Diss. ETH No. 17823

Holography and Wave-Based Image Generation

A dissertation submitted to
ETH Zurich

for the Degree of
Doctor of Sciences

presented by
Remo Ziegler
Eidg.Dipl.Inf.Ing. ETH Zurich,Switzerland
born 17 October 1975
citizen of Switzerland

accepted on the recommendation of
Prof. Dr. Markus Gross, examiner
Prof. Dr. Marc Pollefeys, co-examiner

2008

Abstract

In this thesis, we present a framework to capture and store wavefronts of an arbitrary scene and render images based on a wave-based representation of light. Simulating light as a wave allows for a variety of effects including depth of field, diffraction, interference and features built-in anti-aliasing. We propose a novel framework featuring seamless integration into conventional rendering and display technology, enabling an elegant combination of traditional 3D object or scene representations with wave-based representations.

Two different ways of generating a wavefront from a scene are elaborated, namely a subdivision of artificial scenes into a number of elementary waves as well as a transformation of a light field captured under white light conditions into a wavefront. The proposed framework uses a hologram as the data representation of a wavefront and provides ways of wave propagating through free space as well as propagation through space containing partially planar occluders. Furthermore, images can be computed fairly easily anywhere in space using a physical model of a camera taking into account optical elements, such as lenses and apertures. The required view-dependent depth image is computed from the phase information inherently represented in the complex-valued wavefront and allows a combination with traditional rendering methods.

Kurzfassung

In dieser Arbeit präsentieren wir ein System, mit welchem man eine Wellenfront einer beliebigen Szene aufnehmen, speichern und als Bild wiedergeben kann. Die Bildgenerierung basiert auf einer Repräsentation des Lichtes als Wellen. Durch die Wellensimulation des Lichtes, können wir eine Vielzahl von Effekten erzeugen. Dazu gehören Tiefenschärfe, Beugung, Interferenz und eine implizite Lösung des Aliasing Problems. Wir präsentieren ein neuartiges System, welches sich nahtlos in bestehende Grafiksysteme und Bildwiedergebungssysteme integrieren lässt. Dabei werden traditionelle 3D Objekt Repräsentationen, mit wellenbasierten Repräsentationen kombiniert.

Zwei unterschiedliche Möglichkeiten zur Generierung einer Wellenfront von einer Szene wurden ausgearbeitet. Einerseits wird eine künstliche Szene in mehrere Elementarwellen aufgeteilt, welche als Überlagerung eine Wellenfront ergeben. Andererseits transformieren wir ein "Light Field", welches unter Weisslicht Konditionen aufgenommen wurde, in eine Wellenfront. Unser System verwendet ein Hologramm als Datenrepräsentation für eine Wellenfront. Wir simulieren die Propagation der Wellenfront durch den freien Raum, als auch durch einen Raum mit planaren Objekten. Durch die Verwendung eines physikalisch motivierten Kameramodelles, welches optische Elemente, wie eine Linse und Apertur beinhaltet, können wir aus der Wellenfront Bilder erzeugen. Das notwendige Tiefenbild wird von der der Wellenfront inhärenten Phaseninformation erzeugt. Basierend auf diesem Tiefenbild lassen sich anschliessend wellenbasierte als auch traditionelle strahlenbasierte Methoden kombinieren.

Acknowledgments

First and foremost, I want to especially thank my advisor Markus Gross. He approached me with the idea of using holography in computer graphics, and provided a lot of guidance, particularly during the first steps. His constant encouragement and very broad knowledge in computer graphics, as well as his expertise in paper writing, allowed me to focus on scientific questions without getting lost in details. Furthermore, I appreciate his open personality and sense of humor, which made working at the lab a pleasant experience.

Throughout this thesis, I had the chance to work with many excellent researchers and students. One of the things that stroke me most, is that not the best answers have the biggest potential, but the best questions do. One collaborator, who has brought this to perfection, and with whom I enjoyed many very detailed and long intellectual discussions, is Marc Levoy. Working with him has been a very enriching and inspiring experience. Additionally, his precision in structuring ideas and clear argumentation in paper writing has taught me a lot. I also want to thank Lukas Ahrenberg and Marcus Magnor, with whom I shared stimulating discussions about holography in computer graphics.

Since I chose a topic, which is not that common in computer graphics, I struggled from time to time to find a good sparring partner for in-depth discussions. However, I had the luck to work with some excellent master students, who showed an exceptional performance and enthusiasm during their master thesis. I want to especially thank Peter Kaufmann, Simone Croci, and Simon Bucheli. Furthermore, I had the chance to work together with Rolf Adelsberger, a new PhD student in our lab, who showed sincere commitment in my last years project.

I am grateful to Pascal Müller and Leonard McMillan, who provided useful data shortly before a tight deadline. Deadlines always required long working days, sometimes even until the early morning hours. Thanks to fantastic people in the lab, in-depth discussions, mutual help and an excellent social time has been a standard anytime. Having such a good environment has certainly helped me a lot. Let me especially thank my two office mates, Tim Weyrich and Simon Heinzle, with whom I had a blast sharing the office.

Last but not least, I would like to thank my wife Nathalie, who was always very understanding, when I spent nights and weekends at the lab. Not having this pressure made my life a lot easier.

Contents

CHAPTER 1

Introduction

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

1.1 Motivation

Traditionally, graphics objects or scenes have been represented through geometry and appearance. Most often, the former is described by geometric primitives including triangles, points, basis functions, and others, while the latter is encoded in texture maps. Due to the difficulty of representing some real-world objects, such as fur, hair, or trees, with traditional techniques, research has also focused on image based rendering using light-fields [LH96, WAA$^+$00, IMG00], lumigraphs [GGSC96, BBM$^+$01], reflectance fields [DHT$^+$00a, MPN$^+$02, Wey05], sprites, and other models. All of these graphics representations share the property that their treatment of light is motivated through ray optics and that rendering involves projection and rasterization, or ray tracing.

In recent years, significant progress has been made in the field of display technology with the goal to replace the 2D flat screen by more advanced and immersive 3-dimensional displays. Besides various autostereoscopic displays [Oko76, A.B01], there has been an increased interest in holographic displays [SCS05] which reconstruct the 3D object or scene wavefront directly from the underlying holographic representation. In spite of some current limitations, such as computational costs and sampling rates, the rapid development of compute power and data storage makes holographic representation, image generation, and display, a technology of greatest potential for the future of computer graphics, and interactive visual simulation.

We propose a novel framework for graphics representation, processing, and display which is entirely based on digital holography and wave optics, while leaving the possibility of combining it with traditional rasterized rendering. Holograms are elegant structures capturing the phase and amplitude of an object or scene wavefront as seen from all possible views through a window of given aperture. Based on wave optics, holography is mathematically considerably more demanding than geometric optics. The need for monochromatic, coherent illumination during acquisition and speckle patterns during display additionally seem to argue against considering holograms in the context of computer graphics rendering. On the other hand, holograms represent visual scene appearance in the most elegant way, containing any possible view from a continuous viewport region without aliasing.

A central feature of the proposed framework is its seamless integration into conventional, framebuffer-oriented 2D display technology as well as its compliance with future 3D holographic displays. The core component of our novel representation is a digital hologram which can be recorded both from real-world objects and from conventional, triangle or point-based graphics objects. Real-world data can be captured under white light conditions due to the introduction of a novel bidirectional light field - hologram transform, alleviating many illumination constraints of the recording step.

In order to reconstruct 2D images at a given camera position, the original wavefront has to be reconstructed from the hologram and propagated through space. To this end, we utilize fast, discrete approximation methods based on Fourier theory and angular spectrum, which have been described in a continuous form in [DH98] and in a refined way in [Mat05b]. We specifically study required resampling steps for this discrete angular spectrum propagation. A set of occluders, such as opaque planes, semitransparent planes or diffrative optical elements, interfering with free space propagation, can be modeled by a complex-valued attenuation function combined with a complex valued emissive wavefield. This provides the possibility of modeling fairly complex scenes, which can be rendered entirely wave-based. In addition, we will show that this wave-based approach allows for an elegant modeling of the camera optics. The introduction of a physically based camera model including a thin lens, allow optical effects, such as depth of field, varying focal length, chromatic abberation or diffraction, which would not easily be possible using a theoretical camera model as it is used in traditional rasterization systems.

1.2 Principal Contributions

The goal of this thesis consists of the creation of a wave-based image generation framework, which is based on a holographic data structure, providing possibilities to record, store, reconstruct and render entire wave-based scenes as well as combine wave-based and ray-based representations in the same scene. The contributions to this framework can be divided into four areas:

Recording and rendering of holograms The currently most flexible way of capturing a complex-valued wavefront, is the hologram. It allows to physically record and reconstruct a wavefront from a scene. We introduce the hologram as a data representation in a novel holographic graphics pipeline. This pipeline is the first one allowing the rendering of holographic representations as well as traditional geometric objects in the same scene. A novel depth reconstruction for the data stored on the hologram, allows for a composition of different scene objects using the depth buffer. This heterogeneous rendering of wave-based and ray-based representations can, therefore, be handled partially on the graphics card leading to a better performance.

Lighting and occlusion for wave-based objects Previous work handles occlusions for wave-based evaluation in a ray-based manner. Often, the bending of light is ignored or handled separately. We introduce a novel complex-valued wave-based occlusion simulating opaque, semi-transparent, and refractive objects based on scalar diffraction theory. This representation is useful not only for scene evaluation on a hologram, but also for holographic rendering, where different aperture dependent glare effects can easily be rendered. Furthermore, we show ways to evaluate the lighting for wave-based objects as well as the real-time lighting of a hologram rendering.

White light recording The hologram requires a static interference pattern to encode the wavefront reflected or emitted by the scene. A static interference pattern can be generated using monochromatic light. This requirement strongly restricts the acquisition setting for new scenes. Therefore, we introduce a novel bidirectional light field to hologram transform, which allows to indirectly capture holograms under white light conditions. As a part of this transformation, we present a novel and robust depth field reconstruction from the given light field. The introduction of a frequency minimization criteria leads to more robust depth reconstruction even for highly specular objects.

Spatially adaptive illumination Holograms can be captured under white light conditions, by capturing a light field and transforming it into a hologram. The light field can be captured in a single shot with a set of cameras or a modified camera featuring a lenslet array or mask in front of the

CCD. Low light conditions prevent a single shot hand held light field capture, without flash. An external photographic flash suffers from an uneven illumination of the scene, due to the quadratic light falloff. In this thesis, we present a spatially adaptive photographic flash, which compensates the low flash light for distant regions. A real-time depth acquisition is implemented by a time-of-flight camera. To reduce the noise in the depth values, we present a novel joint adaptive trilateral filter. Furthermore, we introduce a flash-tone-mapping, which enhances contrast in dark image regions, without overexposing highly reflective parts of the scene.

1.3 Thesis Outline

The thesis is organized in two main parts. The first part deals with the holographic graphics pipeline and wave-based rendering. The second part focuses on the acquisition of holograms under white light conditions, which can then be rendered using a method presented in the first part.

In Chapter 2, we give an overview of related work. Holography has been used and studied in many different ways, such that we are focussing the related work on computer generated holograms and digital holograms.

Throughout this thesis, light is being represented as rays and waves. There are different ways of representing the light as waves. We are using the scalar diffraction theory. This representation models the light accurately enough in order to simulate the recording and reconstruction of holograms. For reasons of completeness, a brief introduction into wave-based representation and propagation of light is given in Chapter 3.

In Chapter 4, we introduce the holographic rendering pipeline, which forms the core of this thesis. The holographic rendering pipeline consists of a hologram recording, a hologram reconstruction and a rendering step. The hologram is used as a data representation for wave-based objects. The pipeline is, however, not restricted to objects represented by a hologram, but also allows conventional object representations, such as triangular meshes or point-based objects. These conventional objects are evaluated based on ray-based optics and combined with wave-based rendered objects.

In Chapter 5, we present novel ways to evaluate wave-based objects featuring occlusion and lighting. We mainly differ between three types of evaluations, namely computer generated holography, wave-based rendering and hologram rendering. The computer generated holography (CGH) evaluates the light waves from an artificial object on the entire hologram. Since a hologram can be observed from various directions, the lighting and occlusion has to consider the variation for all the possible paraxial views. The wave-based rendering evaluates a scene given by artificial objects directly for one specific camera view. This allows for approximations of the evaluation without showing visible artifacts. The hologram rendering evaluates a scene, which is represented by a hologram. The lighting for such a scene can be adjusted in real-time using graphics hardware.

In Chapter 6, we describe a bidirectional light field to hologram transform. The light field and the hologram capture the amount of light, which propagates through a plane in space. The light field captures all the rays propagating through such a plane, while the hologram captures all the waves propagated through it. By the introduction of a bidirectional light field to hologram transform, we can switch from one representation to the other in order to profit from their respective advantages. To be able to capture a scene in a single shot under white light conditions, is one of the most dominant advantages of the light field, since it indirectly allows to capture a hologram under white light conditions. Thus, additionally to the CGH, which gives the possibility to record artificial objects on a hologram, we can also capture real scenes onto a hologram.

