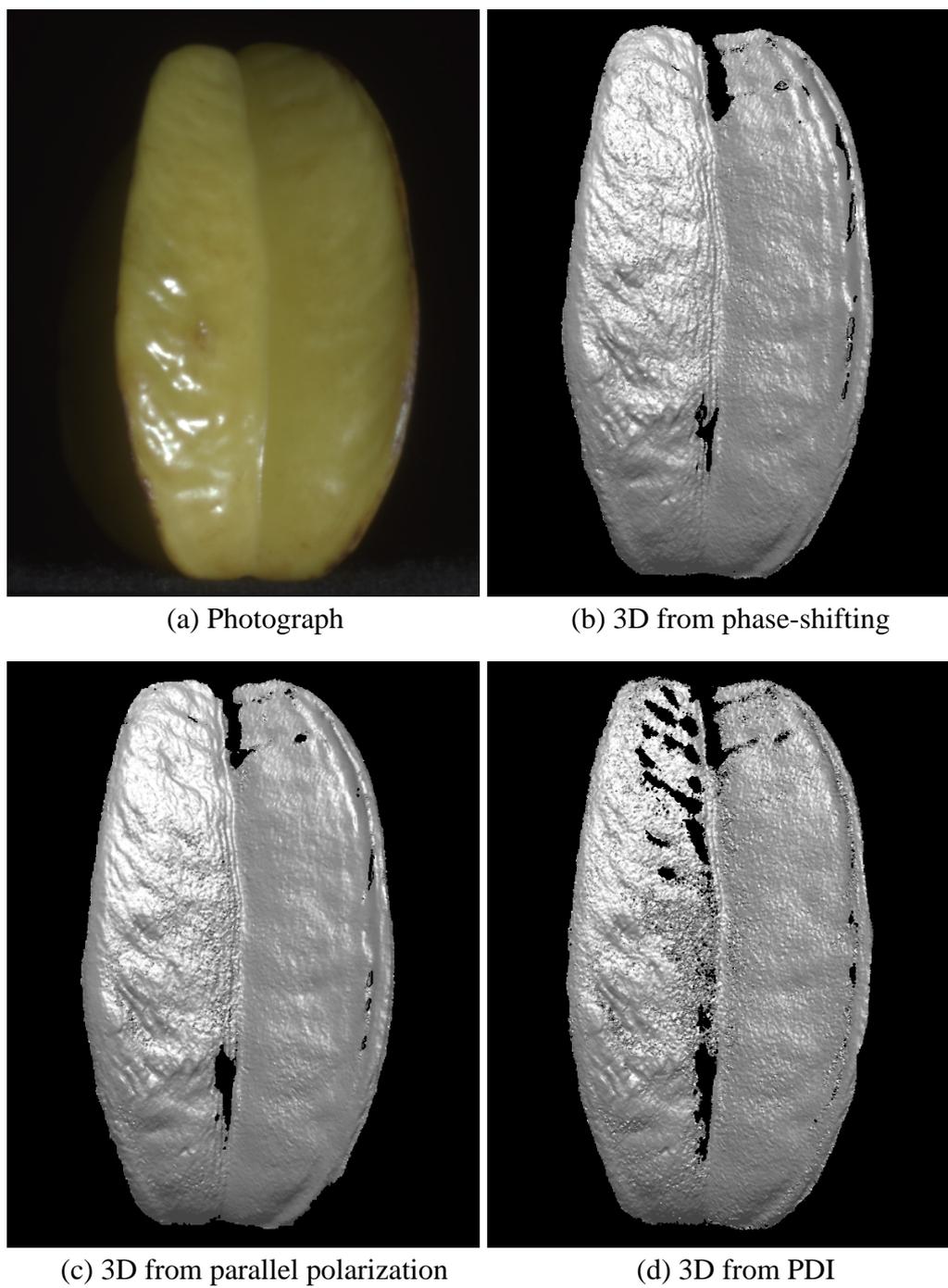
Figure 5.13: *3D reconstructions of a starfruit using different methods.*

filled, translucent vase (Figure 5.6), a heterogeneous planar slab of alabaster (Figure 5.10), some grapes and a starfruit (Figure 5.13). Except for removing spurious background pixels and pixels having a too weak signal no further processing, i.e., noise removal or smoothing has been applied to the results.

5.8.1 Structured Light

Descattering based on phase-shifting without polarization can deal pretty well with translucent objects and clearly removes some amount of the subsurface scattering, as predicted by Nayar et al. [Nayar06]. The phase-unwrapping, however, relies on low frequency patterns which clearly suffer from global effects (see Figure 5.5 and Figure 5.6). Furthermore, structures beneath but close to the surface will have some influence on the estimated direct component which is unwanted in the context of 3D scanning (Figure 5.6 and Figure 5.10).

Sweeping a single line is an alternative to phase-shifting and performs surprisingly similar. As can be seen in Figure 5.10 even for highly translucent objects one obtains a reasonable 3D scan if high quality equipment and HDR sequences are used. The noise in the figure indicates that the SNR of line sweeping compared to phase-shifting is considerably lower. Although line sweeping is still sensitive to the bias introduced by subsurface scattering or subsurface structures global effects are minimized by the comparably small amount of incident light concentrated on a small region.

5.8.2 Polarization

Polarization difference imaging also separates the direct from the global component very well. It faithfully removes all traces of subsurface structures. At grazing angles PDI however filters out too much of the direct reflection (see Figures 5.12, 5.13, 5.14, and 5.15). It is worthwhile to note that depending on the surface properties also some fraction of the direct reflection might be depolarized. This fraction will also be removed in the polarization difference image. For some scenes, we actually observed a better contrast of direct vs. global reflection in the parallel polarization setting $I_j^{\parallel}$ producing smoother 3D scans (see Figures 5.12, 5.13, 5.14, and 5.15). Even though parallel polarization in theory only removes some fraction of multiple scattering effects (compare images in Figure 5.9), combining it with phase-shifting adds the descattering capabilities of both techniques. Figure 5.10 further shows that parallel polarization also renders phase-shifting slightly more robust against subsurface structures, though not as robust as PDI. On the other hand parallel polarization is much easier to acquire since it requires only half the amount of images and a fixed orientation of the filters.

Photograph   3D from phase-shifting

Figure 5.14*: 3D reconstructions of an alabaster figurine of Venus.*

5.9   Limitations

The ordinary N-step phase-shifting is only high frequency in one dimension, the x direction in our case. A two dimensional high frequency pattern will perform better in terms of descattering and further 3D reconstruction. Furthermore, the projector can only generate accurate patterns at limited frequencies because of

3D from parallel polarization   3D from PDI

Figure 5.15: *3D reconstructions of an alabaster figurine of Venus.*

the optics and internal image processing of the projector. The limits of the frequency of the projected pattern constrain the ability of using phase-shifting for 3D scanning.

Polarization enhanced methods, parallel polarization imaging or polarization difference imaging, can alleviate the problem inherent in the ordinary phase-shifting method by separating the direct reflection components. However, the po-

larization based methods also kill direct reflection information when filtering the global reflection components, which lowers significantly the signal-noise-ratio. Furthermore, polarization is heavily depending on the surface orientation. It varies dramatically with the changing surface normal. This limitation cause the polarization based method's failure during scanning highly curved surface or scene with strong interreflection caused by multiple specular reflections.

For both temporal-unwrapping based N-step phase-shifting and polarization enhanced phase-shifting, there is a speed problem. In order to do high quality scanning, usually dozens of HDR images are required. Hence, all the methods presented in this chapter are only designed for static objects or scenes.

Chapter 6

Modulated Phase-Shifting for 3D Scanning

In structured light based 3D scanning, the 3D information is obtained by triangulation after establishing the correspondence between the camera pixels and the sub-pixel locations of the illuminating projector pixels. In order to establish this correspondence a large variety of patterns have been proposed [Salvi04], encoding the projector pixel locations. The decoding algorithms then assume that the measured camera pixels only locally depend on the projected patterns. This assumption is violated if global illumination effects such as interreflections or subsurface scattering are strong, leading to the observations of multiple overlaid signals. Most often these global illumination effects dampen high frequencies in the input patterns.

By choosing appropriate high frequency patterns, Nayar et al. [Nayar06] have demonstrated how global illumination effects can be separated from the direct illumination. In Chapter 5 we exploited this method for 3D scanning by using phase-shifting where the use of high-frequency sinusoid patterns drastically reduced the influence of the global illumination. The employed one-dimensional patterns (sinusoidal in x but constant in y), however, do not fully remove the global effects. It gets even worse for the lower frequency patterns that are required in temporal phase-unwrapping [Huntley93, Huntley97a, Huntley97b]. For these reasons, our method was combined with polarization difference imaging (PDI), exploiting the fact that multiply scattered light becomes depolarized. PDI on the other hand adds to the complexity of the setup and the acquisition. Furthermore, the combined effect of polarized illumination and the polarization characteristics of reflections at surfaces leads to a drastically weakened signal, especially at grazing angles [Born99].

In this chapter we introduce modulation based separation, where the direct

(a)   (b)
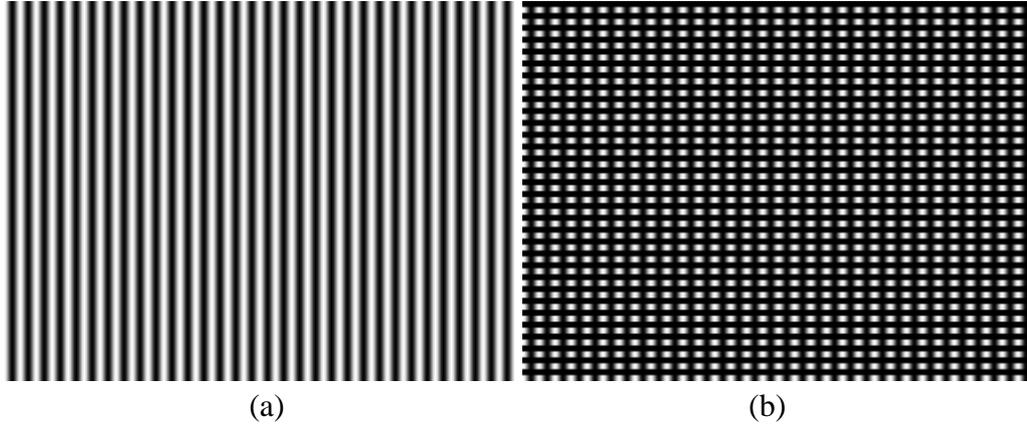
Figure 6.1*: 1D modulation example. (a) Sine pattern in x direction. (b) Modulation in y direction with binary modulation function.*

illumination component even of a low-frequency projected signal is correctly estimated. On top of the low-frequency signal we multiply a two-dimensional high frequency signal. We analyze the separation performance using high-frequency patterns in 1D or 2D, demonstrating that 2D separation patterns result in a much clearer separation. We further show that due to the non-linearity in the separation analysis, a sequential analysis of a 2D pattern once horizontally and once vertically produces an even better separation. Moreover, the modulation and separation scheme can be exploited for general low frequency patterns, such as an ordinary photograph.