A light field and, due to the transform, also a hologram can be captured in a single shot, like a conventional photograph. To accommodate dark scenes, most digital cameras include electronic flash - either integrated with the camera or in a detachable unit. The same flash could be used to capture a light field of a scene under low light conditions. More expensive units offer control over brightness, duration, and angular spread. Regardless of how these parameters are chosen, flash illumination always falls off sharply with distance from the camera, making most flash pictures look unnatural. In Chapter 7, we describe a spatially adaptive photographic flash, which allows to compensate for the falloff on a per pixel basis. Although, we tested our prototype for conventional photography, it could also be used for the illumination of the scene of light field acquisition.

Finally, Chapter 8 concludes this thesis with a discussion of the entire pipeline and the results, and provides an outlook to possible future work.

1.4 Publications

In the context of this thesis, the following publications have been accepted.

R. ZIEGLER, S. CROCI, M. GROSS. Lighting and Occlusion in a Wave-Based Framework. In *Computer Graphics Forum (Proceedings of Eurographics 2008)*, Hersonissos (Crete), Greece, April 2008.

This paper presents novel methods to enhance Computer Generated Holography (CGH) by introducing a complex-valued wave-based occlusion handling method. This offers a very intuitive and efficient interface to introduce optical elements featuring physically-based light interaction exhibiting depth-of-field, diffraction, and glare effects.

R. ZIEGLER, S. BUCHELI, L. AHRENBERG, M. MAGNOR, M. GROSS. A Bidirectional Light Field - Hologram Transform. In *Computer Graphics Forum (Proceedings of Eurographics 2007)*, Prag, Czech Republic, September 2007.

In this work, we propose a novel framework to represent visual information. Extending the notion of conventional image-based rendering, our framework makes joint use of both light fields and holograms as complementary representations. We demonstrate how light fields can be transformed into holograms, and vice versa. *This paper received the Best Paper Award and the Best Student Paper Award.*

R. ZIEGLER, P. KAUFMANN, M. GROSS. A Framework for Holographic Scene Representation and Image Synthesis. In *IEEE Transactions on Visualization and Computer Graphics*, pages 403-415, March 2007.

This paper presents a novel holographic graphics pipeline consisting of several stages including the digital recording of a full-parallax hologram, the reconstruction and propagation of its wavefront, and rendering of the final image onto conventional, framebuffer-based displays.

R. ZIEGLER, P. KAUFMANN, M. GROSS. A Framework for Holographic Scene Representation and Image Synthesis. In *SIGGRAPH '06: Material presented at the ACM SIGGRAPH 2006 conference*, page 108, August 2006.

This publication sketches the holographic graphics pipeline.

R. ZIEGLER, P. KAUFMANN, M. GROSS. A Framework for Holographic Scene Representation and Image Synthesis. In *Technical Report#522, Computer Science Department ETH Zurich, Switzerland*, March 2006.

This publication presents a novel holographic graphics pipeline consisting of several stages including the digital recording of a full-parallax hologram, the reconstruction and propagation of its wavefront, and rendering of the final image onto conventional, framebuffer-based displays.

Additional publication written during this time period, but not directly related to this thesis.

S.-M. RHEE, R. ZIEGLER, J. PARK, M. NAEF, M. GROSS, M.-H. KIM. Low-Cost Telepresence for Collaborative Virtual Environments. In *IEEE Transactions on Visualization and Computer Graphics*, pages 156-166, January–February 2007.

This work presents a novel low-cost method for visual communication and telepresence in a CAVE™-like environment, relying on 2D stereo-based video avatars. The system combines a selection of proven efficient algorithms and approximations in a unique way, resulting in a convincing stereoscopic real-time representation of a remote user acquired in a spatially immersive display.

D. Cotting, R. Ziegler, M. Gross, H. Fuchs. Adaptive Instant Displays: Continuously Calibrated Projections Using Per-Pixel Light Control. In *Proceedings of Eurographics 2005*, pages 705-714, September 2005.

This paper presents a framework for achieving user-defined on-demand displays in setups containing bricks of movable cameras and DLP-projectors. A dynamic calibration procedure is introduced, which handles cameras and projectors in a unified way and allows continuous flexible setup changes, while seamless projection alignment and blending is performed simultaneously.

CHAPTER 2

Related Work

A lot of work has been done in holography since Dennis Gabor invented ”wavefront reconstruction” in 1948 [Gab48], which today is known as holography. Since then, holography has found an application in many different fields. Some applications include but are not limited to: holographic interferometry (i.e. measuring stress or subtle distance changes), optical devices (i.e. diffraction gratings or holographic lenses), security (i.e. CD-labeling, packaging or credit cards), holographic art (i.e. advertising or holographic museum), microscopy, storage or holographic displays.

Recently, the progress in computer technology provided the possibility to propagate light waves in space efficiently. This allows to combine holography and computer technology. The interference pattern usually captured by the holographic film, can be evaluated through computer generated holography (CGH) [MT00, KDS01, RBD$^+$99]. In CGH the scene is subsampled into different primitives such as points, lines, triangles or planes, who's wavefront is evaluated at the hologram. The computing power and the improvement of the resolution of LCD's and other spatial light modulators, allows to visualize the hologram on a digital holographic screen in real-time [LG95, Luc97]. Furthermore, the quickly improving resolution of digital cameras provides the possibility to digitally capture a hologram using a CCD-camera [SJ02]. This is known as digital holography. The maximum sampling of the interference pattern, and therefore the maximal frequency, is determined by the CCD's pixel pitch. The limitation of the maximal frequency of the interference pattern has a direct influence on the maximal field of view. It is limited to about 4° degrees with current commercial CCD-cameras. Digital holograms can be reconstructed numerically [Lie04] allowing to measure certain properties of the objects very precisely.

The techniques described above involve either a physical hologram for display, or a real laser light for recording. In this thesis we generate only *virtual holograms*, from which different viewpoints are being rendered using our novel pipeline. The chosen representation of the virtual hologram could, however, directly be used as an input for a physical holographic display.

The following sections discuss related work concerning the holography pipeline, wave-based rendering, light field - hologram transform and pixel-wise illumination control during recording.

2.1 Holography Pipeline

Although limited to horizontal parallax only holograms, Pappu et al. in [PSU+97] as well as Janda et al. [JHS06] presented a hologram rendering pipeline focused on hologram recording and image generation at the holographic plane. More recently, computer graphics hardware has been used to improve the evaluation time of the interference pattern for full-parallax holograms [ABMW06, IMV07, MIV07]. These pipelines focused on the evaluation of the hologram, and did not present ways of rendering the virtual hologram from multiple viewpoints in combination with conventional computer graphics objects.

2.1.1 Wave Propagation

Goodman [Goo68] gives a good introduction to fast paraxial light wave propagation of planar wavefronts based on the Fourier spectrum. Moravec presented some early work for the simulation of paraxial wave propagation for non-tilted planes in [Mor81]. Restrictions concerning tilted planes are lifted in [TB92, TB93, MSW03] while still applying paraxial propagation of waves. Delen and Hooker [DH98] extend Tommasi's system such that propagation to parallel planes becomes feasible while having the precision of Rayleigh-Sommerfeld. Fourier analysis has also successfully been applied to light transport based on ray optics by Durand et al. [DHS+05] where the influence of shading, occlusion and transport on the frequency content of radiance is studied. We compare two approaches for wave propagation. One is based on the fourier shift theorem, while the other is based on a propagation method presented in [SJ05]. The method presented by Schnars et al. [SJ05] consist of a convolution of the wavefield with a propagation function

$g(\cdot)$. We use the latter since the required padding before the Fourier Transform does not depend on the propagation distance as it does for the first method.

2.1.2 Hologram Recording

Fourier spectrum propagation has recently been applied by Matsushima [Mat05b, Mat05a] for digital hologram recording of textured objects. Similarly to the method presented in [TL93], the 3-D Fourier spectrum has been used for fast 2D slice projection for computer generated holography in [YMT04]. Further efficient propagation of objects consisting of points, lines or triangles were presented in [MT00, KDS01, RBD$^+$99, ABMW06, IMV07, MIV07] showing a clear quality loss in the hardware based approaches. Lucente et al. [LG95, Luc97] and Halle and Kropp [HK97] take advantage of graphics hardware to render higher quality holographic stereograms to a holographic screen achieving interactive rates while loosening the constraint of full-parallax. An excellent overview of holographic screens is given in [SCS05]. Our novel pipeline subdivides the given objects in point samples or partially planar patches. A fast propagation method for spherical waves has been implemented in hardware similar to [ABMW06]. Furthermore, the propagation method for planar patches presented in [SJ05] has been sped up by evaluating the Fourier transform in hardware using the CUDAFFT^{TM}.

2.1.3 Image Reconstruction

Digital reconstructions of captured holograms allow mathematical filtering [CMD00, KJ97] improving the final image quality. Filtered holograms can adopt negative or complex values and therefore not be usable for holographic screens anymore. Still, the digital reconstruction of the filtered hologram can be achieved as shown in [SJ02]. However, Schnars and Jüptner did not simulate a camera model with a lens showing limited depth of field. Recent interest in realistic camera models with limited depth of field has been shown in computer graphics in [IMG00, VWJL04, LCV$^+$04]. Ng [Ng05] even built a prototype of a hand-held plenoptic camera capturing a 4D light field of which a refocused 2D image can be calculated. Although, previous work for hologram reconstruction and rendering exists, and various models for ray-based camera renderings have been presented, our work is the first one to render images from a scene consisting of objects represented by holograms and conventional graphics representations. The images can be generated from arbitrary viewpoints using a wave-based camera.

2.2 Lighting and Occlusion

CGH moves towards a more realistic hologram creation by simulating scenes including complex geometry, texture, lighting and occlusion. The requirement of having to evaluate the lighting as well as the occlusion of the entire scene for every pixel of the hologram, still forms a major bottleneck for CGH.

2.2.1 Occlusion

Occlusion is typically either not handled at all, or the light is treated as rays in order to precompute the occlusion in the scene, such as in [JHS06]. Diffraction due to sharp occlusion boundaries is not simulated by the ray-based approach. Furthermore, a visibility query has to be computed for every new viewing position in order to best approximate the occlusion by ray-based light simulation. Moravec presented a wave-based occlusion approach for paraxial propagation through parallel planes [Mor81], while still using a hybrid wave/ray-based method for image creation. In 1990, Hilaire [HBL+90] mentions an occlusion handling of point sources by polygonal planes laying closer to the viewer. Matsushima [Mat05a] introduces a binary occluder for tilted planes in order to correctly treat scenes built from opaque planes. We extend the binary occluder to a novel complex-valued occluder, allowing the simulation of semi-transparent or refractive objects. Additionally, we can generate glare effects simply by applying a semi-transparent occluder at the aperture during wave-based image creation. Nakamae et al. [NKON90] as well as Kakimoto et al. in [KMN+04] presented a way of simulating glare effects. In [KMN+04] the glare effects are, however, only generated for bright points of the scene, which are composited with the rendered scene using billboards.

For objects with high geometric detail, subdivision into planes makes the wave-based occlusion computation very complex. Therefore, we propose to use point-sampled objects represented by surfels as in [PZvBG00] and in order to derive a better visibility evaluation. The scene is rendered by placing point sources at every visible surfel.

2.2.2 Lighting

In 1995, Lucente and Galyean introduced the holographic stereograms [LG95], where the lighting for different views is taken from the rendered parallax views. Matsushima precomputed the shading of the scene [Mat05b] by using a fixed viewpoint in the middle of the hologram, and adjusted the amplitudes of the primitives accordingly. More recently, Hanak et al. [IMV07] proposed an angular sampling of the scene using the GPU, where a local lighting model is evaluated per hologram pixel. Since we evaluate the scene using wave-based occlusion handling, the lighting model has to be evaluated for every scene point and not every angular sample. This being rather time consuming, we evaluate the lighting for a subset of the hologram pixels and interpolate for values in between. This approach could not be applied [IMV07], since the correspondences between angular samples of different viewpoints are not given.

In our proposed lighting method for hologram renderings, we make the assumptions that the BRDF can be freely chosen and the holographic rendering shows an ambient color. Although these are strong assumptions, they could be alleviated in future work by estimating the BRDF from different parallax views and geometry such as in [SWI97, YDMH99, LGK$^+$01, MGW01]. To the best of our knowledge, there has not been any previous work about lighting of hologram renderings.

2.2.3 Speckle Reduction

Speckle patterns are formed when fairly coherent light is either reflected from a rough surface, or propagated through medium with a random refractive index distribution. A thorough analysis of speckle pattern statistics is given in [DEF$^+$75] and more recently in [Goo06]. Reducing speckles by filtering a high resolution image has been proposed by [KL04] and [Goo68]. Different kind of filters such as local statistic filters [Lee81], sigma filters [Lee83], homomorphic filters [FP95], adaptive linear smoothing filters [FS81], filters based on wavelets [Lan99], morphology-based filters [Sch95] can be applied to reduce artifacts. Other methods control the phase randomization of the wave sources limiting the frequency band, while trying to achieve a smooth spectrum. One method is called the Iterative Fourier Transform Algorithm (IFTA) and has been subject of [WB88, WB89, WB91, BWB91, CSYS02, CSS03]. [WB88] discusses some problems inherent to the IFTA, namely the so-called stagnation problem and the choice of an adequate initial phase distribution. In [CSYS02] a version of the IFTA similar to the one in [WB88] is presented together with an effective technique for the generation of the initial phase distribution, called Binary Phase Difference Diffuser (BPDD). The generation of the initial phase distribution is the topic of [CSS03] where a new technique is presented and compared to BPDD.

Although the BPDD reduces the amount of speckles for a given setup, the minimization directly depends on the propagation distance of the wavefield. This distance is dependent of the camera distance to the hologram, which can vary depending on the chosen viewpoint. Therefore, we used either a phase randomization approach, a filtering approach or a combination thereof, to reduce the speckle artifacts.

2.3 Light Field - Hologram Transform

The light field and the hologram are related, as such as they are capturing all the light passing through a plane in space. The light field is capturing all the rays passing through the plane, ignoring any depth information. The hologram captures the propagated, complex-valued wavefront through the plane, while encoding the depth in the phase term. To fully understand the related work it is important to note, that the *light field* is also known as *integral photography*.

2.3.1 Light Field

In a paper in 1936 [Ger36], Gershun introduced the concept of light fields for the first time. He described it as the amount of light traveling in every direction through every point in space using light vectors. In 1996 Levoy and Hanrahan in [LH96] and Gortler et al. in [GGSC96] presented two similar practical ways of capturing and representing scenes for computer graphics independently, based on Gershun's theory. Many publications have drawn their work upon the light field as well as the lumigraph representation. Various publications focussing on sampling requirements [CCST00], rendering and filtering [IMG00, SYGM03], reducing ghosting artifacts and aliasing, as well as numerous capturing setups consisting of a camera array [WJV$^+$05, YEBM02] or one shot capturing devices such as in [Ng05, LNA$^+$06] keep exploiting the big potential of this field in various ways. A good overview of recent work in this field is presented in [Lev06].

2.3.2 Hologram from Images

Holograms can be computer generated from synthetic data and rendered either on a conventional display as in [ZKG07a] or rendered onto holographic displays as presented in [LG95, Mat05b]. A real scene can be captured on a holographic film or digitally captured by a CCD camera only if illuminated by monochromatic laser light. This is a severe restriction, since for many scenes the light cannot be controlled in such a meticulous way. DeBitetto presented a two-step model to record holographic stereograms under white light illumination in [DeB69]. A method involving the acquisition of objects from different points of view was suggested by Yatagi

[Yat76]. Halle studied the characterization of sampling-related image artifacts and presented different ways of reducing or eliminating aliasing artifacts of holographic stereograms in [Hal94]. The artifacts originate from using a plane as the depth approximation of the scene. More recently, the recording of holograms based on integral photography has found a wide interest [AR03, MOO06, SRS07, KSR07, SR08]. Since these methods do not reconstruct the scene's geometry, they can suffer from ghosting artifacts as described in [Hal94], if the sampling of the views from the integral photograph is not high enough. Lee et al. present two methods [LHKK65, LKLK06], which are based on a depth reconstruction from multiple images. They reduce ghosting artifacts due to a more accurate depth approximation, with the cost of increasing evaluation time. However, their depth reconstruction is based on a simple stereo algorithm, which can fail for non trivial cases. We present a novel and robust depth reconstruction for light fields, which allows the generation of high quality holograms.

2.3.3 Depth Reconstruction

Numerous publications deal with the problem of depth reconstruction from multi-view input. Many algorithms are based on the Epipolar-Plane Image (EPI) representation or on the related Epipolar Volumes, which were first introduced by Bolles et al. in [BBH87]. Although most of the work assumes Lambertian surfaces, various approaches remove specular effects such as [CKS+05, BN95, LLL+02] while few publications [DYW05, STM06] reconstruct surfaces with almost arbitrary BRDF. However, these methods require additional information about the reflection properties, assume light transport constancy requiring multiple acquisitions under different illumination, or are not using the full redundancy of a camera array used to capture a light field.

2.4 Illumination Control

Holograms can be captured under white light conditions, by capturing a light field and transforming it into a hologram. Since the light field can be captured in a single shot using a camera, the acquisition of holograms faces similar problems as conventional photography. One of these problems, is the illumination of the scene. Graphics and vision researchers have long recognized the importance of controlling illumination in computational photography methods. We can classify these methods according to the kind of illumination they employ:

Single Spatially Invariant Flash Petschnigg et al. [PSA+04] and Eisemann et al. [ED04] take two images - one with flash and one without, then use the low noise of the flash image to reduce noise in the noflash image. For this purpose they employ a joint bilateral filter (a.k.a. cross-bilateral filter). However, using a spatially invariant flash on a scene with a large depth extent can lead to underexposed regions in the flash image. These dark regions are noisy, thereby diminishing the quality of the detail that can be transferred to the corresponding regions of the noflash image. Using our novel spatially varying flash, we can ameliorate this problem.

Single Spatially Invariant Flash using Multiple Shots Several problems of flash images, such as strong highlights due to reflection of the flash by glossy surfaces, saturation of nearby objects, and poor illumination of distant objects are addressed by Agarwal et. al [ARNL05]. They capture a flash-noflash image pair, estimate depth as the ratio of the flash and noflash images, then inversely scale the gradients of each pixel in the flash image by this depth. The final image is computed by integrating this gradient field. However, separately scaling the intensities of individual pixels does not account for possible interreflections among features in the scene. By contrast, our spatially adaptive flash applies different illumination along each ray. This produces a physically valid scene, and it requires capturing only one image rather than a pair of images.

Programmable Illumination Using a projector-like light source is a powerful alternative to spatially invariant flash illumination. Nayar et. al [NKGR06] show how a programmable flash can be used to separate an image into direct and indirect reflections. However, they cannot compute how these reflections would change if only a part of the scene were lit - a necessary capability if one wishes to reduce illumination on nearby objects. Mohan et al. [MBW$^+$07] presents an inexpensive method for acquiring high-quality photographic lighting of desktop-sized static objects such as museum artifacts. They take multiple pictures of the scene, while a moving-head spotlight is scanning the inside of a foam-box enclosing the scene. This requires a controlled environment, which is typically not possible for candid shots. Other systems for programmatically controlled illumination include time-multiplexed lighting [Bel07] and Paul Debevec's Light Stages [DHT$^+$00b]. However, these methods are not suitable for candid photography.

Dynamic Range Many methods have been proposed for improving the dynamic range of an imaging system. These systems typically compromise either temporal or spatial resolution, or they employ multiple detectors. Nayar et. al [NB03] provides a good summary of previous work. They also present a way to enlarge the dynamic range of a system by adding a spatial light modulator (LCD) or DMD [NBB06] to the optical path. This enables them to extend the dynamic range of the sensor by varying the exposure on a per-pixel basis. Our work is complementary to this method. It permits us to attenuate light that is too bright, while also providing a way to add light to regions that would otherwise be too dark.

2.5 Holography and Wave-Based Image Generation

Despite some earlier work on holographic representations in graphics [LG95, Luc97, HK97] and computer generated holography (CGH) [SJ05, SCS05] this thesis for the first time presents a full framework for a holography-inspired graphics pipeline, which allows to generate, process and render holograms from synthetic and real 3D objects. Rendering holograms from synthesized objects allows using small holograms, since the wavelength can be chosen outside of the visible spectrum. Our main contributions stated above are both fundamental and practical. They include a summary of the relevant theory as well as a variety of algorithms and methods for the digital recording, wave propagation and 2D rendering of holograms, while elaborating the inherently aliasing free image generation by wave-based rendering. Some rendering and compositing examples demonstrate the versatility and quality of the presented methods.

CHAPTER 3

Fundamentals of Holography

This chapter gives a brief history of light studies, presents different light models and provides an introduction to scalar diffraction theory.

3.1 History

Light has always been an important source of life and has been studied from very early on already. Around 500 BC, Pythagoras thought of light as "particles", which were emitted from the objects themselves and travel in straight lines. Roughly a century later, Plato supposed, that vision was produced by rays of light, which originate in the eye and hit the object when being viewed. Aristotle was the first to conclude that light travels in "something like waves", while trying to compare light and sound. Since then, many philosophers and scientists such as Hero of Alexandria, Alhazen Ibn al Haitam, Roger Bacon, Leonardo da Vinci, Galileo Galilei, and many more, have devoted parts of their life to study reflection, refraction, image creation and invented the camera obscura.

According to [Goo68], the first accurate report of an experiment showing the phenomena of diffraction, was written by Grimaldi and published after his death in 1665. The corposcular theory was the accepted way of modeling the propagation of light at that time. A big step was made by Christian Huygens in 1678, where he stated the following principle: *every point that is hit by a wave is the origin of a spherical elementary wave*. Further development concerning wave theory stagnated during the 18^{th} century, due to the great influence of Isaac Newton, who favored the corposcular theory. However, in 1804 Thomas Young introduced the concept of interference and strengthened the wave theory of light. Augustin Jean Fresnel found a mathematical formulation of the findings of Huygens and Young in 1818. Various people, such as Gustav Kirchhoff, Rayleigh Sommerfeld, and William Lawrence Bragg have worked on the mathematical formulation of diffraction theory. It was not until 1948, that holography was invented by Dennis Gabor [Gab48], who called it "wavefront reconstruction" at that time. Relying on monochromatic light, further development of holography was almost non existent until the invention and improvement of laser technology. Emmet Leith and Juris Upatnieks presented a way of off-axis holograms in 1963, opening up many possibilities for practical holography. A further major breakthrough was realized by Stephen A. Benton in 1968. He presented a white-light transmission hologram, which could be mass-produced using an embossing technique.

3.2 Light Model

Most applications in computer graphics apply a *ray-based representation* of light (cf. Fig. 3.1). Objects are being modeled as triangles, points, splats, voxels, and implicit surfaces, which are projected based on ray-based optics onto the image plane. Modeling light with a simple ray-based representation, allows to simulate many effects of reflection and refraction.

Another, more complex representation, is to model the light based on the *scalar wave theory*. It models light as a wave with a scalar amplitude and phase, which are depending on space and time. Although this model is more complicated to evaluate, it allows to simulate diffraction and interference, which are required when simulating holograms. However, the scalar wave theory allows only to simulate the propagation of light through a linear, isotropic, homogeneous, nondispersive and nonmagnetic dielectric medium Sect. 3.3.

A more complete description of light is given by the electromagnetic field, which was introduced by Maxwell in 1860. His mathematical description of the electromagnetic field is known as the Maxwell's equations. They describe the light as *vectorial electromagnetic waves*, which allow to simulate effects of polarization, an effect, which cannot be simulated with the scalar wave theory.

The currently most accurate description of light is the one of *quantum optics*. The light is modeled by a stream of photons and hence inherently quantized.

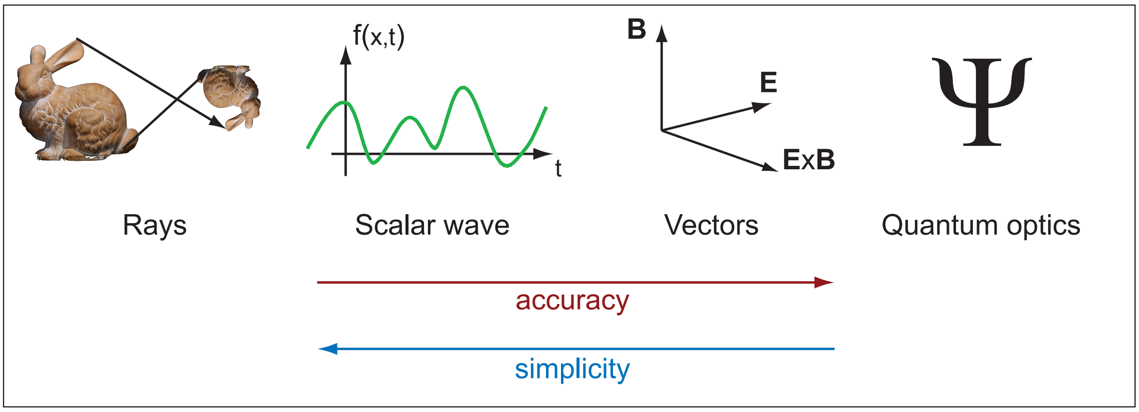

Figure 3.1: *Different physical representations of light. (The Bunny dataset is courtesy of Stanford University, Stanford 3D Scanning Repository .)*

This very complex representation of light is not required to simulate holography, but it is setting physical limitations of the setup used for hologram recording. For instance, it determines that the coherence length of the laser used for recording, has to be at least as big as to contain the object and the holographic plate.