The improved separation based on modulation with sequential analysis leads to a much better performance of 3D scanning. The proposed algorithm is more efficient with regard to subsurface scattering as well as interreflections than previous methods.

6.1 Modulation

In telecommunication, modulation is the process of varying a periodic waveform in order to carry another signal. Usually, a high-frequency sinusoid waveform is used as carrier signal. We perform modulation by multiplying an input signal L (sine waves for phase-shifting, see Eq. 5.1) with other shifted high frequency patterns. The process can be formally expressed as

$$\tilde{L}_{ij}(x,y) = L(x,y)M_x^i(x,y)M_y^j(x,y), \tag{6.1}$$

where $\tilde{L}$ is the modulated function to be projected. M_x is the modulation function in x dimension with frequency f_x and N_x shifts. Accordingly, M_y is a modulation function in the y dimension with frequency f_y and N_y shifts. The total number of projected images is $N_x N_y$. The highest frequency of the modulated signal is now given by either f_x or f_y and is no longer limited by the original signal's frequency since a multiplication in the spatial domain corresponds to a convolution in the frequency domain. Thus the modulated function $\tilde{L}$ has higher frequency than the non-modulated function L in both dimensions, and we can overall obtain a better separation of the direct/global components.

In principal, all the high frequency functions proposed in [Nayar06] for the separation can be used as modulation functions, M_x and M_y. We use either periodic binary functions or sinusoid functions as the modulation functions in all our experiments. The sinusoid modulation function is defined in the same way as L (see Eq. 5.1). The periodic binary function in x dimension is defined as

$$\mathcal{B}_x(x,y) = \lfloor 2xf \rfloor \bmod 2. \tag{6.2}$$

Figure 6.1 shows an 1D modulation example. The sine pattern in x direction is modulated in y direction with binary modulation function. The modulated pattern is of high frequency in both x and y direction and thereby possesses better separation property.

6.2 Multi-Pass Separation

In principle, a modulation in one direction is sufficient to perform some (possibly limited) separation. In this case, the separation into the direct and global components can be done using sinusoid functions and the analysis given in Eq. 5.8. One can also use a binary pattern. The separation is then computed as follows:

$$\begin{aligned} L_d(x,y) &= I_{max}(x,y) - I_{min}(x,y), \quad \text{and} \\ L_g(x.y) &= I_{min}(x,y). \end{aligned} \tag{6.3}$$

where $I_{max}(x,y)$ and $I_{min}(x,y)$ are the maximum and minimum intensities of pixel (x,y) for the N_x images.

In order to increase the separation performance one should perform the modulation in both dimensions to remove global illumination in both directions. Applying the modulation in 2D, one can simply apply the minimum-maximum approach on the $N_x N_y$ images and then evaluate Eq. 6.3, as that has been proposed by Nayar et al. [Nayar06].

It turns out, however, that a much better separation can be achieved by using sequential separation in each dimension. One can see the N_y images for a

(a)   (b)

(c)   (d)
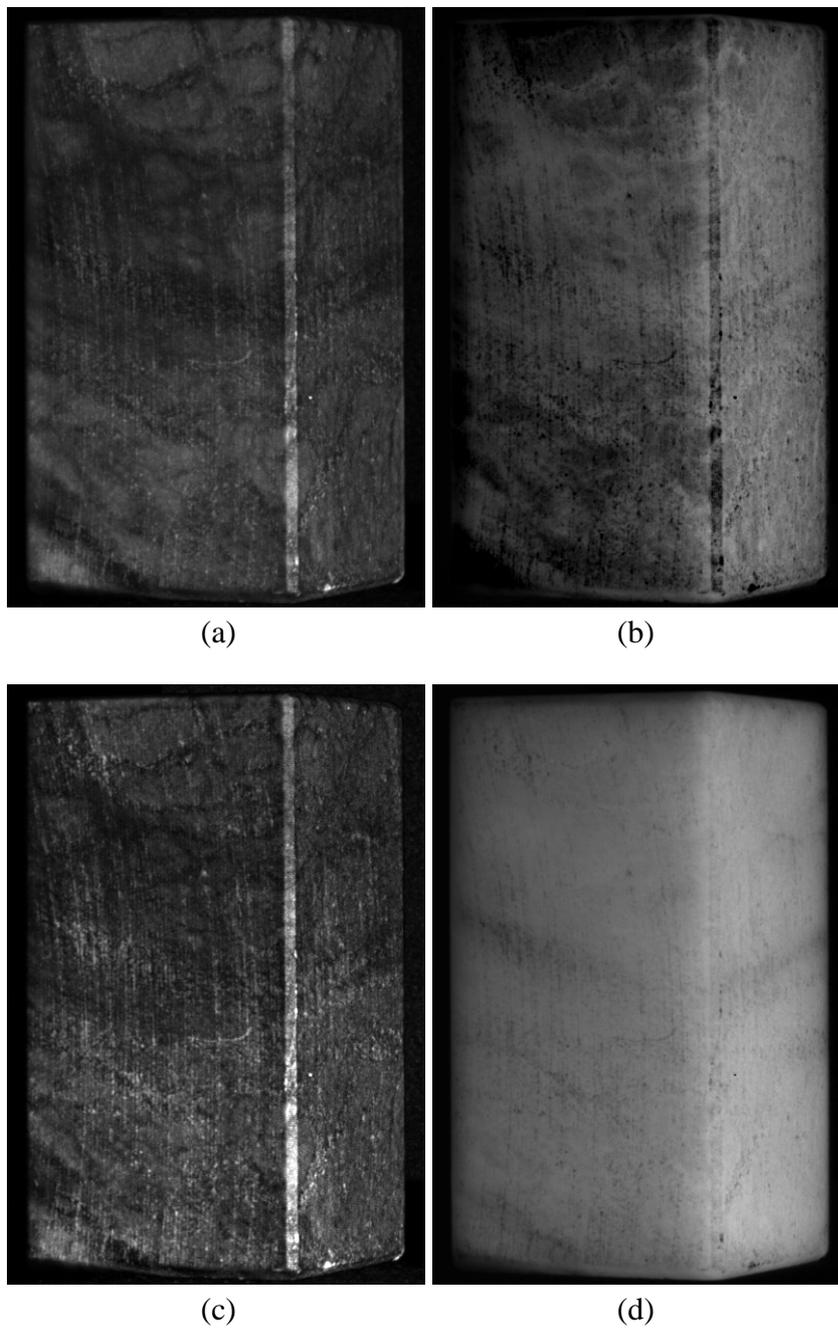
Figure 6.2: *Two-pass separation vs. one-pass 2D separation. Direct component (a) and global component (b) from the one-pass algorithm. Direct component (c) and global component (d) from the two-pass separation. Note, while working on exactly the same input data the two-pass separation method produces a much clearer separation.*
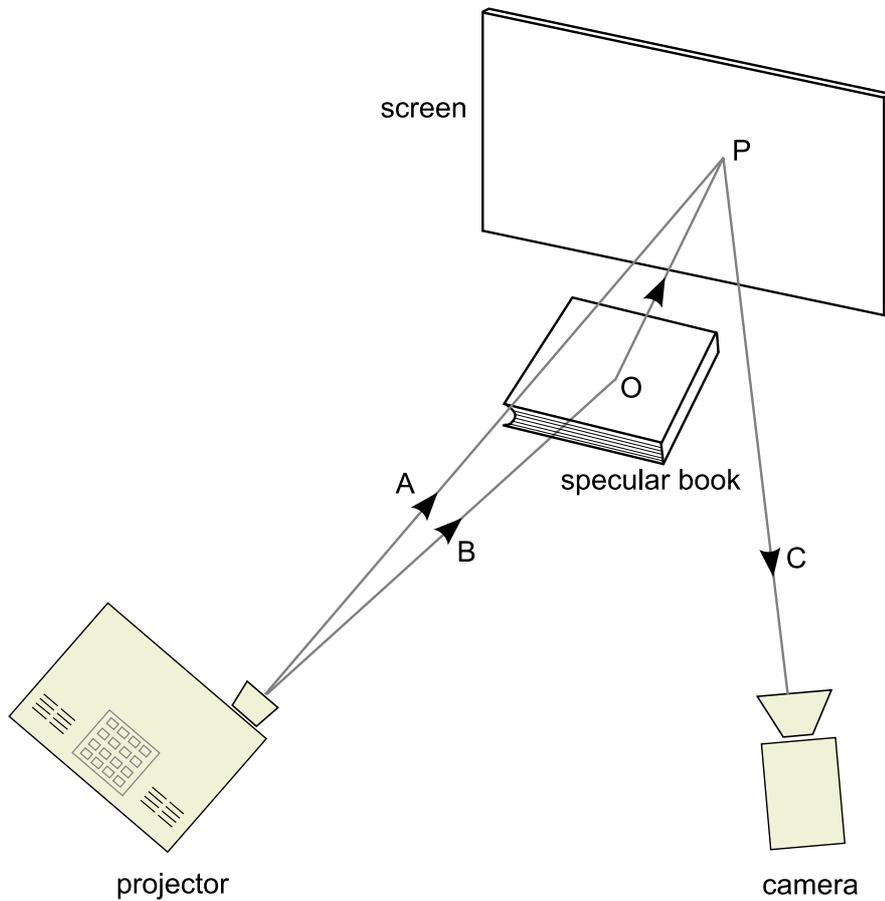
Figure 6.3*: Setup sketch. The projector projects an ordinary photograph onto a white diffuse paper screen. On the way of the projection, there is a specular book which reflects the projection onto the screen. Consider a point P on the screen, its radiance is composed of both the direct projection ray A and the reflected ray B, which hits the book at O and reflects. Therefore, the final camera ray C is a hybrid of both direct reflection and interreflection.*

fixed step i in x as the 1D modulation of the "original" signal M_x^i with the modulation pattern M_y. Performing the 1D separation for every individual step in x then results in N_x direct component images L_d^i which act as the input to the second separation phase, this time in x. Note that we have used exactly the same input images as before but have a two-stage filtering where global illumination effects that might pass the first separation are filtered out in the second pass. The improved performance of the two-pass separation vs. the one-pass 2D minimum-maximum separation is demonstrated in Figure 6.2.

(a)   (b)

(c)   (d)
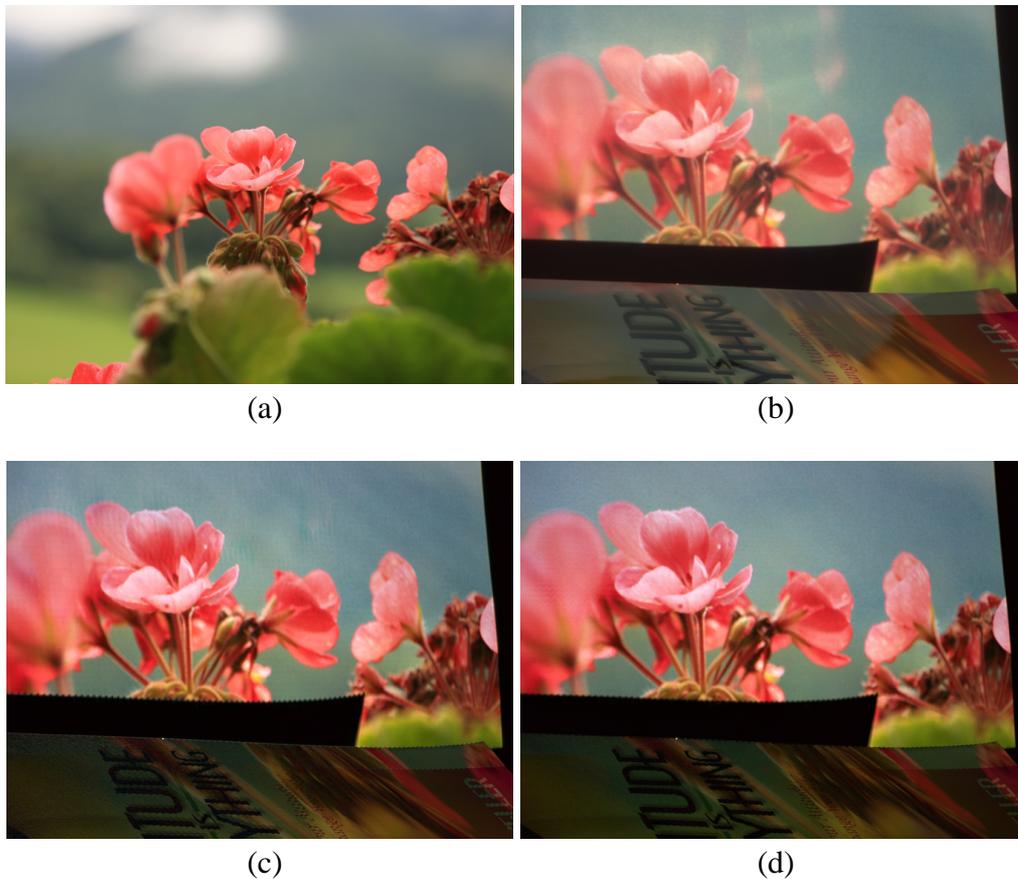
Figure 6.4: *An ordinary photograph (a) is projected into a scene containing a specular surface at the bottom and a diffuse paper screen in the back. (b) The projected image in the back is polluted by the ghost image caused by the reflection from the specular book cover. (c) The direct component separated using 1D phase-shifting, with $f_x = \frac{1}{8}$ still contains some artifacts. (d) The two-pass separation algorithm, with $f = \frac{1}{8}$ in both dimensions suppresses the ghost image completely. See Figure 6.3 for the setup sketch.*

6.3   Modulation for a General Signal

The idea of modulation based separation can be applied to more general signals. We demonstrate this by projecting an ordinary photograph into a scene inducing interreflections.

Figure 6.3 shows the schematic sketch of a scene with interreflections. The projector projects an ordinary photograph onto a white diffuse paper screen. On the way of the projection, there is a specular book which reflects the projection

onto the screen. Consider a point P on the screen, its radiance is composed of both the direct projection ray A and the reflected ray B, which hits the book at O and reflects. Therefore, the final camera ray C is a hybrid of both direct reflection and interreflection.

Figure 6.4 demonstrates the effects of the setup and our processing. The projected image is reflected off the specular book cover causing a ghost image on the white paper screen in the back. The interreflection further reduces the contrast of the projected image. Moreover, the white paper is moderately translucent and induces subsurface scattering which reduces the sharpness of the projection. All these global illumination effects, interreflection and subsurface scattering cause the degeneration of the projection (see Figure 6.4(b)). In order to suppress these artifacts, we first modulate the image using a sinusoid function with $f_x = \frac{1}{8}$ in x direction. This 1D separation (Figure 6.4(c)) reduces the global illumination effects, but one can still notice some remaining artifacts. The image in Figure 6.4(d) is obtained by using the two-pass separation with a sinusoid of $f_x = \frac{1}{8}$ and a binary pattern with $f_y = \frac{1}{8}$. The additional second pass removes the ghost image completely and improves the contrast.

This example demonstrates the effectiveness of modulation and multi-pass separation in removing interreflection and subsurface scattering effects even for general input signals. More generally, any low frequency input signal could be modulated to inherit the separation properties of the high frequency modulation function.

6.4 Modulated Phase-Shifting and Unwrapping

In the case of phase-shifting, the original patterns L are already one-dimensional sinusoids that inherently perform 1D separation for sufficiently high frequencies f. For temporal phase-unwrapping [Huntley93, Chen07] a series of different frequencies is applied, e.g. $\frac{1}{1024}$, $\frac{1}{512}$, $\frac{1}{256}$, $\frac{1}{128}$, $\frac{1}{64}$, $\frac{1}{32}$, $\frac{1}{16}$, and $\frac{1}{8}$. The highest frequencies are used for estimating the phase while the lower frequencies are required to unwrap the phase.

In order to obtain measurements that are robust to global illumination effects we apply a modulation in the y direction using a shifted high frequency pattern on top of the original phase-shifting signal and then perform the two-pass separation presented in the previous sections. The first pass, operating in the y direction, only performs the direct/global separation. The second pass is then used for both filtering out the global component that might have survived the first pass and for determining the phase or performing the unwrapping.
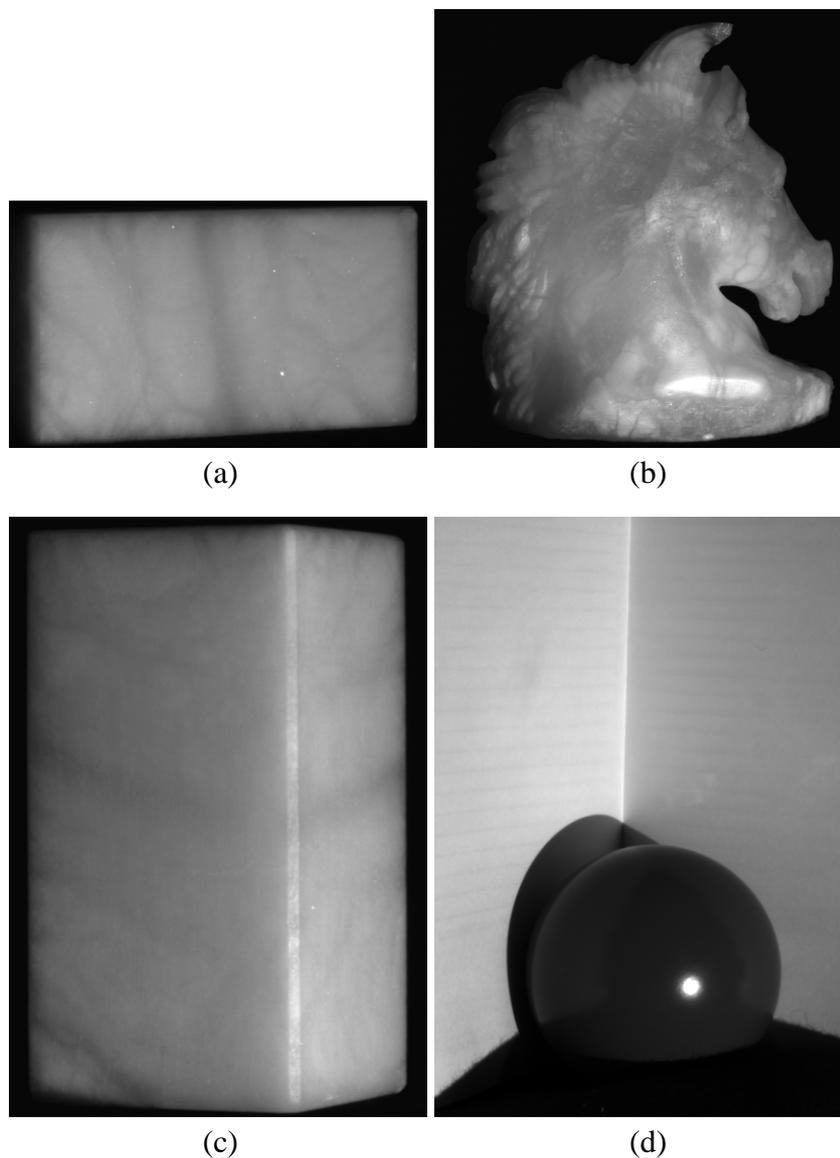
Figure 6.5: *Photographs of the experimental scenes. (a) One face of an alabaster block. (b) An alabaster horse head. (c) One edge of an alabaster block. (d) A glossy sphere in the corner of an open notebook.*

In theory, for the low frequencies a modulation in x and y would be necessary to remove any influence of global illumination effects on the unwrapping. In practice, since the low frequencies are only used to determine the period and not the phase, the unwrapping turns out to be stable enough even without any modulation in all our test cases.

Modulation is therefore only applied to the high frequency signals used to estimate the phase with $f_x = \frac{1}{8}$ or $f_x = \frac{1}{16}$. Since the frequencies in x are high enough to obtain reasonable separation performance, additional 1D modulation in y is sufficient. For the 3D scans presented in this chapter we applied a modulation frequency of $f_y = \frac{1}{6}$ with 6 shifts, leading to six times the number of images for the highest frequency. However, since no modulation is required for the lower frequencies the overhead for the whole 3D scanning using modulated phase-shifting is quite moderate. For PDI enhanced phase-shifting method, we need two times the number of images for the highest frequency.

6.5 Results

In this section we describe the experimental setup and demonstrate the effectiveness of our method in removing subsurface scattering and interreflections during 3D scanning. We show results for both planar (a face of an alabaster block) or non-planar objects (an edge of an alabaster block and an alabaster horse head). We also show results for a scene including high frequency interreflection (a glossy sphere in a corner). The four scenes are presented in Figure 6.5.

Our basic acquisition system consists of only one projector and one camera capturing images. In this setup, we again use a Mitsubishi XD490U XGA DLP projector (with 1024×768 pixels) and a 14-bit 1360×1024-pixel ProgRes®C14*plus* CCD camera. For comparison with polarization difference imaging enhanced phase-shifting [Chen07], we add two additional linear polarizers, one for the camera and the other for the projector. The camera's response curve is recovered using the method proposed by Debevec and Malik [Debevec97] and HDR images are taken by fusing multiple images with different exposure times. We calibrate the projector-camera system geometrically using Zhang's method [Zhang00].