Throughout the thesis, we are using the scalar wave theory to evaluate the propagation of waves Sect. 3.3. We also refer to it as *wave-based simulation*, or *wave-space simulation*. The visibility of objects in the scene are either evaluated using wave-based, or ray-based approaches. Finally, the illumination is evaluated using an optimized ray-based approach in order to speed up the required computing time.

3.3 Scalar Wave Representation

The Maxwell equations describe the relationship between the electric field $\mathbf{E} = (E_x, E_y, E_z)$ and the magnetic field $\mathbf{H} = (H_x, H_y, H_z)$. The scalar wave representation used in this thesis, can be interpreted as a special case of the Maxwell equations, if the propagation of the wave has to be simulated in a linear, isotropic, homogeneous, nondispersive and nonmagnetic medium. Important definitions:

Linear Medium A medium is linear if the response to several stimuli acting simultaneously is identical to the sum of the responses that each component stimulus would produce individually.

Isotropic Medium A medium is isotropic if its properties are independent of the direction of polarization of the wave.

Homogeneous Medium A medium is homogeneous if the permittivity is constant throughout the region of propagation.

Nondispersive Medium A medium is nondispersive if the permittivity is independent of the wavelength over the wavelength region occupied by the propagating wave.

Nonmagnetic Medium A medium is nonmagnetic if the magnetic permeability is always equal to μ_0, the vacuum permeability.

By imposing these constraints on the scene, we can simplify the Maxwell equations, and summarize the behavior of all components of **E** and **H** through a single scalar wave equation:

$$\nabla^2 u(P,t) - \frac{n^2}{c^2}\frac{\partial^2 u(P,t)}{\partial t^2} = 0. \tag{3.1}$$

$u(P,t)$ represents any of the scalar field components $E_x, E_y, E_z, H_x, H_y, or H_z$ dependent on the position P and time t. The refractive index of the medium is defined by n, while c is the velocity of propagation of light in vacuum.

If boundary conditions are imposed, a coupling between **E** and **H** is introduced, leading to some error even in a homogeneous medium. Fortunately, this error is only non-negligible in the vicinity of the aperture, but can be neglected completely many wavelength away from the aperture. This is the case when simulating holograms in our pipeline.

Based on the scalar wave theory, we describe the evaluation of spherical waves and plane waves at arbitrary points in space in Sect. 3.3.2, and elaborate the efficient propagation of a wavefront given on a surface S to an arbitrary point P' or to an arbitrarily placed surface $S_{A'}$ in the remainder of this section. This allows to simulate all the objects, given as an input to the holographic pipeline, as a set of spherical waves (i.e. point sampled objects) or partially planar surfaces (i.e. triangular meshes or apertures). Issues arising when evaluating the continuous equations in discrete space are discussed in Sect. 3.3.6.

3.3.1 Coherence

The best quality holograms are created using a coherent laser as an illumination source. There is a coherence in time and a coherence in space. We refer to a wave to be coherent in time if the phase difference between two arbitrary points along a ray path is independent of time. This is the case with a continuous, monochromatic wave radiated by a point source. We refer to a coherence in space, if the phase difference between two arbitrary points perpendicular to the ray path is independent of time. The envelope of the interference and therefore the quality of the hologram is dependent on the degree of coherence.

3.3.2 Elementary Waves

Without further specifications, we will present the wave equations for the case of purely monochromatic waves. A general time harmonic scalar wave $u(P,t)$ at position P and time t can be represented as

$$
\begin{aligned}
u(P,t) &= U(P)cos(\omega t) \qquad (3.2)\\
&= \Re\{U(P)e^{-\mathbf{i}\omega t}\} \qquad (3.3)
\end{aligned}
$$

with $\omega = 2\pi\nu$ being the angular frequency while ν describes the frequency or oscillation of the wave. The complex function $U(P)$, with the real-valued amplitude $A(P)$, and phase $\varphi(P)$ at position P is defined as

$$U(P) = A(P)e^{\mathbf{i}\varphi(P)} \ . \qquad (3.4)$$

The time-dependent term $e^{-\mathbf{i}\omega t}$ of Eq.(3.3) can be neglected, since the simulated light of our pipeline is of monochromatic nature. Therefore, the complex function $U(P)$ (sometimes called *phasor*) is sufficient to adequately describe $u(P,t)$ and is used as the wave representation in this thesis.

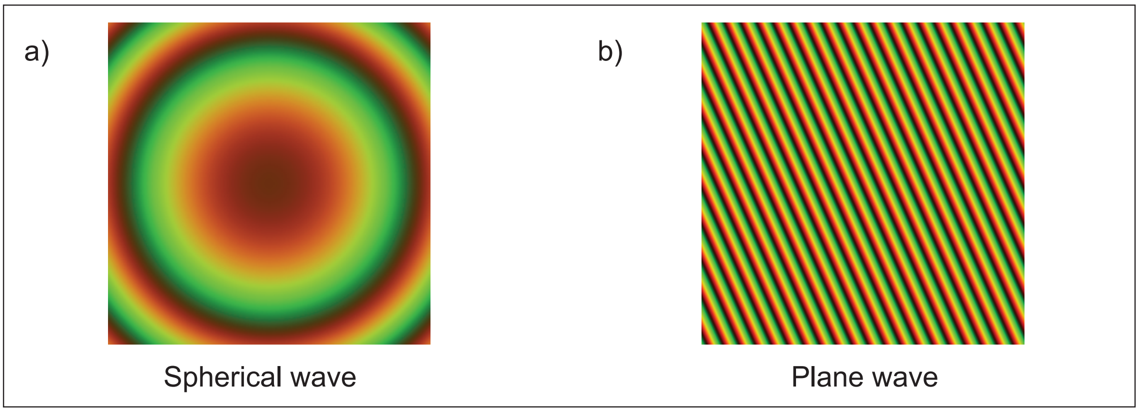

Figure 3.2: *a) Complex-valued amplitude of the evaluation of a spherical wave on a plane. b) Complex-valued amplitude of the evaluation of a plane wave on a plane. The real component is visualized in the green channel, while the complex component is shown in the red channel.*

A spherical wave (cf. Fig. 3.2a), originating at point P_0, is being represented as

$$U(P) = A_0 \frac{e^{\mathbf{i}kr+\varphi_0}}{r} \ , \qquad (3.5)$$

where A_0 is the real-valued amplitude at P_0, k is the wave number, φ_0 the initial phase at P_0, P is the position of evaluation and $r = ||P - P_0||_2$. The wave number k is defined as $k = \frac{2\pi}{\lambda}$ with λ being the wavelength.

A monochromatic plane wave (cf. Fig. 3.2b) can be represented as

$$U(P) = A_0 e^{\mathbf{ik \bullet r} + \varphi_0} \tag{3.6}$$

where A_0 and φ_0 are the real-valued amplitude and phase at the origin P_0 and $\mathbf{r}$ is defined as $\mathbf{r} = P - P_0$. The vector $\mathbf{k}$ is defined as $\mathbf{k} = k * (\alpha, \beta, \gamma)$ with k being the wave number and (α, β, γ) being the unit vector pointing in the direction of the wave propagation. The components of the vector are called *directional cosines*.

3.3.3 Interference

As shown in Sect. 3.3.2, we can represent the wave by a time-independent equation Eq.(3.4). If we substitute this representation into Eq.(3.1) we obtain the following formula:

$$(\nabla + k^2)U = 0, \tag{3.7}$$

known as the *Helmholtz equation*.

A scene can consist of many elementary waves, which can superimpose each other. The superposition is called interference. If we take two different sources $U_1(P)$ and $U_2(P)$, the interference $U(P)$ can be described as the sum of the two as $U(P) = U_1(P) + U_2(P)$. We can easily verify, that this sum still fulfills the Helmholtz equation Eq.(3.7). Even the scaling of any of the two components by a constant complex value fulfills Eq.(3.7). This allows to create entire scenes as a set of weighted elementary waves, which superpose each other.

Due to the monochromatic nature of the laser light, the interference pattern results in a standing wave, which can be recorded by a photosensitive film, such as the hologram. The spatially dependent amplitude of the interference pattern varies depending on the amplitude and phase of the interfering sources. It can reach from 0, in the case of *destructive interference*, to the sum of all the amplitudes of the contributing sources, in case of *constructive interference*. The constructive interference occurs, when all the sources are in phase, while the destructive interference occurs, when the sources are out of phase, such that the waves cancel each other out.

3.3.4 Wavefront Propagation

In 1678 Christian Huygens [Goo68] presented a way of describing the propagation of a wavefront as the "envelope" of "secondary" spherical waves being emitted by every point of the primary wavefront at time $t = 0$ (cf. Fig. 3.3). Later Gustav Kirchhoff put the theory on firmer mathematical

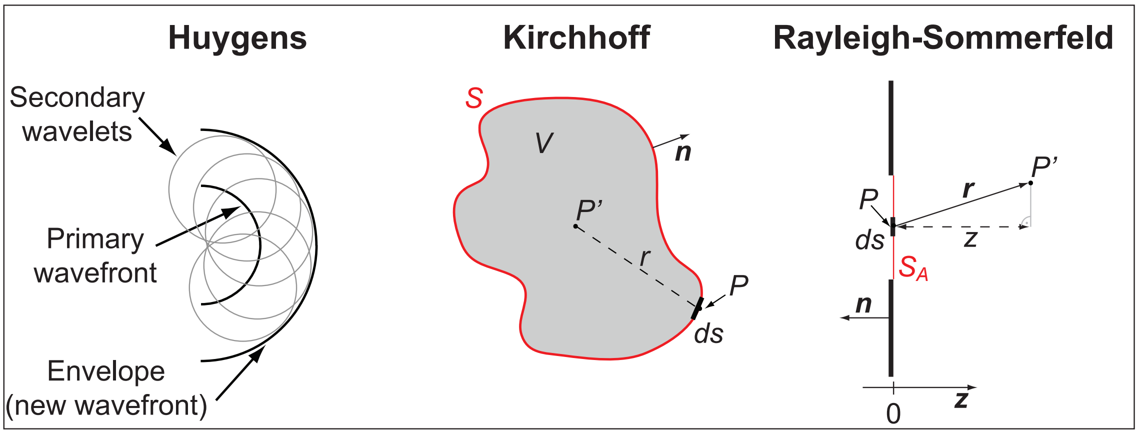

Figure 3.3: *Propagation of a wavefront based on Huygens, Kirchhoff's and Rayleigh-Sommerfeld's formulations.*

ground by finding a formulation to evaluate the wave at an arbitrary point P' in space knowing the evaluation and the derivative of the homogeneous wave equation on a closed surface S, enclosing volume V containing P' (cf. Fig. 3.3):

$$U(P') = \frac{1}{4\pi} \iint_S \frac{e^{ikr}}{r} \frac{\partial U(P)}{\partial \mathbf{n}} - U(P) \frac{\partial}{\partial \mathbf{n}} \left(\frac{e^{ikr}}{r} \right) ds \tag{3.8}$$

$$U(P') = \begin{cases} U(P') & \text{if } P' \in V \\ 0 & \text{if } P' \notin V \end{cases}$$

$\mathbf{n}$ is the surface normal, $U(P')$ is the evaluation at point P', whereas $U(P)$ is the evaluation of the wave at $P \in S$. The so-called *Rayleigh-Sommerfeld* theory can be derived from the Kirchhoff formula with the difference of only imposing boundary conditions on the evaluation of the wave function or on its derivative. The limitation of only being valid for planar geometries is not a real restriction, since the hologram as well as different primitives of synthetic objects like triangles and rectangles are planar as well. Assuming point P' away from the immediate vicinity ($k \gg 1/r$) of the aperture S_A , a propagation in the positive $\mathbf{z}$ direction perpendicular to S_A with S_A at $z = 0$

and $P \in S_A$ leads to the Rayleigh-Sommerfeld formula

$$U(P') = \frac{1}{2\pi} \iint_{S_A} U(P) \frac{\partial}{\partial \mathbf{n}} \left(\frac{e^{\mathbf{i}kr}}{r} \right) ds \tag{3.9}$$

$$U(P') = \frac{\mathbf{i}}{\lambda} \iint_{S_A} U(P) \frac{e^{\mathbf{i}kr}}{r} (\mathbf{r} \bullet \mathbf{n}) ds \quad \text{, with } \frac{k}{2\pi} = \frac{1}{\lambda} \tag{3.10}$$

The Rayleigh-Sommerfeld equation Eq.(3.10) can be interpreted from a physical point of view as a superposition of spherical waves $\frac{e^{\mathbf{i}kr}}{r}$ located at the aperture S_A with amplitude $\frac{U(P)}{\lambda}$multiplied by a phase shift of 90° caused by the multiplication of $\mathbf{i}$. Additionally the spherical waves are multiplied by a directional factor $(\mathbf{r} \bullet \mathbf{n})$. For further details and derivations please see [Goo68] and included references.
Evaluating these general forms of scalar diffraction theory can be very time consuming. There exist two approximations: one for the *near field* called the *Fresnel approximation* and one for the *far field* called the *Fraunhofer approximation*. As shown in [Goo68] the Fresnel approximation can be efficiently calculated by a convolution, while the Fraunhofer approximation can be interpreted as a Fourier Transform.