6.5.1 Alabaster Plane

In the first example we perform 3D scanning on a planar slab of highly translucent alabaster with significant subsurface structures (Figure 6.6). Performing phase-shifting with 1D patterns of frequency $f_x = \frac{1}{8}$ separates out some of the subsurface scattering but gets corrupted by structures close to the surface. Applying polarization difference imaging filters out the depolarized multiply scattered components but also weakens the signal, resulting in a less biased but slightly noisier reconstructed geometry. An even better separation with a higher frequency signal is obtained by modulating the phase-shifting signal by a second sinusoid in the y direction of frequency $f = \frac{1}{6}$, following the approach outlined in the previous

(a)   (b)   (c)
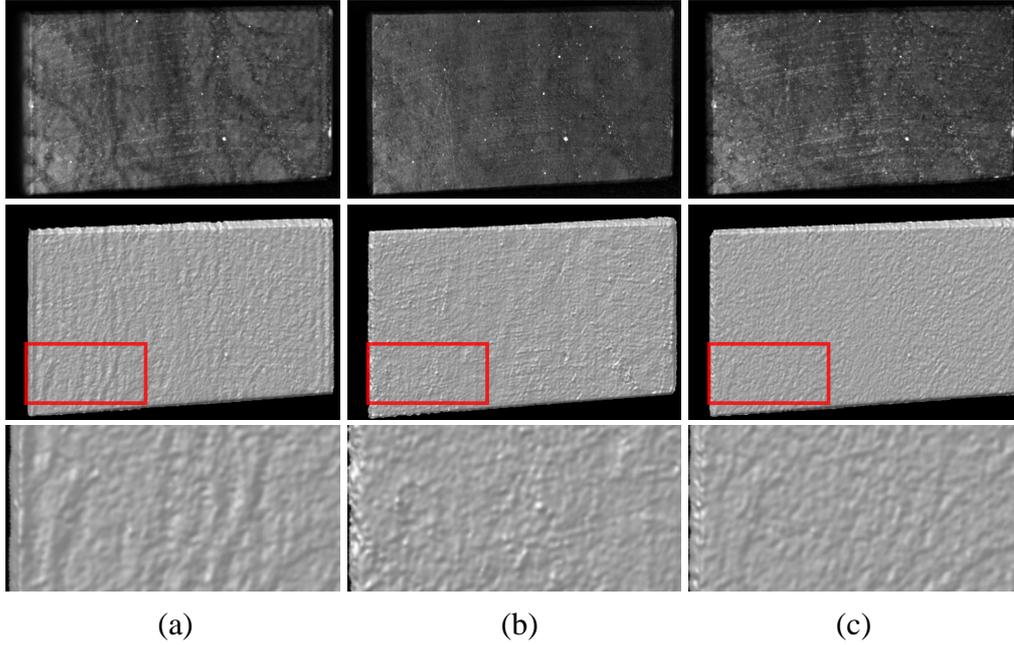
Figure 6.6: *Separated direct components (first row) and reconstructed 3D shapes (second row) for the alabaster slab. (a) Standard phase-shifting (PS) with $f_x = \frac{1}{8}$. (b) Polarization difference imaging (PDI) with $f_x = \frac{1}{8}$. (c) Modulated phase-shifting (MPS) with $f_x = \frac{1}{8}$ and $f_y = \frac{1}{6}$. The subsurface structures that are corrupting the 3D reconstruction in (a) have been successfully filtered out by both PDI and MPS.*

sections. The separated direct component and the resulting 3D geometry is almost free of subsurface structures and slightly smoother than the reconstruction based on PDI.

This increase in quality comes at a small cost in the time required for acquisition: For 1D phase-shifting, we applied 8 different shifts. For PDI this 8-image sequence is actually captured twice (16 images in total) for two different orientations of the camera polarization filter. In the modulated phase-shifting we added three phase shifts in the y direction amounting to $3 \times 8 = 24$ images.

The benefit of modulating a traditional phase-shifting pattern gets even more apparent if the original signal has lower frequency. In Figure 6.7, the dark veins of the alabaster cause some clear marks in the reconstructed geometry when applying traditional 1D phase-shifting with $f_x = \frac{1}{16}$. Modulating with a binary pattern of frequency $f_y = \frac{1}{6}$ again brings out a much better separated direct component and an unbiased 3D shape. The effectiveness of modulated phase-shifting (MPS) to obtain accurate results even for low frequencies is of high importance for phase-

(a)   (b)
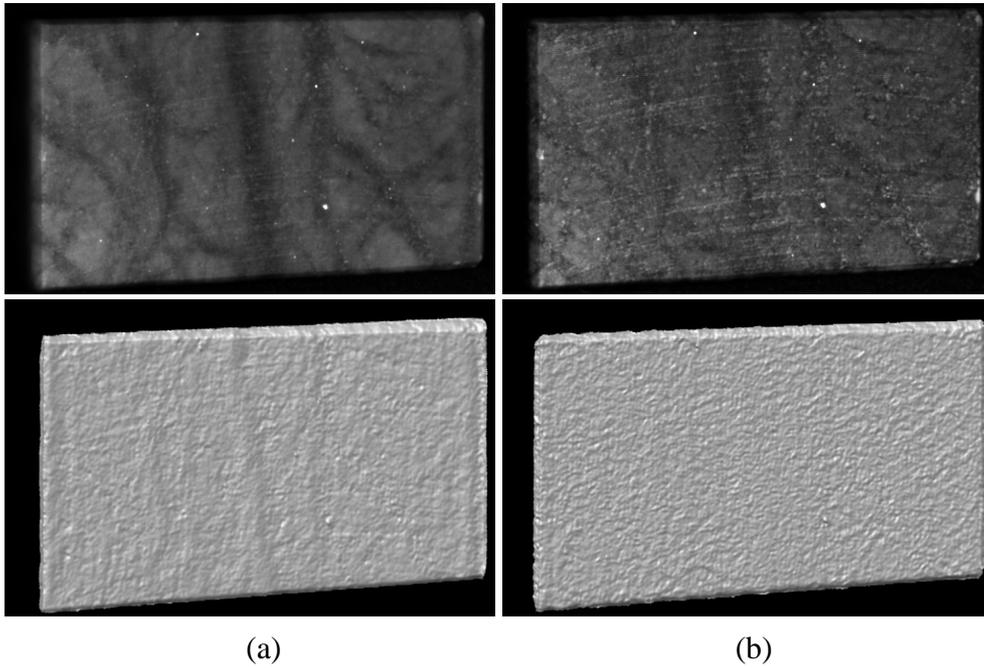
Figure 6.7: *Direct components and 3D shapes for the alabaster block with lower frequencies. (a) PS with $f_x = \frac{1}{16}$. (b) MPS with $f_x = \frac{1}{16}$ and $f_y = \frac{1}{6}$. The subsurface structures influence the estimated geometry from PS.*

shifting with temporal phase unwrapping where it brings two benefits: Phase-maps generated with lower frequencies typically carry slightly less noise than those obtained from higher frequencies. Furthermore, they are easier to unwrap because there are less ambiguous periods.

6.5.2 Alabaster Edge

While in the previous example PDI and our MPS have demonstrated rather similar performance we obtained rather different results on the second experiment where we rotated the alabaster block to capture one of its edges (Figure 6.8). Due to the different orientation of the normals which influence the polarization even of the direct reflection the measured direct component in the PDI is much darker. The noise in the 3D reconstruction by PDI is clearly noticeable. Standard phase-shifting still contains too much subsurface scattering which can be removed effectively using additional modulation.

(a)   (b)   (c)
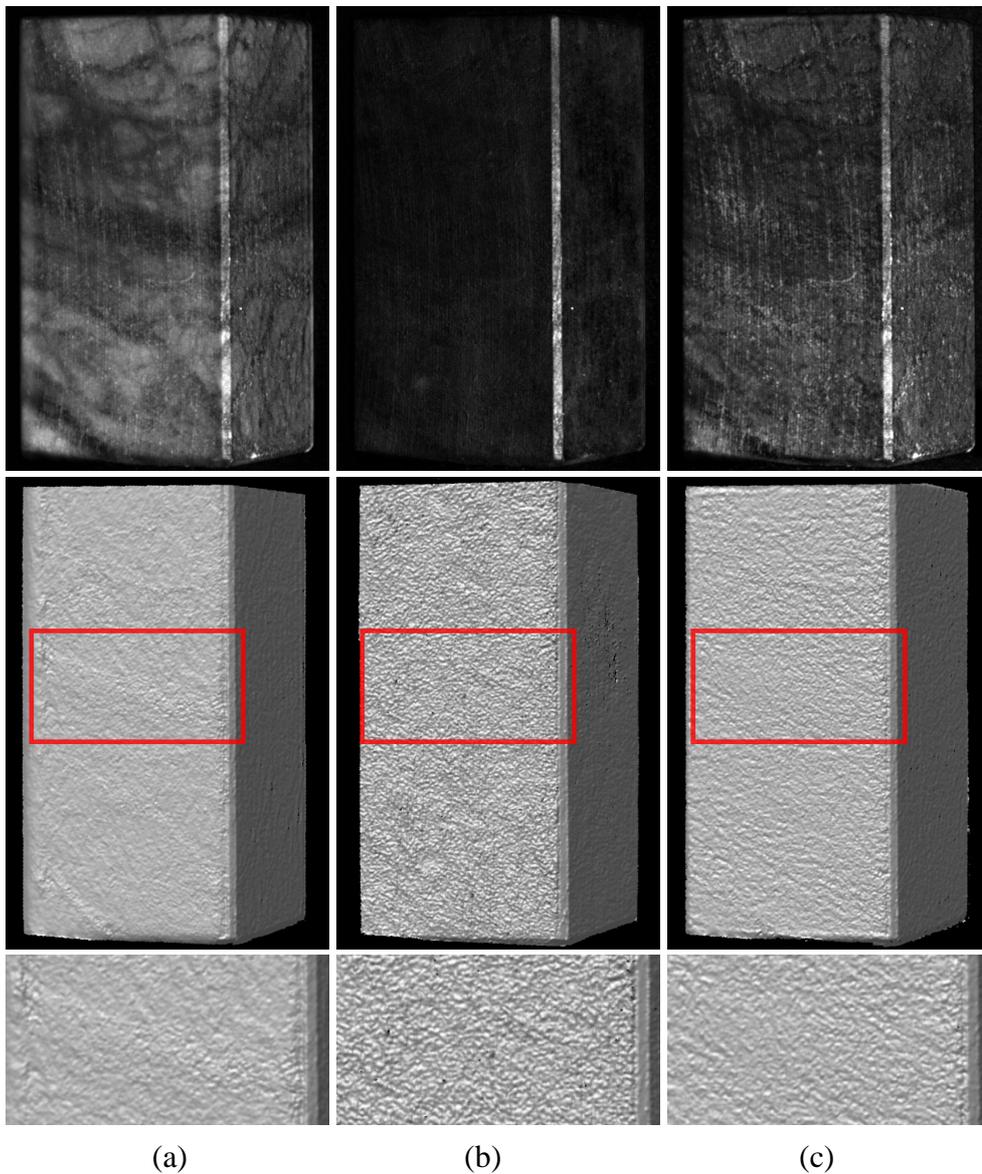
Figure 6.8: *Direct components and reconstructed 3D shapes for the edge of the alabaster block. (a) PS with $f_x = \frac{1}{8}$ is not sufficient to remove all subsurface scattering. (b) For PDI, $f_x = \frac{1}{8}$, the strength of the recovered direct component is largely dependent on the surface normal resulting in a noisy reconstruction. (c) MPS with $f_x = \frac{1}{8}$ and $f_y = \frac{1}{6}$ yields good results independent of the surface orientation.*

6.5.3   Alabaster Horse Head

A slightly better performance of the PDI approach is visible in the results we obtained for the even more translucent alabaster horse head in Figure 6.9. Due to its
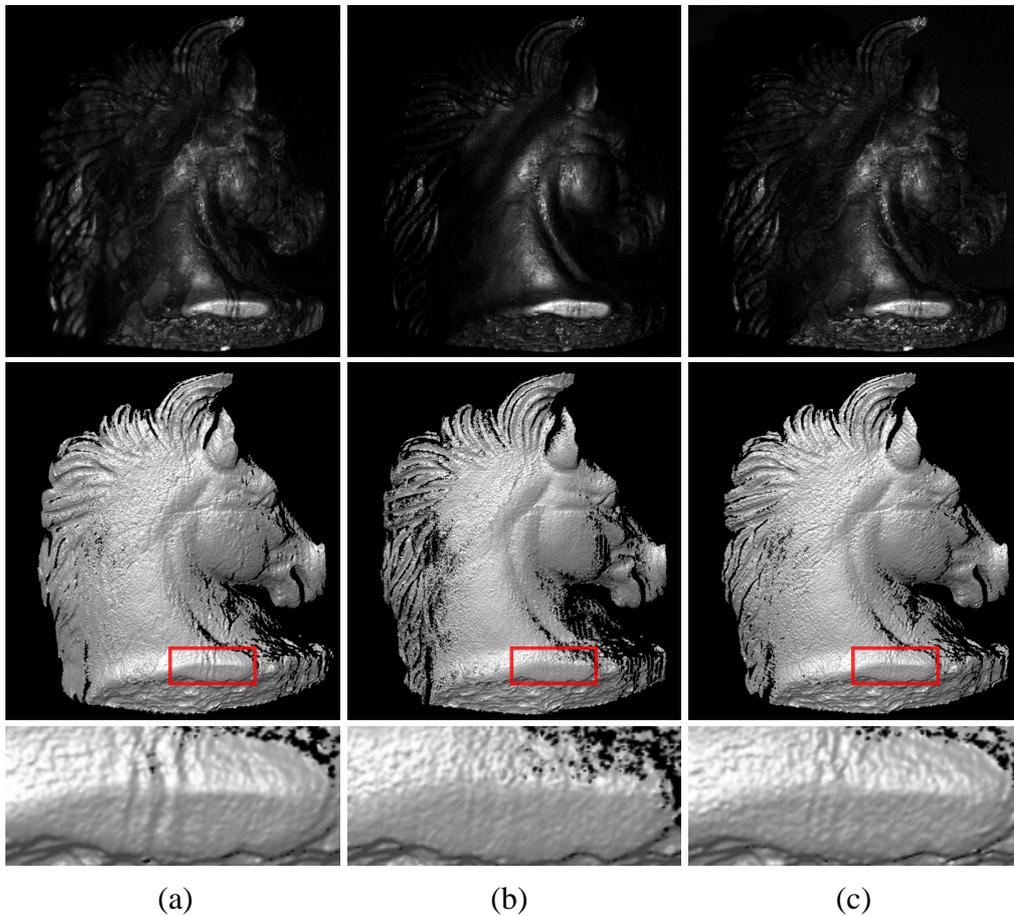
Figure 6.9: *Direct components and reconstructed 3D shapes for the alabaster horse head. (a) PS with $f_x = \frac{1}{8}$. (b) PDI with $f_x = \frac{1}{8}$. (c) MPS with $f_x = \frac{1}{8}$ and $f_y = \frac{1}{6}$. PDI reconstructs fewer surface points but produces a smoother surface for some parts. The applied frequency for MPS is not sufficient for a perfect separation in this case.*

dependence on the surface normal PDI cannot reconstruct as much of the surface as our MPS method but the reconstructed surface is slightly smoother, indicating that in this case the MPS could not filter out subsurface structures completely.

6.5.4 Interreflection Scene

The performance of the three approaches on a different global illumination effect is demonstrated in Figure 6.10. The illumination in this scene contains reflections of a glossy sphere as well as interreflections between the two pages of a book,

(a)   (b)   (c)
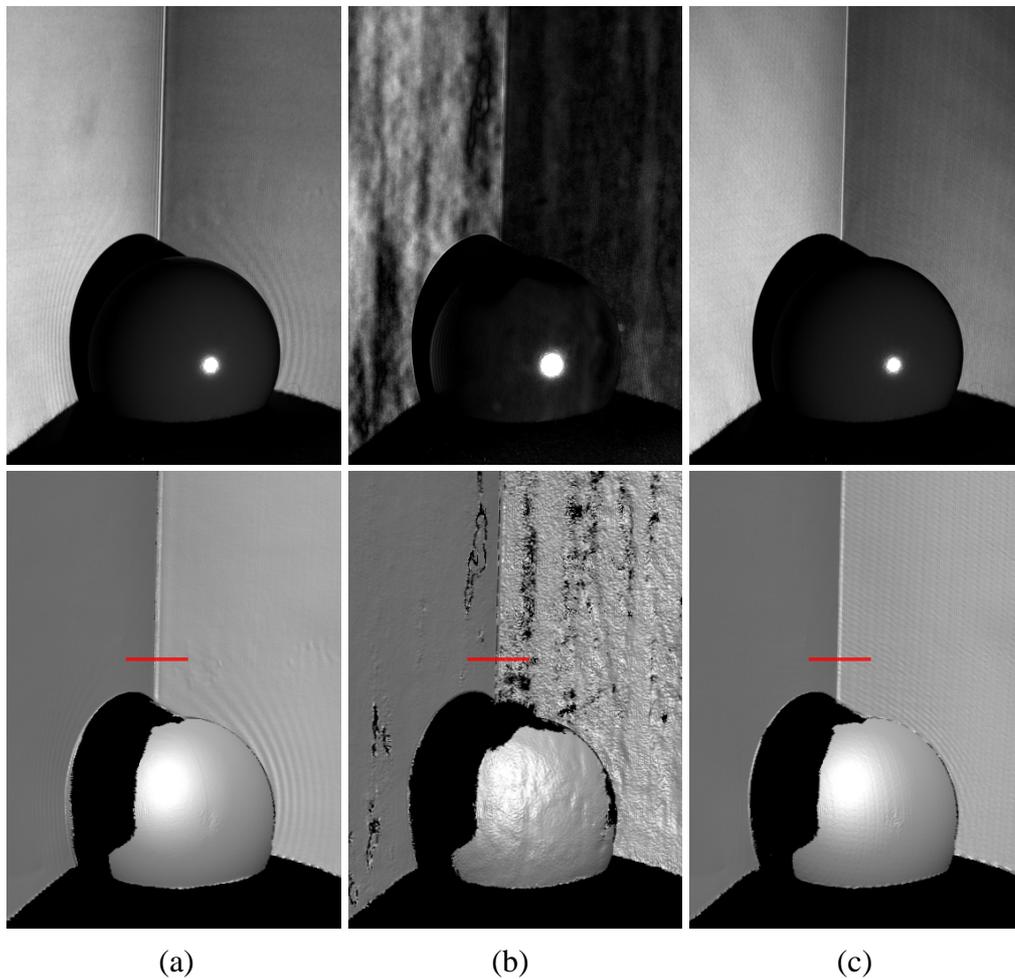
Figure 6.10: *A scene with strong interreflections caused by a glossy sphere be-
tween the two pages of a book. We apply the same frequencies as before. (a) PS
produces ripples in the final geometry due to the reflections of the sphere. (b) The
result by PDI is very noisy due to the modified polarization after interreflections.
(c) The most accurate reconstruction is achieved using MPS. See Figure 6.11 for
the profiles along the red lines.*

mostly noticeable near the edge. These interreflections heavily corrupt the 3D
geometry estimated by standard phase-shifting. The reconstruction based on PDI
is so noisy that it is hard to determine how much of the interreflections actually
have been filtered out. Applying modulated phase-shifting removes most of the
interreflections and produces the most accurate 3D shape. For comparison, the
3D profiles of the edge are shown in Figure 6.11. Both the PS and the PDI meth-
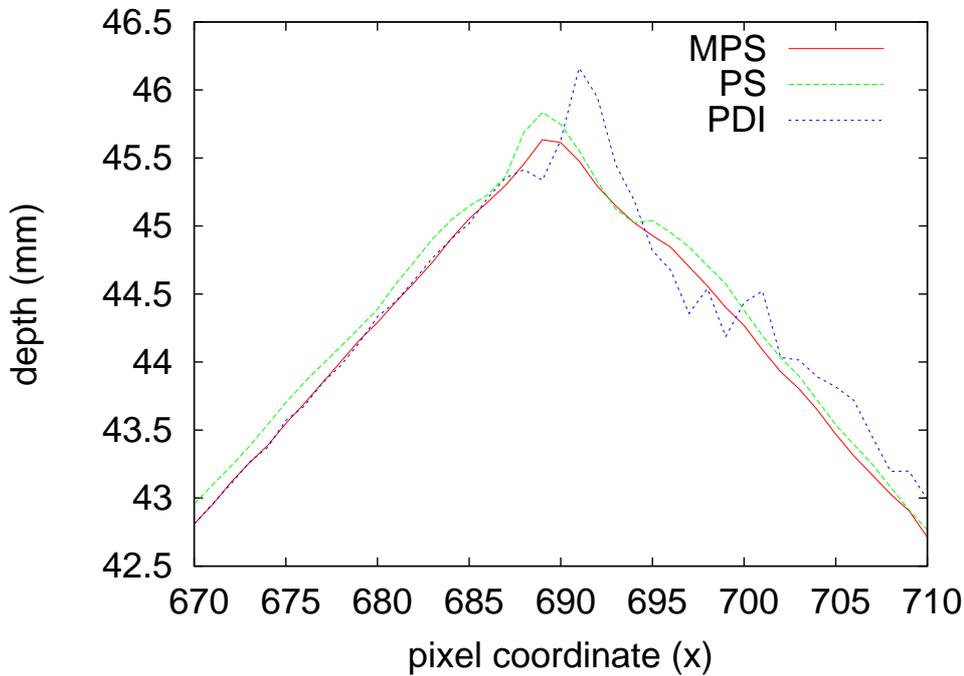
Figure 6.11: *Corner profiles of the scene in Figure 6.10. Both PS and PDI show significant artifacts due to the interreflections.*

ods suffer from the strong interreflection, while the MPS method reconstructs the corner geometry very well.

6.5.5 Two-Pass vs. Single-Pass 2D Separation

For the edge of the alabaster block, Figure 6.2 shows the difference between two-pass separation we used in all the presented experiments and the minimum-maximum algorithm on 2D patterns [Nayar06]. In the 2D minimum-maximum approach, the direct component is estimated from the minimum and maximum intensity values for all measurements of each pixel, while in the two-pass method the separation is performed sequentially along the y and x directions. Although exactly the same input images are used, our two-pass method by exploiting the non-linearity in the analysis results in a much clearer separation where the direct component contains far less of the global subsurface scattering component.

6.6   Limitations

As demonstrated in Figure 6.6 and 6.7 the separation performance of the structured patterns depends on the highest frequency employed. While modulated phase-shifting introduces a principled way of increasing the separation performance for a low-frequency input signal, the carrier frequency is still limited by the projection system. The maximum reliable frequency we could project is $\frac{1}{6}$ since our projector cannot project very high-frequency patterns accurately. The projector introduces serious artifacts (interference-like wave artifacts) when projecting checkerboard with checker size smaller than 3. In Figure 6.9(c) a higher frequency would have been needed to completely remove all subsurface scattering. Even at size 3, the projected patterns contains noticeable artifacts, which cause high frequency noise in the result scans. Most of the presented results to a small extent suffer from this problem. It is mostly noticeable in Figure 6.10(c), where the smooth paper surface is polluted by small scale noise. This limitation renders modulated phase-shifting inferior to traditional phase-shifting when scanning opaque objects without global illumination effects.