However, direct integration as well as the approximations have the deficiency of either being inefficient or inaccurate. Therefore, we compare two accurate and efficient methods of propagation. The first method was firstly presented by Delen and Hooker [DH98] and refined by Matsushima [Mat05b]. They present an accurate and fast propagation based on angular spectrum propagation, presented in [Rat56, Goo68], and the Fourier shift theorem. We will refer to this propagation as the *Fourier shift propagation*. The second method was presented in [SJ05] and is based on a convolution of the wavefront with a propagation function. We will refer to this propagation as the *convolution propagation*. Since the angular spectrum plays an important role in both propagation methods, it is introduced in Sect. 3.3.5.

3.3.5 Angular Spectrum

The complex monochromatic planar field $U(x,y,0)$ given at an aperture S_A can be split into multiple uniform infinite plane waves traveling in different directions (cf.Fig. 3.4a). The amplitude of everyone of those plane waves can be found by applying a Fourier Transform on $U(x,y,0)$, leading to the angular spectrum $\mathcal{A}(\nu_x,\nu_y,0)$. Therefore, the angular spectrum $\mathcal{A}(\nu_x,\nu_y,0)$ is an assembly of all plane waves defined as $e^{-2\pi\mathbf{i}(\nu_x x+\nu_y y)}$ in Eq.(3.11). The frequencies ν_x and ν_y can be substituted by the directional cosines α and β

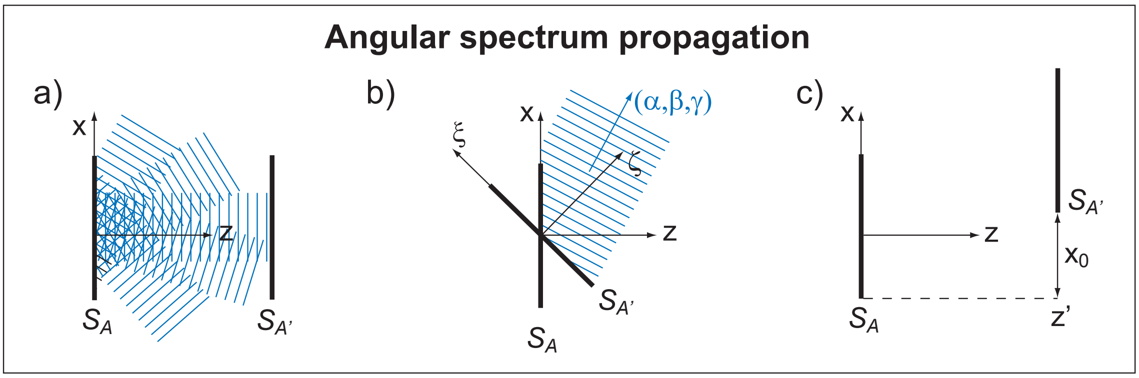

Figure 3.4: *The angular spectrum splits a wave field in multiple plane waves. This can be used to propagate a wave from an aperture S_A to an aperture $S_{A'}$. For an arbitrary placement of $S_{A'}$ "straight" propagation a), rotation of the spectrum b) and propagation on parallel offset planes c) is needed. The figure shows the 2D case, where the third dimension can be imagined to be perpendicular to x and z.*

of the plane wave (Eq.(3.6)) as $\nu_x = \frac{\alpha}{\lambda}$ and $\nu_y = \frac{\beta}{\lambda}$ leading to Eq.(3.12).

$$\begin{aligned} \mathcal{A}(\nu_x,\nu_y,0) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} U(x,y,0)e^{-2\pi\mathbf{i}(\nu_x x+\nu_y y)}dxdy \\ &= \mathcal{F}\{U(x,y,0)\} \end{aligned} \tag{3.11}$$

$$\mathcal{A}(\frac{\alpha}{\lambda},\frac{\beta}{\lambda},0) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} U(x,y,0)e^{-2\pi\mathbf{i}(\frac{\alpha}{\lambda}x+\frac{\beta}{\lambda}y)}dxdy \tag{3.12}$$

Propagation to Paraxial Parallel Plane

To propagate the wave field along $\mathbf{z}$ to a parallel plane at distance z, the angular spectrum $\mathcal{A}(\nu_x,\nu_y,0)$ has to be multiplied by a propagation phase term:

$$\mathcal{A}(\frac{\alpha}{\lambda},\frac{\beta}{\lambda},z) = \mathcal{A}(\frac{\alpha}{\lambda},\frac{\beta}{\lambda},0)e^{\frac{2\pi}{\lambda}z\mathbf{i}\sqrt{1-\alpha^2-\beta^2}} \quad \text{for } \alpha^2+\beta^2<1 \tag{3.13}$$

In case of $\alpha^2+\beta^2>1$ the wave would evaluate to a real-valued so-called evanescent wave, and therefore, not carry any energy away from the plane at $z=0$. The propagation defined in [Goo68] is restricted to propagation between parallel planes sharing a common center. This is also known as a paraxial propagation.

Angular Spectrum Rotation

Using the rotation of the angular spectrum from plane S_A with coordinate system (x,y,z) to $S_{A'}$ with coordinate system (ξ,η,ζ) (cf.Fig. 3.4b) as proposed by [TB92], lifts the restriction of propagation onto parallel planes. The rotational matrix M transforming (ξ,η,ζ) into (x,y,z) also relates the corresponding spatial frequencies $(\nu_\xi,\nu_\eta,\nu_\zeta)$ to (ν_x,ν_y,ν_z) by

$$(\nu_\xi,\nu_\eta,\nu_\zeta)^T = M^T(\nu_x,\nu_y,\nu_z)^T \; . \tag{3.14}$$

The coordinate transformation causes a shift of the sampling points requiring interpolation of the values for the new coordinate system. Additionally, backward propagating waves are set to 0. According to [DH98] linear interpolation is appropriate to find amplitudes at intermediate points.

Propagation to Parallel Plane With Offset

The propagation between two non-paraxial planes can be done by two different methods, as mentioned in Sect. 3.3.4. One is the Fourier shift propagation and the other is the convolution propagation.

Fourier Shift Propagation If the parallel plane $S_{A'}$ is offset to the propagation direction of S_A by $(x_0,y_0,0)$ as depicted in Fig. 3.4c), the propagation can be calculated by means of the Fourier shift theorem [DH98] as

$$U(x,y,z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{\mathcal{A}}(\nu_x,\nu_y,0)e^{\mathbf{i}2\pi z\sqrt{\frac{1}{\lambda^2}-\nu_x^2-\nu_y^2}} \times$$
$$\times\, e^{\mathbf{i}2\pi(\nu_x x+\nu_y y)}d\nu_x d\nu_y \; , \tag{3.15}$$

where

$$\hat{\mathcal{A}}(\nu_x,\nu_y,0) = \mathcal{A}(\nu_x,\nu_y,0)e^{\mathbf{i}2\pi(\nu_x x_0+\nu_y y_0)} \tag{3.16}$$

expresses the Fourier shift theorem. Using the propagation of the angular spectrum in continuous space yields to the identical result as the Rayleigh-Sommerfeld (RS) formula Eq.(3.9). A practical implementation affords a discretization, leading to different problems which are treated in Sect. 3.3.6.

Convolution Propagation If the plane S_A is chosen to lie in the $z=0$ plane, and the wave is propagating in the positive z-direction, we can simplify Eq.(3.10) to:

$$U(P') = -\frac{ik}{2\pi}z\iint_{S_A} U(P)\frac{e^{ikr}}{r^2}ds. \tag{3.17}$$

Let us rewrite this equation with $P = (x,y,0)$ and $P' = (x',y',z')$:

$$U(x',y',z') = -\frac{ik}{2\pi} z \iint_{S_A} \underbrace{U(x,y,0)}_{U} \underbrace{\frac{e^{ik\sqrt{(x'-x)^2+(y'-y)^2+z'^2}}}{(x'-x)^2+(y'-y)^2+z'^2}}_{G} dxdy. \tag{3.18}$$

As shown in [SJ05], Eq.(3.18) can be interpreted as a convolution, where the wave function U is convolved with the propagation function G. A practical implementation of RS using convolution affords a discretization leading to different problems which are treated in Sect. 3.3.6.

3.3.6 Discrete Propagation

The Fourier Transform Eq.(3.11) and the continuous propagation Eq.(3.15) are defined as an integral over an infinite plane, whereas in the discrete case the Discrete Fourier Transform (DFT) and the summation is restricted to a finite size leading to artifacts. Furthermore, the energy conservation does not hold, since the wave is not being evaluated over an infinite plane after propagation, and therefore, decreases with an increasing distance z [MSW03]. By zero-padding the plane before propagation those artifacts can be limited.

For the sake of simplicity the derivation for discrete propagation to parallel planes with offset (DPPO) will be based on a propagation between two squares but the extension to rectangles is straightforward. We present the discrete case of the Fourier Shift propagation and the convolution propagation in the following paragraphs. Using these two methods for DPPO instead of the direct integration of RS, allows to reduce the complexity of the propagation from $O(N^4)$ to $O(N^2 \log N)$. The complexity of DPPO differs solely by a constant factor from the Fresnel and Fraunhofer approximation while still evaluating the propagation with the precision of RS. For general propagation between non-parallel planes, we can neglect the complexity of the rotation step described in Sect. 3.3.5.

Discrete Fourier Shift Propagation The aperture S_A is sampled using a uniform grid of $N \times N$. The discretized wave field on S_A is given as U_{xy}. It is Fourier transformed using DFT leading to the angular spectrum $\mathcal{A}_{ml}$. The propagated wave field U'_{xy} can be evaluated at distance z based on Eq.(3.15) in the discrete case as

$$U'_{xy} = \frac{1}{N^2} \sum_{m=0}^{N-1} \sum_{l=0}^{N-1} \hat{\mathcal{A}}_{ml} e^{\mathbf{i}2\pi z \sqrt{\frac{1}{\lambda^2} - \left(\frac{m}{N}\right)^2 - \left(\frac{l}{N}\right)^2}} e^{\mathbf{i}2\pi(\frac{m}{N}x + \frac{l}{N}y)} \tag{3.19}$$

for $x = 0..N-1$ and $y = 0..N-1$ with

$$\hat{\mathcal{A}}_{ml} = \mathcal{A}_{ml} e^{\mathbf{i}2\pi(\frac{m}{N}x_0 + \frac{l}{N}y_0)}. \tag{3.20}$$

Since the term of multiplication to shift the spectrum is not zero-padded and does not wrap around, repetitive artifacts can be observed some distance away from the aperture (cf. Fig. 3.5a). By increasing the zero-padding of the initial wavefield with increasing propagation distance, these artifacts can be minimized (cf. Fig. 3.5b). The efficiency of propagation is, however, decreasing with increasing propagation distance due to the required bigger zero-padding and DFT.