Chapter 7

Conclusion and Future Work

3D scanning of complex objects, rich of interreflection, subsurface scattering, fine-scale geometric details etc., has general and continuous interests for both computer graphics and computer vision. In this thesis, we presented three classes of 3D scanning methods.

- We presented a mesostructure reconstruction method that can successfully handle complex real-world objects. The main observation we build our approach on is that specular reflection reveals effectively fine-scale surface details and yields reliable visual information for mesostructure reconstruction. To exploit this observation, we developed a very simple system, including a camera, an LED light source, four specular spheres, and a checkerboard. The user can interactively control the light source to incrementally capture the specularity field. The resulting dense specularity field is used for mesostructure reconstruction. We demonstrated the effectiveness of our method by efficiently capturing the mesostructure of a variety of complex real-world objects, including human skin, dried apricot, orange skin, jelly candy, black leather, and dark chocolate.

  In the future, we would like to use multiple cameras to capture a complete mesostructure model of 3D object with complex reflectance properties. Another extension of our method is a hybrid reconstruction using shape from specularity when specular highlights are detected, and shape from photometric stereo for pixels below the threshold for specular highlights. Using the reconstructed mesostructure to improve the compactness of BTF representations and to realistically render highly detailed surfaces is a further research direction.

- The second method is a hybrid using both phase-shifting and polarization. A careful analysis of phase-shifting without polarization, combined with

parallel polarization and with polarization difference imaging has shown that some of the shortcomings of pure phase-shifting such as its sensitivity to subsurface structures can be overcome. Even though PDI has the ability to robustly remove all global effects due to subsurface scattering which otherwise renders accurate 3D scanning a hard problem, parallel polarization sometimes provides a better SNR resulting in less noisy range maps. Depending on the richness of subsurface structure, the planarity of the object and the sensitivity of the camera one might choose one method over the other.

The descattering property of polarization can also be used by other structured light techniques, such as gray code, binary code, De Bruijn sequences etc. [Salvi04]. It would be also interesting to exploit the hybrid method that can intelligently choose the right method to do reliable 3D scanning for general translucent objects. We expect there will be some further descattering based methods that can help accurate and robust 3D reconstruction of translucent objects or even general scenes including a wide range of materials.

- Traditional phase-shifting for profilometry usually employs sinusoid patterns where the signal varies only in one direction. Phase-shifting is one of the most robust 3D scanning techniques but the results might still be corrupted by global illumination effects such as subsurface scattering or inter-reflections. We introduced the concept of modulated phase-shifting where the original one-dimensional pattern is multiplied by a second shifted high-frequency pattern in the other direction. This modulation requires capturing more images but results in significantly improved robustness against pollution due to these global illumination effects.

  Compared to 3D scanning techniques based on polarization difference imaging (PDI), the modulation based phase-shifting is less demanding with regard to the acquisition setup. It clearly outperforms PDI for objects with strongly varying surface normals as well as for scenes with strong inter-reflections.

  Moreover, we extended the modulation scheme to general low-frequency functions, e.g. an ordinary photograph. This provides a new possibility to enhance the robustness of applications where the frequency of the original signal is rather limited.

  An interesting direction for future investigation is a systematic evaluation of different modulation functions and modulation schemes. Another valuable extension is to exploit the general modulation for novel applications.

Bibliography

[Adato07]   YAIR ADATO, YURIY VASILYEV, OHAD BEN-SHAHAR,
                 AND TODD ZICKLER. Toward a Theory of Shape from Spec-
                 ular Flow. In *Proceedings of IEEE Int. Conference on Com-
                 puter Vision*, 2007.

[Alldrin08]   NEIL ALLDRIN, TODD ZICKLER, AND DAVID KRIEG-
                 MAN. Photometric Stereo with Non-Parametric and Spatially-
                 Varying Refletance. In *Proceedings of IEEE Conf. on Com-
                 puter Vision and Pattern Recognition*, 2008.

[Bernardini02]   F. BERNARDINI, H. RUSHMEIER, I. M. MARTIN, J. MIT-
                 TLEMAN, AND G. TAUBIN. Building a Digital Model of
                 Michelangelo's Florentine Pietà. *IEEE Computer Graphics
                 and Applications*, 22(1):59–67, 2002.

[Besl88]   PAUL J. BESL. Active, Optical Range Imaging Sensors.
                 Mach. Vision Appl., 1(2):127–152, 1988.

[Blake90]   A. BLAKE AND H. H. BÜLTHOFF. Does the brain know the
                 physics of specular reflection? *Nature*, 343:165–168, 1990.

[Blake91]   A. BLAKE AND H. H. BÜLTHOFF. Shape from Specularities:
                 Computation and Psychophysics. *Philosophical Trans. of the
                 Royal Society of London*, B 331:237–252, 1991.

[Blinn78]   J. F. BLINN. Simulation of Wrinkled Surfaces. *Computer
                 Graphics*, 12(3):286–292, 1978.

[Born99]   M. BORN, E. WOLF, AND A. B. BHATIA. *Principles of
                 Optics*. Cambridge University Press, 1999.

[Bouguet]   J.-Y. BOUGUET. MATLAB Camera Calibration Toolbox.
                 http://www.vision.caltech.edu/bouguetj/calib_doc/.

[Bruning74]   J. H. BRUNING, D. R. HERRIOTT, J. E. GALLAGHER, D. P. ROSENFELD, A. D. WHITE, AND D. J. BRANGACCIO. Digital Wavefront Measuring Interferometer for Testing Optical Surfaces and Lenses. *Appl. Opt.*, 13:2693–2703, 1974.

[Chen06]   TONGBO CHEN, MICHAEL GOESELE, AND HANS-PETER SEIDEL. Mesostructure from Specularity. In *Proceedings of IEEE Conf. on Computer Vision and Pattern Recognition*, pages 1825–1832, 2006.

[Chen07]   TONGBO CHEN, HENDRIK LENSCH, CHRISTIAN FUCHS, AND HANS-PETER SEIDEL. Polarization and Phase-Shifting for 3D Scanning of Translucent Objects. In *Proceedings of IEEE Conf. on Computer Vision and Pattern Recognition*, 2007.

[Chen08]   TONGBO CHEN, HANS-PETER SEIDEL, AND HENDRIK LENSCH. Modulated Phase-Shifting for 3D Scanning. In *Proceedings of IEEE Conf. on Computer Vision and Pattern Recognition*, 2008.

[Clark97]   J. CLARK, E. TRUCCO, AND L. B. WOLFF. Using Light Polarization in Laser Scanning. *Image and Vision Computing*, 15(2):107–117, 1997.

[Curless95]   BRIAN CURLESS AND MARC LEVOY. Better Optical Triangulation through Spacetime Analysis. In *Proceedings of International Conf. on Computer Vision*, pages 987–994, 1995.

[Curless97]   B. CURLESS. *New Methods for Surface Reconstruction from Range Images*. PhD thesis, Stanford University, 1997.

[Dana99]   K. J. DANA, B. VAN GINNEKEN, S. K. NAYAR, AND J. J. KOENDERINK. Reflectance and Texture of Real World Surfaces. *ACM Transactions on Graphics*, 18(1):1–34, 1999.

[Davis05]   J. DAVIS, D. NEHAB, R. RAMAMOORTHI, AND S. RUSINKIEWICZ. Spacetime Stereo: a Unifying Framework for Depth from Triangulation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 27(2):296–302, 2005.

[Debevec97]   P. E. DEBEVEC AND J. MALIK. Recovering High Dynamic Range Radiance Maps from Photographs. In *Proceedings of SIGGRAPH*, pages 369–378, 1997.

[Fleming04]   R. FLEMING, A. TORRALBA, AND E. H. ADELSON. Specular Reflections and the Perception of Shape. *Journal of Vision*, 4(9):798–820, 2004.

[Fleming05]   R. FLEMING AND H. BUELTHOFF. Low-Level Images Cues in the Perception of Translucent Materials. *ACM Transactions on Applied Perception*, 2(3):346–382, 2005.

[Francken08]   YANNICK FRANCKEN, TOM CUYPERS, TOM MERTENS, JO GIELIS, AND PHILIPPE BEKAERT. High Quality Mesostructure Acquisition Using Specularities. In *Proceedings of IEEE Conf. on Computer Vision and Pattern Recognition*, 2008.

[Frankot88]   R. T. FRANKOT AND R. CHELLAPPA. A Method for Enforcing Integrability in Shape from Shading Algorithms. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 10(4):439–451, 1988.

[Ghiglia98]   DENNIS C. GHIGLIA AND MARK D. PRITT. *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*. Wiley, New York, 1998.

[Godin01]   GUY GODIN, J.-ANGELO BERALDIN, MARC RIOUX, MARC LEVOY, LUC COURNOYER, AND FRANCOIS BLAIS. An Assessment of Laser Range Measurement of Marble Surfaces. In *Proceedings of Fifth Conference on Optical 3-D Measurement Techniques*, pages 49–56, 2001.

[Goldman05]   D. B. GOLDMAN, B. CURLESS, A. HERTZMANN, AND S. M. SEITZ. Shape and Spatially-Varying BRDFs from Photometric Stereo. In *Proceedings of International Conf. on Computer Vision*, pages 341–348, 2005.

[Hall-Holt01]   OLAF HALL-HOLT AND SZYMON RUSINKIEWICZ. Stripe Boundary Codes for Real-Time Structured-Light Range Scanning of Moving Objects. In *Proceedings of International Conf. on Computer Vision*, pages 359–366, 2001.

[Hasinoff06]   S. W. HASINOFF AND K. N. KUTULAKOS. Confocal Stereo. In *Proceedings of European Conference on Computer Vision*, pages 620–634, 2006.

[Hertzmann05]   AARON HERTZMANN AND STEVEN M. SEITZ. Example-Based Photometric Stereo: Shape Reconstruction with General, Varying BRDFs. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 27(8):1254–1264, 2005.

[Huang06]   PEISEN S. HUANG AND SONG ZHANG. Fast Three-Step Phase-Shifting Algorithm. *Applied Optics*, 45(21):5086–5091, 2006.

[Hullin08]   MATTHIAS B. HULLIN, MARTIN FUCHS, IVO IHRKE, HANS-PETER SEIDEL, AND HENDRIK P. A. LENSCH. Fluorescent Immersion Range Scanning. In *Proceedings of SIGGRAPH*, 2008.

[Huntley93]   J. M. HUNTLEY AND H. O. SALDNER. Temporal Phase Unwrapping Algorithm for Automated Interferogram Analysis. *Appl. Opt.*, 32:3047–3052, 1993.

[Huntley97a]   J. M. HUNTLEY AND H. O. SALDNER. Phase Unwrapping: Application to Surface Profiling of Discontinuous Objects. *Appl. Opt.*, 36:2770–2775, 1997.