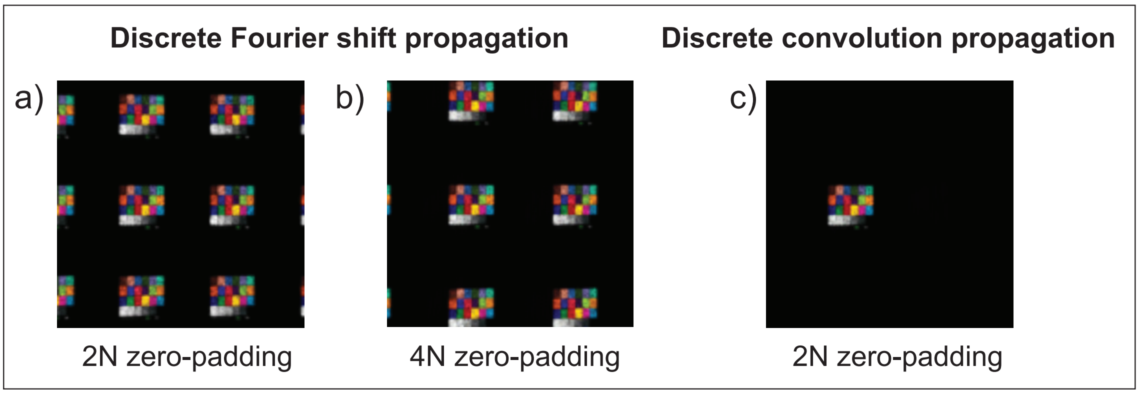

Figure 3.5: *a) Repetition artifacts are visible when using a* $2N$ *zero-padding of the aperture before propagation. b) The artifacts are smaller with* $4N$ *zero-padding and could be eliminated with an even bigger zero-padding with the cost of a decreasing efficiency. c) The discrete convolution propagation handles the propagation as a cyclic convolution requiring only a* $2N$ *zero-padding independent on propagation distance.*

Discrete Convolution Propagation The propagation of an evaluation $U[x,y]$ given on a uniform sampled grid $N \times N$ with a sample size $\delta = \frac{W}{N}$ to a parallel plane with the evaluation $U_{z'}[x',y']$, can be written as in Eq.(3.21) based on Eq.(3.17). The vector r can be written as $r = \sqrt{(\delta(x'-x)+x_0)^2 + (\delta(y'-y)+y_0)^2 + z'^2}$ considering a displacement x_0 and y_0 between the source and the target square. Equation 3.21 can be rewritten as Eq.(3.22).

$$U_{z'}[x',y'] = \frac{z'}{i\lambda}\delta^2 \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} U[x,y]\frac{e^{ikr}}{r^2} \tag{3.21}$$

$$U_{z'}[x',y'] = \frac{z'}{i\lambda}\delta^2 \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} U[x,y]G[x'-x,y'-y] \tag{3.22}$$

This can be interpreted as a discrete convolution of U and G resulting in $U_{z'} = \frac{z'i}{\lambda}\delta^2(U * G)$ with δ being the size of a sample. The convolution can be efficiently evaluated by a multiplication in the Fourier domain. However a multiplication in the discrete Fourier domain corresponds to a cyclic convolution imposing a zero-padding of a width of at least $(N-1)$ of U and G leading to $\hat{U}$ and $\hat{G}$. The propagation based on discrete convolution propagation can therefore be written as Eq.(3.23).

$$\hat{U}_{z'} = \frac{z'}{i\lambda}\delta^2 N \mathcal{F}^{-1}\{\mathcal{F}\{\hat{U}\} \cdot \mathcal{F}\{\hat{G}\}\} \tag{3.23}$$

The zero-padding is not dependent on the propagation distance, since the propagation function is also being zero-padded, and therefore, suppressing a wrap-around of the signal. For the rest of the thesis, we are using the convolution propagation for propagations of wavefields given on parallel planes.

Discrete Propagation Pipeline So far we summarized one method for rotation or tilting in Sect. 3.3.5 and one for propagation in Sect. 3.3.6. These two methods can now be combined to evaluate a wave originating on S_A on any arbitrarily placed aperture $S_{A'}$. The accuracy of the result relies on the sampling of S_A and $S_{A'}$ and on the order of the tilt and the propagation step. In case propagation and tilt would be needed, the tilt operation has to be evaluated first. For minimal sampling requirements refer to Sect. 3.5.1.

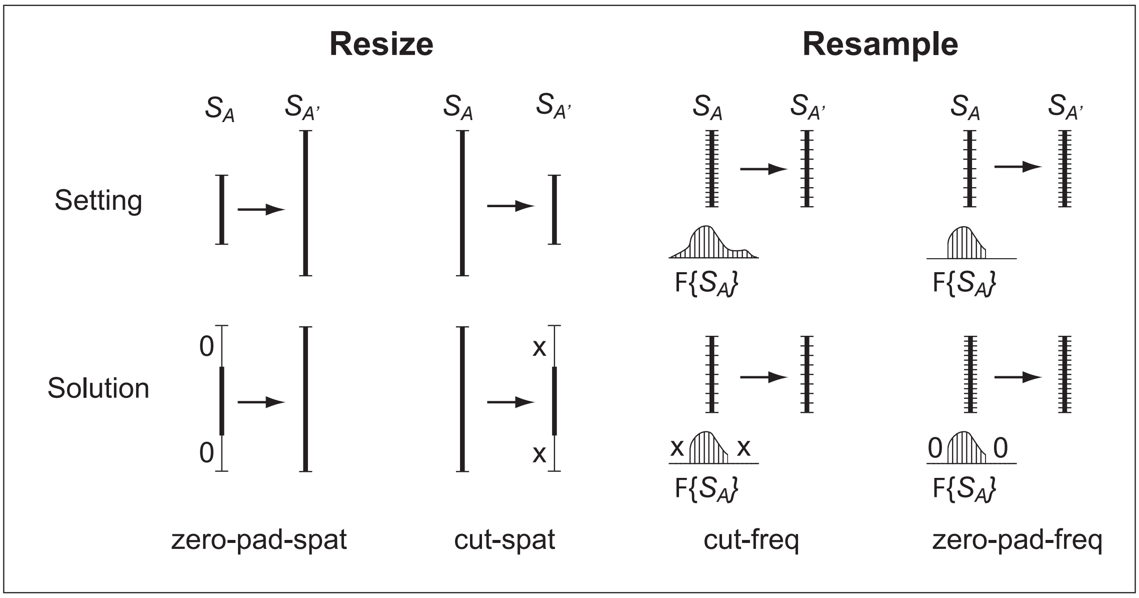

Figure 3.6: *The top row depicts the setting of S_A and $S_{A'}$ before modification. The bottom row shows the solution for the matched size after zero-padding in the spatial domain, the cutting in the spatial domain, the adjusted sampling by zero-padding in the frequency domain, and cutting in the frequency domain.*

Since the source aperture S_A does not necessarily have the same resolution and size as $S_{A'}$, re-sampling or resizing might be necessary. The four possible cases and their solution are depicted in Fig. 3.6. The sizes of the apertures S_A and $S_{A'}$ are preferably chosen as a power of 2 in order to guarantee an efficient FFT.

3.4 Hologram

A full-parallax hologram can be imagined as a window revealing limited viewing possibilities from a 3D scene (cf. Fig. 3.7). There exist many different holograms, being able to capture different properties of complex waves, such as phase, amplitude or intensity. Different setups and different shapes increase the variety even further [Goo68]. We are simulating a transmission hologram as in Fig. 3.7b, a setup, where the laser and the object are placed on the same side of the finite planar hologram during recording. The setup for a transmission hologram is depicted schematically in Fig. 3.8. The reconstruction of this type of hologram requires a laser placed at the same relative position to the hologram and having the same wavelength as for recording. The reconstruction of a hologram is shown in Fig. 3.9c. For the remainder of the thesis we will refer to a transmission hologram simply as *hologram*. The hologram is a thin real-valued and densely sampled plane measuring the static wavefield intensity, created by interference of multiple monochromatic waves. The hologram can be imagined as a high resolution photosensitive film. In a physical setup of a hologram, the laser light source is used to illuminate the object as well as the hologram, in order to guarantee monochromatic nature of the light waves. The laser can be split up in these two parts using a beam-splitter as depicted in Fig. 3.8. The light from the object is called the *object wave*, whereas the light from the laser during recording, is called *reference wave*, and during reconstruction *reconstruction wave*. The following two sections will conceptually describe the recording of a hologram in Sect. 3.4.1 as well as the reconstruction of the object wave in Sect. 3.5. Capturing the interference pattern affords a very high sampling rate of the hologram. Sampling rate requirements will be discussed in Sect. 3.5.1.

3.4.1 Hologram Recording

Two incident light waves and a holographic plate are needed to generate a hologram. One of the light waves is a plane wave Eq.(3.6) and stems from the laser source. We refer to this wave as reference wave R. Another wave stems from the object and is referred to as object wave O. In a physical setup, the object wave has to be created by the reflection of the illuminating laser light as depicted in Fig. 3.8, in order to guarantee coherent lighting of the entire scene. Remember that this is the requirement to generate a

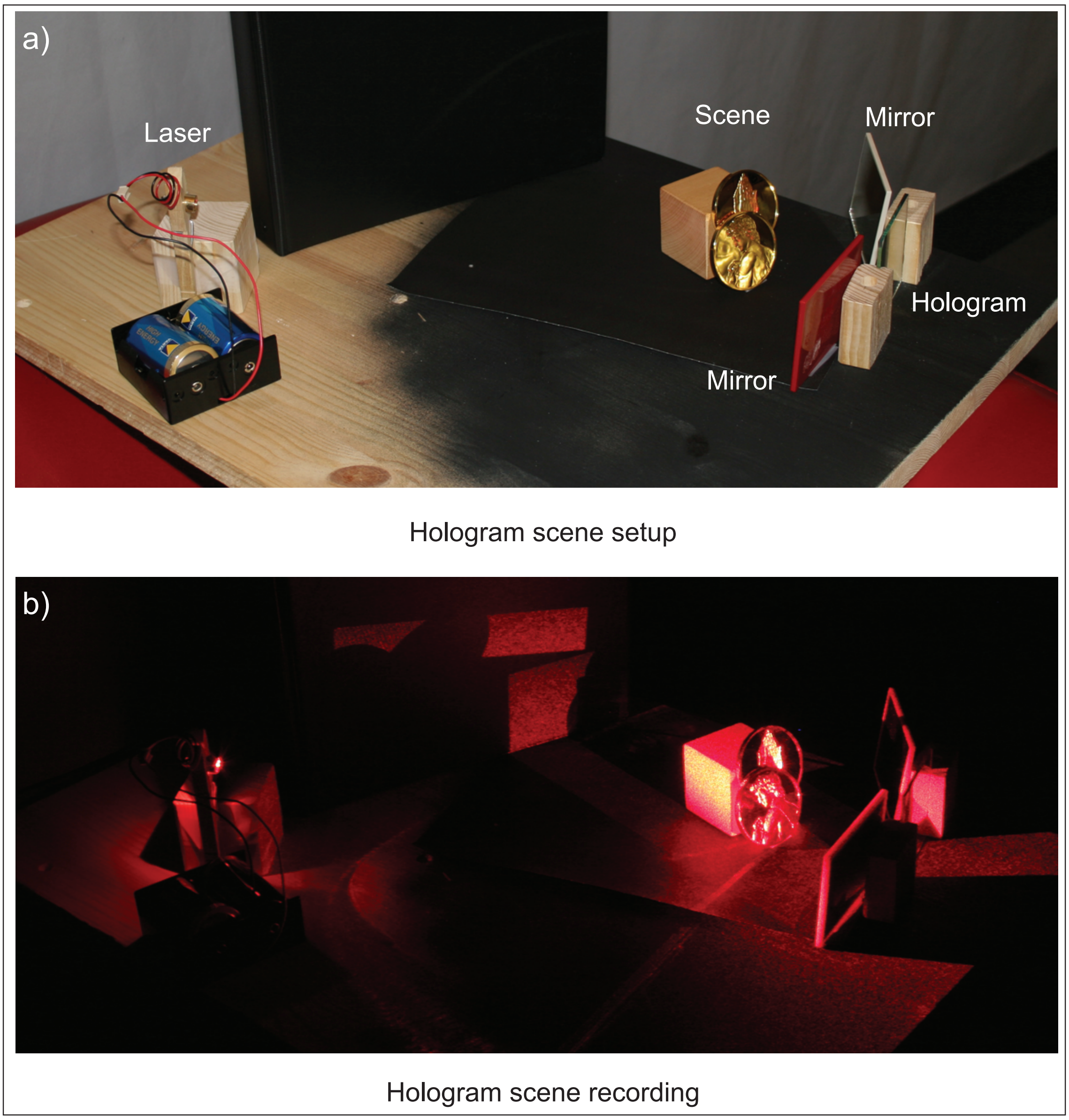

Figure 3.7: *a) A very simple setup to capture a hologram. Instead of using a beam splitter we used two mirrors, reflecting the laser light onto the object. b) This is the same setup as in a), but seen as it appears during recording. Only the laser light is illuminating the scene and the hologram.*

static interference pattern, which can be captured by the holographic film. In a theoretical setup however, we do not have to simulate the reflection of the laser light at the object. The theoretical approach allows assumptions of monochromatic laser properties implying spatial and temporal coherence and no motion artifacts. Monochromatic light can easily be simulated by choosing the same wavelength λ for O and R. The object wave O can be simulated by the superposition of a set of emitting light sources, which can be elementary waves Sect. 3.3.2 or wavefields Sect. 3.3.4. Figure 3.10a shows

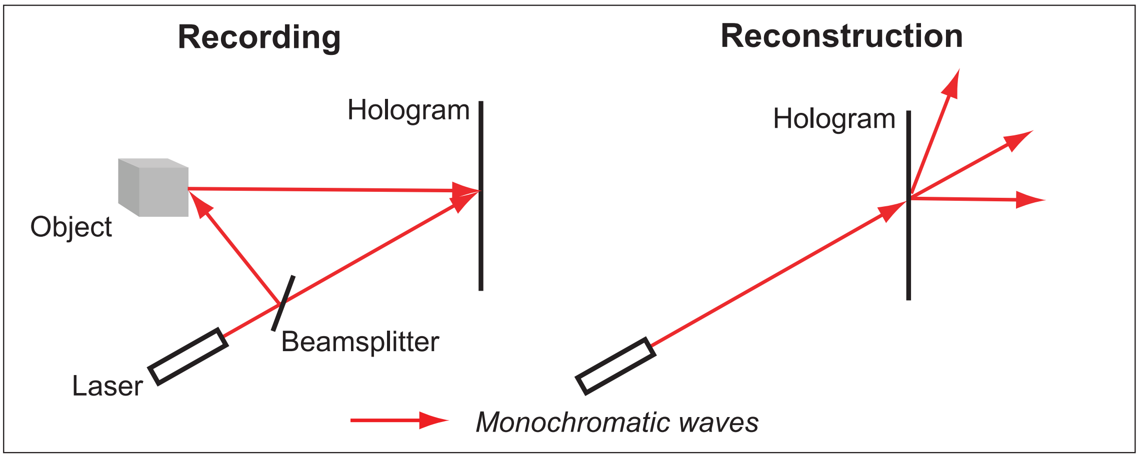

Figure 3.8: *Possible setup of a real transmission hologram using a beam-splitter to warrant monochromatic waves.*