[Huntley97b]   J. M. HUNTLEY AND H. O. SALDNER. Shape Measurement by Temporal Phase Unwrapping: Comparison of Unwrapping Algorithms. *Meas. Sci. Technol.*, 8:986–992, 1997.

[Ikeuchi81]   K. IKEUCHI. Recognition of 3D Objects Using the Extended Gaussian Image. In *Proceedings of International Joint Conf. on Artificial Intelligence*, pages 595–600, 1981.

[Ishimaru78]   A. ISHIMARU. *Wave Propagation and Scattering in Random Media*. Academic Press, 1978.

[Jensen01]   HENRIK WANN JENSEN, STEPHEN R. MARSCHNER, MARC LEVOY, AND PAT HANRAHAN. A Practical Model for Subsurface Light Transport. In *Proceedings of SIGGRAPH 2001*, pages 511–518, 2001.

[Klinker88]   G. KLINKER, STEVEN SHAFER, AND TAKEO KANADE. Color Image Analysis with an Intrinsic Reflection Model. In *Proceedings of International Conf. on Computer Vision*, pages 292–296, 1988.

[Koenderink96] J. J. KOENDERINK AND A. DOORN. Illuminance Texture Due to Surface Mesostructure. *J. Opt. Soc. Am.*, 13(3):452–463, 1996.

[Kutulakos05] K. N. KUTULAKOS AND E. STEGER. A Theory of Refractive and Specular 3D Shape by Light-Path Triangulation. In *Proceedings of International Conf. on Computer Vision*, pages 1448–1455, 2005.

[Larson92] G. W. LARSON. Measuring and Modeling Anisotropic Reflection. *Comp. Graphics*, 26(2):265–272, 1992.

[Levoy00] MARC LEVOY, KARI PULLI, BRIAN CURLESS, SZYMON RUSINKIEWICZ, DAVID KOLLER, LUCAS PEREIRA, MATT GINZTON, SEAN ANDERSON, JAMES DAVIS, JEREMY GINSBERG, JONATHAN SHADE, AND DUANE FULK. The Digital Michelangelo Project: 3D Scanning of Large Statues. In *Proceedings of SIGGRAPH*, pages 131–144, 2000.

[Liu01] X. LIU, Y. YU, AND H.-Y. SHUM. Synthesizing Bidirectional Texture Functions for Real-World Surfaces. In *Proceedings of SIGGRAPH*, pages 97–106, 2001.

[Lu00] RONG LU, J. J. KOENDERINK, AND A. M. L KAPPERS. Specularities on Surfaces with Tangent Hairs or Grooves. *CVIU*, 78:320–335, 2000.

[Lythgoe79] J. N. LYTHGOE. *The Ecology of Vision*. Oxfrod Uni. Press, 1979.

[Ma07] WAN-CHUN MA, TIM HAWKINS, PIETER PEERS, CHARLES-FELIX CHABERT, MALTE WEISS, AND PAUL DEBEVEC. Rapid Acquisition of Specular and Diffuse Normal Maps from Polarized Spherical Gradient Illumination. In *Proceedings of EGSR*, 2007.

[Magda01] S. MAGDA AND T. ZICKLER. Beyond Lambert: Reconstructing Surfaces with Arbitrary BRDFs. In *Proceedings of International Conf. on Computer Vision*, pages 391–398, 2001.

[Masselus02] V. MASSELUS, P. DUTRÉ, AND F. ANRYS. The Free-Form Light Stage. In *Proceedings of EGRW*, pages 247–256, 2002.

[Miyazaki02]   D. MIYAZAKI, M. SAITO, Y. SATO, AND K. IKEUCHI. De-
                  termining Surface Orientations of Transparent Objects Based
                  on Polarization Degrees in Visible and Infrared Wavelengths.
                  J. Opt. Soc. Am. A, 19(4):687–694, 2002.

[Miyazaki03a]   DAISUKE MIYAZAKI, MASATAKA KAGESAWA, AND KAT-
                  SUSHI IKEUCHI.   Polarization-Based Transparent Surface
                  Modeling from Two Views. In *Proceedings of International
                  Conf. on Computer Vision*, page 1381, 2003.

[Miyazaki03b]   DAISUKE MIYAZAKI, R. T. TAN, K. HARA, , AND KAT-
                  SUSHI IKEUCHI. Polarization-Based Inverse Rendering from
                  a Single View. In *Proceedings of IEEE International Conf.
                  on Computer Vision*, 2003.

[Miyazaki05]   D. MIYAZAKI AND K. IKEUCHI. Inverse Polarization Ray-
                  tracing: Estimating Surface Shape of Transparent Objects. In
                  *Proceedings of IEEE Conf. on Computer Vision and Pattern
                  Recognition*, pages 910–917, 2005.

[Morel06]   O. MOREL, C. STOLZ, F. MERIAUDEAU, AND P. GORRIA.
                  Active Lighting Applied to Three-Dimensional Reconstruc-
                  tion of Specular Metallic Surfaces by Polarization Imaging.
                  Appl. Opt., 45:4062–4068, 2006.

[Morris07a]   NIGEL J. W. MORRIS AND KIRIAKOS N. KUTULAKOS. Dy-
                  namic Refraction Stereo. In *Proceedings of International
                  Conf. on Computer Vision*, pages 1573–1580, 2007.

[Morris07b]   NIGEL J. W. MORRIS AND KIRIAKOS N. KUTULAKOS.
                  Reconstructing the Surface of Inhomogeneous Transparent
                  Scenes by Scatter Trace Photography. In *Proceedings of In-
                  ternational Conf. on Computer Vision*, 2007.

[Müller05]   G. MÜLLER, J. MESETH, M. SATTLER, R. SARLETTE,
                  AND R. KLEIN. Acquisition, Synthesis and Rendering of
                  Bidirectional Texture Functions. *Computer Graphics Forum*,
                  24(1):83–109, 2005.

[Narasimhan05]   S.G. NARASIMHAN, S.K. NAYAR, B. SUN, AND S.J. KOP-
                  PAL. Structured Light in Scattering Media. *Proceedings of In-
                  ternational Conf. on Computer Vision*, pages 420–427, 2005.

[Nayar97]   S. K. NAYAR, X. S. FANG, AND T. BOULT. Separation of Reflection Components Using Color and Polarization. *International Journal on Computer Vision*, 21(3):163–186, 1997.

[Nayar06]   SHREE K. NAYAR, GURUNANDAN KRISHNAN, MICHAEL D. GROSSBERG, AND RAMESH RASKAR. Fast Separation of Direct and Global Components of a Scene Using High Frequency Illumination. *ACM Transactions on Graphics*, 25(3):935–944, 2006.

[Nehab05]   DIEGO NEHAB, SZYMON RUSINKIEWICZ, JAMES DAVIS, AND RAVI RAMAMOORTHI. Efficiently Combining Positions and Normals for Precise 3D Geometry. *ACM Trans. Graph.*, 24(3):536–543, 2005.

[Nehab08]   D. NEHAB, T. WYRICH, AND S. RUSINKIEWICZ. Dense 3D Reconstruction from Specularity Consistency. In *Proceedings of IEEE Conf. on Computer Vision and Pattern Recognition*, 2008.

[Neubeck05]   A. NEUBECK, A. ZALESNY, AND L. V. GOOL. 3D Texture Reconstruction from Extensive BTF Data. In *Proceedings of International Workshop on Texture Analysis and Synthesis*, pages 13–18, 2005.

[Nicodemus77] F. E. NICODEMUS, J. C. RICHMOND, J. J. HSIA, I. W. GINSBERG, AND T. LIMPERIS. Geometrical Considerations and Nomenclature for Reflectance. National Bureau of Standards, 1977.

[Norman04]   J. F. NORMAN. Perception of Three-Dimensional Shape from Specular Highlights, Deformations of Shading, and other Types of Visual Information. *Psychol. Sci.*, 15(8):565–570, 2004.

[Oren96]   M. OREN AND S. K. NAYAR. A Theory of Specular Surface Geometry. *International Journal on Computer Vision*, 24:105–124, 1996.

[Paterson05]   JAMES A. PATERSON, DAVID CLAUS, AND ANDREW W. FITZGIBBON. BRDF and Geometry Capture from Extended Inhomogeneous Samples Using Flash Photography. *Computer Graphics Forum*, 24(3):383–391, 2005.

[Poussart88]   D. POUSSART AND D. LAURENDEAU. 3-D Sensing for Industrial Computer Vision. pages 122–159, 1988.

[Rahmann01]   S. RAHMANN AND N. CANTERAKIS. Reconstruction of Specular Surfaces Using Polarization Imaging. In *Proceedings of IEEE Conf. on Computer Vision and Pattern Recognition*, pages 149–155, 2001.

[Robertson03]   M. A. ROBERTSON, S. BORMAN, AND R. L. STEVENSON. Estimation-Theoretic Approach to Dynamic Range Enhancement using Multiple Exposures. *J. Electronic Imaging*, 12(2):219–285, 2003.

[Rowe95]   M. P. ROWE AND JR. E. N. PUGH. Polarization-Difference Imaging: a Biologically Inspired Technique for Observation through Scattering Media. *Optics Letters*, 20(6):608–610, 1995.

[Rushmeier97]   HOLLY E. RUSHMEIER, GABRIEL TAUBIN, AND A. GUÉZIEC. Applying Shape from Lighting Variation to Bump Map Capture. In *Proceedings of EGWR*, pages 35–44, 1997.

[Rusinkiewicz02] SZYMON RUSINKIEWICZ, OLAF HALL-HOLT, AND MARC LEVOY. Real-Time 3D Model Acquisition. In *Proceedings of SIGGRAPH*, pages 438–446, 2002.

[Saito99]   M. SAITO, Y. SATO, K. IKEUCHI, AND H. KASHIWAGI. Measurement of Surface Orientations of Transparent Objects by Use of Polarization in Highlight. *J. Opt. Soc. Am. A*, 16(9):2286–2293, 1999.

[Salvi04]   J. SALVI, J. PAGES, AND J. BATLLE. Pattern Codification Strategies in Structured Light Systems. *Pattern Recognition*, 37(4):827–849, 2004.

[Sanderson88]   A. C. SANDERSON, L. E. WEISS, AND S. K. NAYAR. Structured Highlight Inspection of Specular Surfaces. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 10(1):44–55, 1988.

[Savarese01]   SILVIO SAVARESE AND PIETRO PERONA. Local Analysis for 3D Reconstruction of Specular Surfaces - Part I. In

Proceedings of IEEE Conf. on Computer Vision and Pattern Recognition, pages 738–745, 2001.

[Savarese02] SILVIO SAVARESE AND PIETRO PERONA. Local Analysis for 3D Reconstruction of Specular Surfaces - Part II. In *Proceedings of European Conference on Computer Vision*, pages 759–774, 2002.

[Schechner00] Y. Y. SCHECHNER, JOSEPH SHAMIR, AND NAHUM KIRYATI. Polarization and Statistical Analysis of Scenes Containing a Semireflector. *J. Opt. Soc. Am. A*, 17(2):276–284, 2000.

[Schechner03] Y. Y. SCHECHNER, S. G. NARSIMHAN, AND S. K. NAYAR. Polarization-Based Vision through Haze. *Applied Optics*, 42(3):511–525, 2003.

[Schechner04] YOAV Y. SCHECHNER AND N. KARPEL. Clear under water vision. In *Proceeding of IEEE Conf. on Computer Vision and Pattern Recognition*, pages 536–543, 2004.

[Schechner05] YOAV Y. SCHECHNER AND NIR KARPEL. Recovery of Underwater Visibility and Structure by Polarization Analysis. *IEEE Journal of Oceanic Engineering*, 30(3):570–587, 2005.

[Seitz05] STEVEN M. SEITZ, YASUYUKI MATSUSHITA, AND KIRIAKOS N. KUTULAKOS. A Theory of Inverse Light Transport. In *Proceedings of International Conf. on Computer Vision*, pages 1440–1447, 2005.

[Shafer85] S. A. SHAFER. Using Color to Separate Reflection Components. *Color Research Application*, 10:210–218, 1985.

[Solem04] JAN ERIK SOLEM, HENRIK AANAES, AND ANDERS HEYDEN. A Variational Analysis of Shape from Specularities Using Sparse Data. In *Proceedings of 3DPVT*, pages 26–33, 2004.

[Srinivasan85] V. SRINIVASAN, HSIN-CHU LIU, AND MAURICE HALIOUA. Automated Phase-Measuring Profilometry: A Phase Mapping Approach. *Appl. Opt.*, 24:185–188, 1985.

[Talvala07] EINO-VILLE TALVALA, ANDREW ADAMS, MARK HOROWITZ, AND MARC LEVOY. Veiling Glare in High Dynamic Range Imaging. In *Proceedings of SIGGRAPH '07*, page 37, 2007.

[Tan05]   ROBBY T. TAN AND KATSUSHI IKEUCHI. Separating Re-
                   flection Components of Textured Surfaces Using a Single Im-
                   age. *IEEE Trans. on Pattern Analysis and Machine Intelli-
                   gence*, 27(2):178–193, 2005.

[Tarini05]   MARCO TARINI, HENDRIK P. A. LENSCH, MICHAEL GOE-
                   SELE, AND HANS-PETER SEIDEL. 3D Acquisition of Mir-
                   roring Objects. *Graphical Models*, 67(4):233–259, 2005.

[Todd04a]   J. T. TODD. The Visual Perception of 3D Shape. *TRENDS in
                   Cognitive Science*, 8(3):115–121, 2004.

[Todd04b]   J. T. TODD, J. .F. NORMAN, AND E. MINGOLLA. Lightness
                   Constancy in the Presence of Specular highlights. *Psycholog-
                   ical Science*, 15(1):33–39, 2004.

[Tomasi98]   C. TOMASI AND R. MANDUCHI. Bilateral Filtering for Gray
                   and Color Images. In *Proceedings of International Conf. on
                   Computer Vision*, pages 839–846, 1998.

[Treibitz06]   TALI TREIBITZ AND YOAV Y. SCHECHNER. Instant
                   3Descatter. In *Proceedings of IEEE Conf. on Computer Vi-
                   sion and Pattern Recognition*, pages 1861–1868, 2006.

[Treibitz08]   TALI TREIBITZ AND YOAV Y. SCHECHNER. Active Polar-
                   ization Descattering. *IEEE Trans. on Pattern Analysis and
                   Machine Intelligence*, 2008.

[Trifonov06]   B. TRIFONOV, D. BRADLEY, AND W. HEIDRICH. Tomo-
                   graphic Reconstruction of Transparent Objects. In *Proceed-
                   ings of Eurographics Symposium on Rendering*, 2006.

[Tyo96]   J. S. TYO, M. P. ROWE, E. N. PUGH, AND N. ENGHETA.
                   Target Detection in Optically Scattering Media by Polariza-
                   tion Difference Imaging. *App. Opt.*, 35:639–647, 1996.

[Umeyama04]   SHINJI UMEYAMA AND GUY GODIN. Separation of Diffuse
                   and Specular Components of Surface Reflection by Use of
                   Polarization and Statistical Analysis of Images. *IEEE Trans.
                   on Pattern Analysis and Machine Intelligence*, 26(5):639–
                   647, 2004.

[van de Hulst81]   HENDRIK CHRISTOFFEL VAN DE HULST. *Light Scattering
                   by Small Particles*. Courier Dover Publications, New York,
                   Dover, 1981.

[Veeraraghavan07] Ashok Veeraraghavan, Ramesh Raskar, Amit Agrawal, Ankit Mohan, and Jack Tumblin. Dappled Photography: Mask Enhanced Cameras for Heterodyned Light Fields and Coded Aperture Refocusing. In *Proceedings of SIGGRAPH*, page 69, 2007.

[Wallace99] A. M. Wallace, B. Liang, E. Trucco, and J. Clark. Improving Depth Image Acquisition Using Polarized Light. *International Journal of Computer Vision*, 32(2):87–109, 1999.

[Wang03] Jing Wang and Kristin J. Dana. A Novel Approach for Texture Shape Recovery. In *Proceedings of International Conf. on Computer Vision*, pages 1374–1380, 2003.

[Wang05] J. Wang, X. Tong, J. Snyder, Y. Chen, B. Guo, and H.-Y. Shum. Capturing and Rendering Geometry Details for BTF-Mapped Surfaces. *Visual Computer*, 21:559–568, 2005.

[Wang06] J. Wang and K. J. Dana. Relief Texture from Specularities. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 28:446–457, 2006.

[Wei05] Yichen Wei, Eyal Ofek, Long Quan, and Heung-Yeung Shum. Modeling Hair from Multiple Views. *ACM Trans. Graph.*, 24(3):816–820, 2005.

[Weidenbacher06] U. Weidenbacher, P. Bayerl, H. Neumann, and R. Fleming. Sketching Shiny Surfaces: 3D Shape Extraction and Depiction of Specular Surface. *ACM Trans. on Applied Perception*, 3(3):262–285, 2006.

[Weise07] T. Weise, B. Leibe, and L. Van Gool. Fast 3D Scanning with Automatic Motion Compensation. In *Proceedings of IEEE Conf. on Computer Vision and Pattern Recognition*, 2007.

[Wolff90] L. B. Wolff. Polarization-Based Material Classification from Specular Reflection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12(11):1059–1071, 1990.

[Wolff91] L. B. Wolff and T. E.Boult. Constraining Object Features Using a Polarization Reflectance Model. *IEEE Trans. on*

              Pattern Analysis and Machine Intelligence, 13(7):635–657,
              1991.

[Wolff94]   L. B. WOLFF. Polarization Camera for Computer Vision with
              a Beam Splitter. *Journal of the Optical Society of America*,
              11:2935–2945, 1994.

[Woodham80]   R. J. WOODHAM. Photometric Method for Determining Sur-
              face Orientation from Multiple Images. *Optical Engineering*,
              19(1):139–144, 1980.

[Wu04]   TAI-PANG WU AND CHI-KEUNG TANG. Separating Spec-
              ular, Diffuse, and Subsurface Scattering Reflectances from
              Photometric Images. In *Proceedings of European Conference
              on Computer Vision*, pages 419–433, 2004.

[Wu05]   T. P. WU AND C. K TANG. Dense Photometric Stereo Using
              a Mirror Sphere and Graph Cut. In *Proceedings of IEEE Conf.
              on Computer Vision and Pattern Recognition*, pages 140–147,
              2005.

[Wust91]   CLARENCE WUST AND DAVID W. CAPSON. Surface Profile
              Measurement Using Color Fringe Projection. *Journal Ma-
              chine Vision and Application*, 4(3):193–203, 1991.

[Yamazaki07] S. YAMAZAKI, S. NARASIMHAN, S. BAKER, AND
              T. KANADE. Coplanar Shadowgrams for Acquiring Visual
              Hulls of Intricate Objects. In *Proceedings of International
              Conference on Computer Vision*, 2007.

[Yu05]   YIZHOU YU AND JOHNNY T. CHANG. Shadow Graphs and
              3D Texture Reconstruction. *International Journal on Com-
              puter Vision*, 62(1-2):35–60, 2005.

[Zhang00]   ZHENGYOU ZHANG. A Flexible New Technique for Camera
              Calibration. *IEEE Trans. on Pattern Analysis and Machine
              Intelligence*, 22(11):1330–1334, 2000.

[Zhang03]   L. ZHANG, B. CURLESS, AND S. M. SEITZ. Spacetime
              Stereo: Shape Recovery for Dynamic Scenes. In *Proceedings
              of IEEE Conf. on Computer Vision and Pattern Recognition*,
              pages 367–374, 2003.

[Zhang04]   LI ZHANG, NOAH SNAVELY, BRIAN CURLESS, AND
              STEVEN SEITZ. Spacetime Faces: High-Resolution Capture
              for Modeling and Animation. In *Proceedings of SIGGRAPH*,
              pages 548–558, 2004.

[Zhang05]   LI ZHANG. *Spacetime Stereo and Its Applications*. PhD the-
              sis, the University of Washington, Seattle, 2005.

[Zhang06a]   SONG ZHANG AND PEISEN S. HUANG. High-resolution
              real-time 3-D shape measurement, 2006. Opt. Eng., 2006.

[Zhang06b]   SONG ZHANG AND PEISEN S. HUANG. Novel method for
              structured light system calibration. *Opt. Eng.*, 45(8), 2006.

[Zhang06c]   SONG ZHANG, XIAOLIN LI, AND SHING-TUNG YAU.
              Multi-level quality-guided phase unwrapping algorithm for
              real-time 3-D shape reconstruction, 2006. Appl. Opt. 2006.

[Zhang06d]   SONG ZHANG AND SHING-TUNG YAU. Generic nonsinu-
              soidal phase error correction for 3-D shape measurement us-
              ing a digital video projector, 2006. Appl. Opt. 2006.

[Zhang06e]   SONG ZHANG AND SHING-TUNG YAU. High-Resolution,
              Real-Time Absolute 3-D Coordinate Measurement Based on
              the Phase Shifting Method. *Opt. Express*, 14:2644–2649,
              2006.

[Zheng00]   JIANG YU ZHENG AND AKIO MURATA. Acquiring a Com-
              plete 3D Model from Specular Motion under the Illumination
              of Circular-Shaped Light Sources. *IEEE Trans. on Pattern
              Analysis and Machine Intelligence*, 22(8):913–920, 2000.

[Zisserman89] ANDREW ZISSERMAN, PETER GIBLIN, AND ANDREW
              BLAKE. The Information Available to a Moving Observer
              from Specularities. *Image Vision Comput.*, 7(1):38–42, 1989.

Curriculum Vitae – Lebenslauf

Curriculum Vitae

1977	Born in Taizhou, Zhejiang, China
1984 – 1989	Fundamental school, Tianma
1989 – 1992	Junior school, Tianma
1992 – 1995	High school, Chumen
1995 – 1999	Bachelor study in Computer Science, Harbin Institute of Technology, China
1999	BEng in Computer Science
1999 – 2002	Master study in Computer Science, Beijing University of Technology, China
2002	MEng in Computer Science
2002 – 2004	Research assistant, University of Trier, Germany
2004 – 2008	Ph.D. Student at the Max-Planck-Institut für Informatik Saarbrücken, Germany