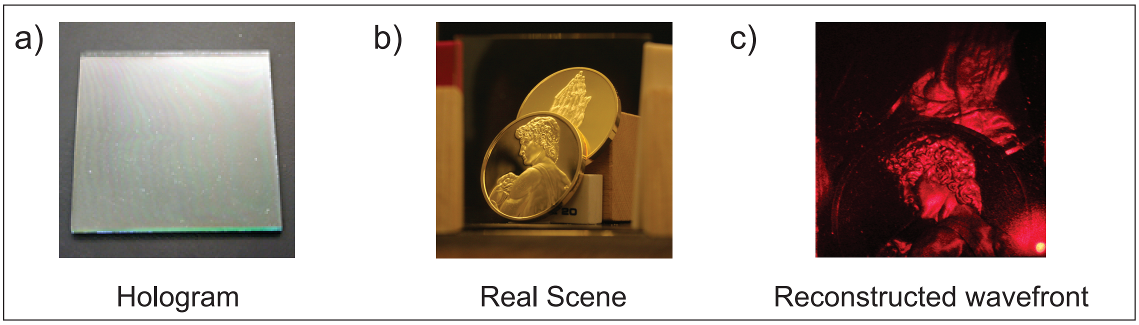

Figure 3.9: *a) Photosensitive holographic plate on which the hologram is being captured. b) The real scene, which is captured on the hologram. c) The object wave is reconstructed using a laser beam.*

the applied setup for hologram recording. The two light waves O and R are mutually coherent, leading to a time independent interference pattern in space. The intensity of this interference pattern is being recorded by a holographic plate as function H:

$$H = |O + R|^2 \tag{3.24}$$

Methods, as presented in Sect. 3.3, can be used to evaluate the object wave at the hologram in an efficient manner. Examples of the evaluation of different object types are described in Sect. 4.3.

Note, that the hologram H is a real-valued function, from which we would assume not to be able to reconstruct the amplitude and phase unambiguously. In the next section, however, we show how to extract the required information we need to reconstruct the complex-valued object wave.

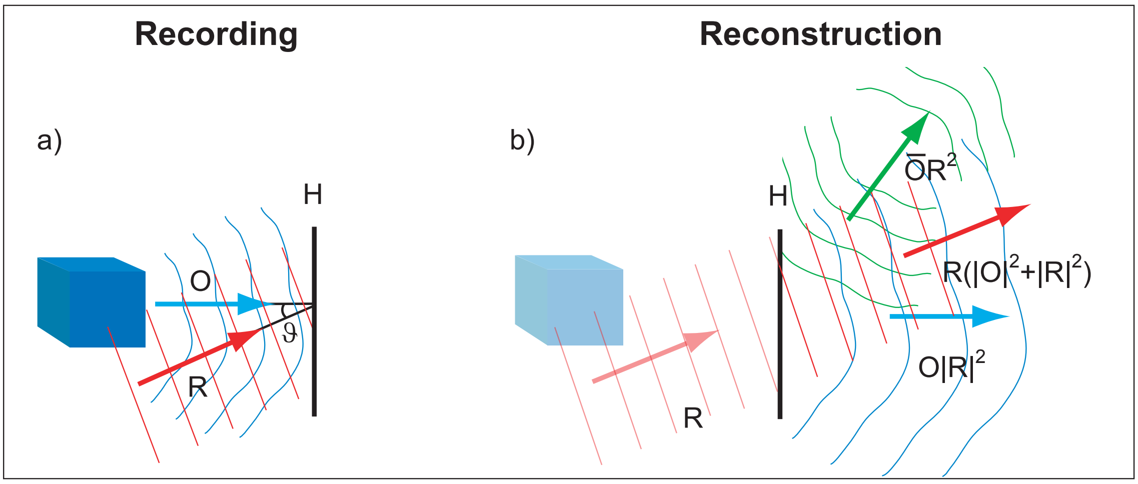

Figure 3.10: (a) *Hologram recording by interference of object wave O and reference wave R.* (b) *Hologram reconstruction using the reconstruction wave R resulting in the reference term* $R(|O|^2 + |R|^2)$*, the real image* $\bar{O}R^2$ *and the virtual image* $O|R|^2$.

3.5 Hologram Reconstruction

The holographic material is assumed to provide a linear mapping between the recorded intensity and the amplitude transmitted or reflected during the reconstruction process. A reconstruction wave R', identical to the former reference wave R, illuminates the holographic plate and is modulated by the transmittance T

$$T = t_b + \beta' H = H, \text{ with } t_b = 0 \text{ and } \beta' = 1\,. \tag{3.25}$$

Thereby, t_b is a "bias" transmittance which can be set to 0 and β' is a scaling factor which will be set to 1 in our theoretical setup leading to $T = H$. The incident reconstruction wave R' is, therefore, modulated by H leading to a reconstructed wave U as:

$$\begin{aligned} U = R' \cdot H = R'(|O|^2 + |R|^2) \quad &+ \quad O\bar{R}R' \quad &+ \quad \bar{O}RR' \\ = \underbrace{R(|O|^2 + |R|^2)}_{reference\ term} \quad &+ \quad \underbrace{O|R|^2}_{virtual\ image} \quad &+ \quad \underbrace{\bar{O}R^2}_{real\ image} \end{aligned} \tag{3.26}$$

Reference Term The reference term $R(|O|^2 + |R|^2)$, can be approximated as a scaled version of the reference wave. This assumption is true if $|O|^2$ and $|R|^2$ are constant over the whole hologram. For R being a planar wave this is trivial. However, for O the assumption is only true if the object is not close to the hologram or if $|O|^2$ is small compared to $|R|^2$ and therefore variations would be negligible. The influence of O is further discussed in Sect. 3.5.2.

Virtual Image The term $O|R|^2$ can be considered proportional to O, since $|R|^2$ is constant over the whole hologram. This is also called the virtual image and is the part we are interested in, since it describes the same wave as has been emitted by the object during recording.

Real Image The last term $\bar{O}R^2$ is the conjugate of the object wave and produces the real image which corresponds to an actual focusing of light in space. The conjugate wave $\bar{O}$ is propagating in the opposite direction of O and therefore focuses in an image in front of the hologram. If R is perpendicularly incident to the hologram plane, R^2 is a constant value and $\bar{O}$ will be mirrored at the hologram.

3.5.1 Sampling

Recording the interference pattern affords a *very high* sampling rate. It depends on the wavelength λ and the angle ϑ between the direction of propagation of O and the propagation direction of R. The sampling distance d can be computed by

$$d = \frac{\lambda}{2sin(\vartheta)} \ . \tag{3.27}$$

The chosen sampling distance during recording also has an influence on the maximum viewing angle of the hologram which has to be considered when positioning a camera in Sect. 4.4. To achieve a maximum viewing angle $\vartheta = \frac{\pi}{2}$, the sampling distance d has to be at least as small as $\frac{\lambda}{2}$ in order to guarantee that the Nyquist frequency of the hologram is equal to the carrier frequency of the interference pattern. Note, that by simulating synthetic objects only three orders of magnitude bigger than the wavelength, we are able to keep the hologram size reasonable and still get nice images. This is true, since the resolution of the final image and not the physical size of it is important when rendering to a screen.

3.5.2 Filtering

The reconstruction leads to three different terms (cf. Eq.(3.26)), whereas we are only interested in the virtual image $O|R|^2$. Therefore, the hologram is filtered in a way to suppress the unwanted terms during reconstruction, while maintaining the quality of the virtual image. To suppress unwanted terms we have to rewrite Eq.(3.24) as

$$H = |O|^2 + |R|^2 + O\bar{R} + \bar{O}R \ . \tag{3.28}$$

Considering Eq.(3.4), O and R can be substituted such that $O = A_O e^{\mathbf{i}\varphi_O}$ and $R = A_R e^{\mathbf{i}\varphi_R}$, where A_O is the amplitude of O, φ_O is the phase of O, A_R is the amplitude of R and φ_R is the phase of R leading to

$$H = A_O^2 + A_R^2 + \underbrace{A_O A_R e^{\mathbf{i}\varphi_O} e^{-\mathbf{i}\varphi_R} + A_O A_R e^{\mathbf{i}\varphi_R} e^{-\mathbf{i}\varphi_O}}$$
$$H = A_O^2 + A_R^2 + \qquad 2A_O A_R cos(\varphi_O - \varphi_R)\,. \tag{3.29}$$

For better readability, the x and y dependency of H,A_O,A_R,φ_O and φ_R have been omitted. The first two terms of Eq.(3.29) are close to be constant over the whole hologram, while the third term is statistically varying from $2A_O A_R$ to $-2A_O A_R$. Thus, the average intensity $\hat{H}$ over the whole hologram H can be approximated by $\hat{H} \approx A_O^2 + A_R^2$. The first two terms can, therefore, be suppressed by subtracting the average intensity $\hat{H}$ from the hologram H leading to $H' = H - \hat{H}$. If H' is used for reconstruction, the reference term will be suppressed. This method is known as DC-term suppression in [KJ97], since the almost constant term $A_O^2 + A_R^2$ leads to the DC-term $\mathcal{H}(0,0)$, where $\mathcal{H}$ is the Fourier Transform of the hologram H. The last picture of the first row of Fig. 3.11 shows the reconstruction of the hologram after DC-term filtering. As a comparison, the result without filtering is shown as the second to last picture.

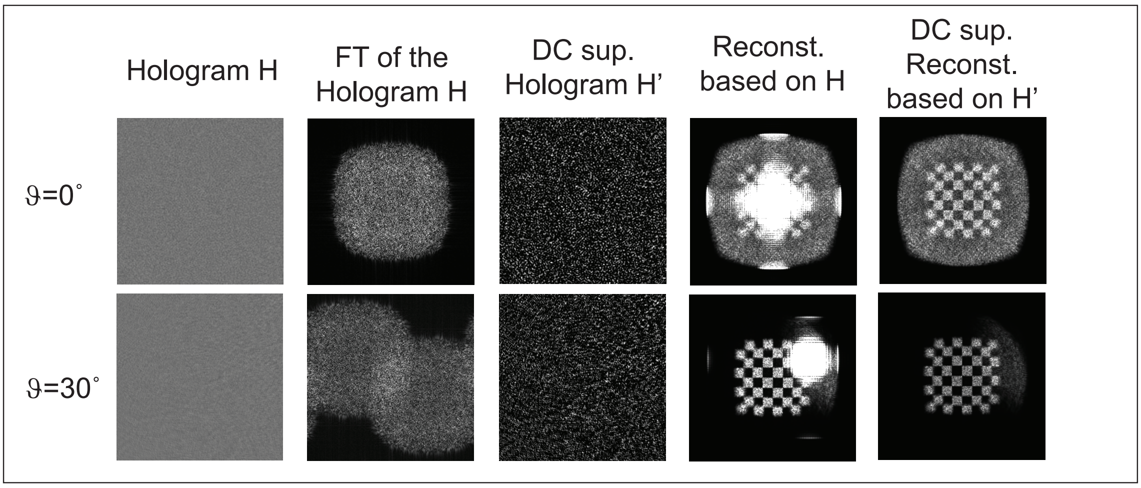

Figure 3.11: *Each row of pictures corresponds to a certain angle ϑ. The DC-Term of the hologram H gets suppressed leading to hologram H' depicted in the third column. In the reconstruction based on H we can clearly see the reference term, while a reconstruction based on H' leads to much better results. Best results are obtained with a big angle ϑ leading however to a high sampling rate of the hologram.*

The DC-term filtering is able to strongly reduce the reference term. However, the real image, represented by $\bar{O}R^2$ still leads to undesired artifacts.

If the angle ϑ between O and R is $0°$ the object wave and its complex conjugate $\bar{O}$ are propagating on the same axis. The propagation direction of $\bar{O}$ can, however, be modified by increasing the angle ϑ, without modifying the propagation direction of O. Therefore, by setting the reference wave sufficiently off-axis, deviates $\bar{O}$ such that it cannot be seen in front of the hologram anymore. The second row of Fig. 3.11 shows the different representations of such an off-axis setting. The resulting image clearly shows an improved image quality compared to the first row. This method of filtering is used for all the hologram renderings in this thesis.

More specific filtering could be applied to the Fourier spectrum $\mathcal{H}$ of the hologram H. Instead of reducing only the DC-term, we could apply a high-pass filtering or an asymmetrical masking of the hologram, as presented by [CMD00], in order to reduce the unwanted terms. However, this method would only improve our results if the spectra of R,O and $\bar{O}$ would not overlap. The spectrum $\mathcal{H}$ is visualized in the second row of Fig. 3.11, where the overlap of the spectra of O, R and $\bar{O}$ is visible.

3.5.3 Conclusion

In this chapter, we have presented the light model, which is used for the holographic pipeline. Based on this model, we have introduced different ways to propagate light waves through free space and evaluate at any point in space. Furthermore, we provided a short description of the principal of hologram recording and reconstruction. Most importantly, we presented a filtering of the hologram, which is highly improving the image quality obtained from the hologram.

In the next chapter, the holographic pipeline is being introduced. Since wave propagation has been discussed in this chapter already, we will focus on the wave-based evaluation of traditional computer graphics representations, such as point-based objects, as well as triangular objects. First rendering results are presented by introduction of an image generation step. The image generation is based on a physical lens model, which allows to focus the wavefront, generates a limited depth of field, while being inherently aliasing free.

CHAPTER 4

Holographic Pipeline

This chapter describes the framework of the holographic graphics pipeline, which is the core of this thesis. The framework is based on the holographic representation and display of graphics objects. As opposed to traditional graphics representations, our approach reconstructs the light wave reflected or emitted by the original object directly from the underlying digital hologram. Our novel pipeline consists of several stages including the digital recording of a full-parallax hologram, the reconstruction and propagation of its wavefront, and rendering of the final image onto conventional, framebuffer-based displays. The required view-dependent depth image is computed from the phase information inherently represented in the complex-valued wavefront. Our model also comprises a correct physical modeling of the camera taking into account optical elements, such as lens and aperture. It thus allows for a variety of effects including depth of field, diffraction, interference and features built-in anti-aliasing. A central feature of our framework is its seamless integration into conventional rendering and display technology which enables us to elegantly combine traditional 3D object or scene representations with holograms. The presented work includes the theoretical foundations and allows for high quality rendering of objects consisting of large numbers of elementary waves while keeping the hologram at a reasonable size.

First, we give an overview of the holographic pipeline in Sect. 4.1. Then, we describe the different steps, such as recording, reconstruction, propagation, and image generation based on a simple example in Sect. 4.2. Since we have already elaborated all the required details for the reconstruction and the propagation step in Chapter 3, we restrict our detailed description on the recording (cf. Sect. 4.3) and the image generation (cf. Sect. 4.4). This chapter concentrates on the theoretical aspects and problems of the wave pipeline and refers to [Goo68] for issues solely relevant to practical holography.

4.1 Overview

In this section a conceptual overview of the complete wave framework including our novel holographic pipeline (highlighted in orange in Fig. 4.1) is described. The first step of this framework involves recording of holo-

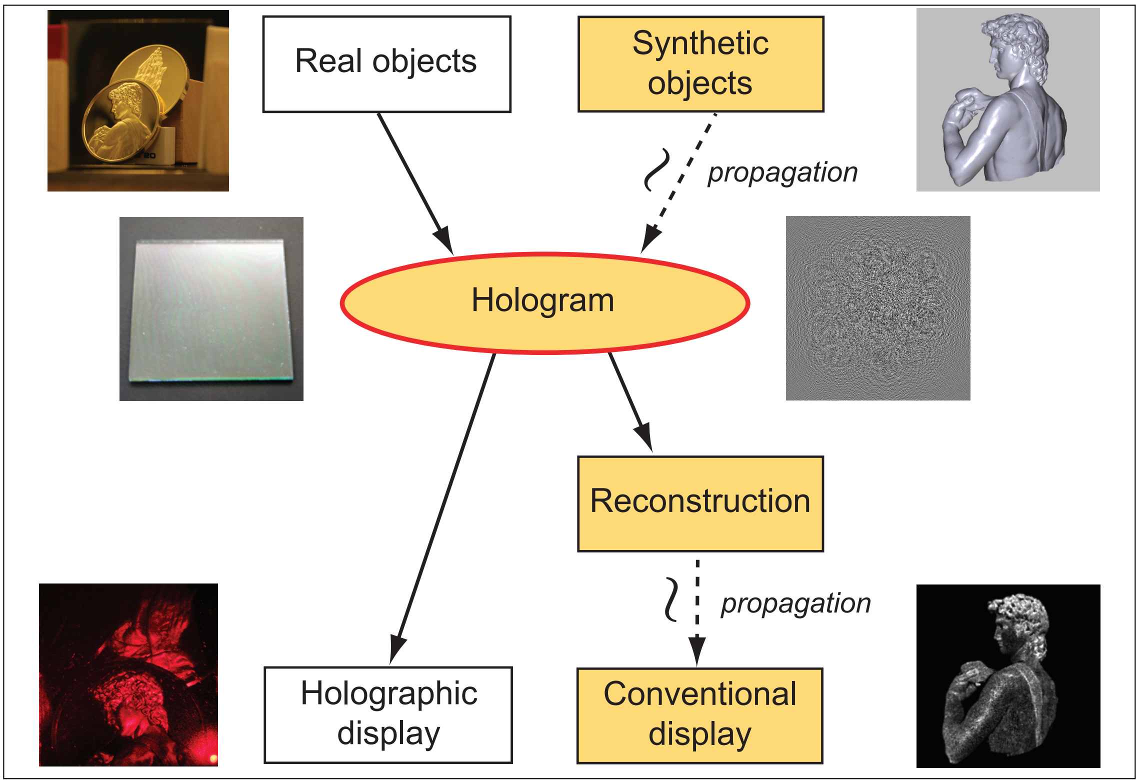

Figure 4.1: *Overview of the wave framework with the components of our new wave pipeline highlighted in orange. (The David dataset is courtesy of Stanford University, Digital Michelangelo Project.)*

grams. We distinguish between recording and acquisition of holograms. Acquisition stands for the physical capturing of a hologram of *real objects* onto a high resolution photographic film or a CCD as in digital holography. Recording (cf.Sect. 4.3) is used for the computation of a hologram from *synthesized objects* as in computer generated holography (CGH). In order to generate a hologram from a wavefront given on the surface of a *synthetic object* we have to be able to propagate this wavefront to any position in space. We explain different techniques of subsampling and propagation of the scene in Sect. 4.3.

Furthermore, the hologram can be used as input to a holographic screen as well as an input for further processing leading to an output on a conventional display. This output requires a discrete reconstruction of the wavefront at the hologram followed by a propagation of this wavefront to an arbitrarily placed camera as described in Sect. 4.4, which provides the input to a conventional display.

4.2 Step by Step Recording and Image Generation

We briefly describe the different steps involved to record an object consisting of N points on a hologram, reconstruct its wavefield, and create an image from the hologram. This high level description shall help to focus on the details in the remaining sections of this chapter.

1. **Recording** The evaluation of the reference wave on the hologram is straight forward. Eq. 4.1 shows the evaluation of a spherical wave for a position $P = (x,y,0)$ on the holographic plane (cf. Fig. 4.2). To evaluate the entire wavefield on the hologram, Eq.(4.1) has to be evaluated for every point P, which is part of the hologram. The object wave is evaluated on the holo-

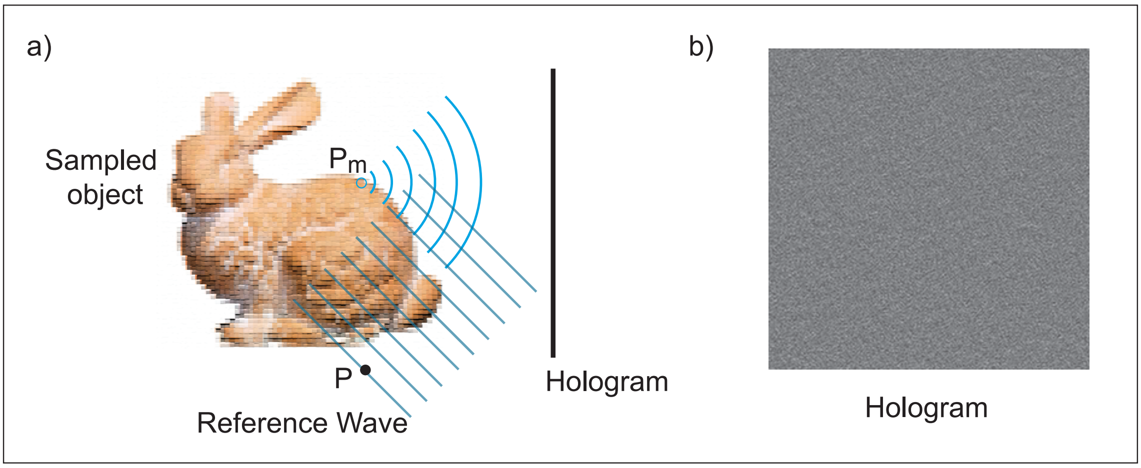

Figure 4.2: *a) The hologram records the interference pattern of the reference wave and the object wave. The object can be subsampled into elementary waves. These elementary waves could be N spherical waves. An example of a spherical wave with origin at P_m is shown in a). A real-valued hologram is depicted in b).*

gram as a sum of elementary waves with the origin at positions P_m and an initial phase of φ_m. All elementary waves are evaluated on every position P

on the hologram. The wavefield of an object sampled into N spherical wave primitives can be described as a superposition as in Eq.(4.2). The absolute value squared of the sum of the object wave and the reference wave is leading to the value of the hologram at every position $(x,y,0)$ (cf Eq.(4.3)). This hologram is real-valued. A more detailed description of the object wave recording is given in Sect. 4.3.

$$R(P) = A_0 e^{\mathbf{ik} \bullet \mathbf{r}_0 + \varphi_0} \quad \text{, with } \mathbf{r} = P - P_0 \tag{4.1}$$

$$O(P) = \sum_{m=1}^{N} A_m \frac{e^{\mathbf{i}k \cdot r_m + \varphi_m}}{r_m} \quad \text{, with } r_m = ||P - P_m||_2 \tag{4.2}$$

$$H(P) = |O(P) + R(P)|^2 \tag{4.3}$$

2. **Reconstruction** We choose a reconstruction wave R' identical to the reference wave R during recording. The evaluation of R' at P is multiplied with every value of the hologram H' at the same position P. H' corresponds to the hologram, which is filtered according to the filtering described in Sect. 3.5.2. The result is a complex-valued wavefield $U(P)$, which is given as a discrete set of complex values (cf. Fig. 4.3). U is given at the hologram on the opposite site of the incident reconstruction wave. It can be described as Eq.(4.4).

$$U(P) = R(P) \cdot H'(P) \approx O(P) \tag{4.4}$$

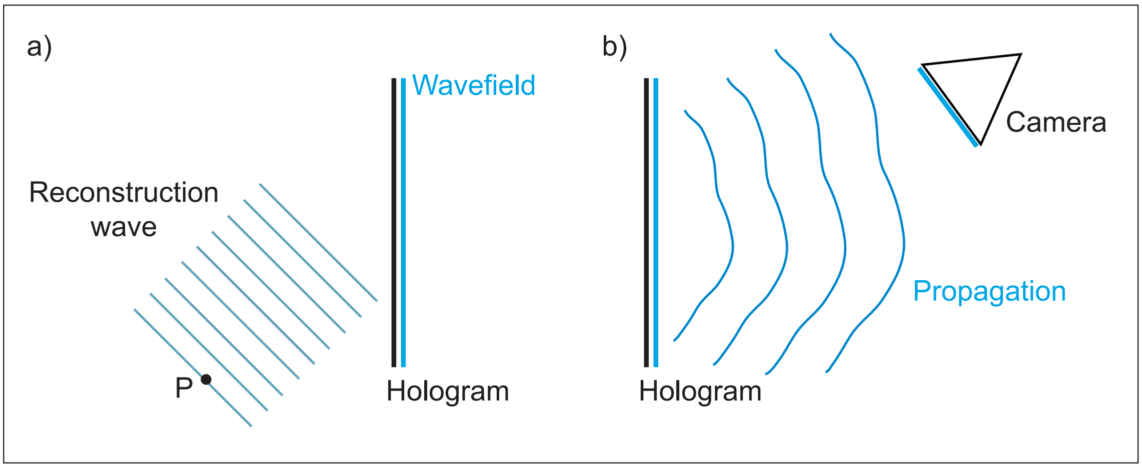

Figure 4.3: *a) The multiplication by the reconstruction wave R' generates a discrete complex-valued wavefield just behind the hologram. b) The reconstructed wavefield is propagated through a homogeneous occlusion free space to the observing camera.*

3. **Propagation** Since $U(P)$ is not given in an explicit form anymore, we are using the propagation of a wavefield given on a plane as described in Sect. 3.3.6. This method allows to propagate the wavefield $U(P)$ to the observing camera, leading to wavefield $U(P')$. The set of points P' lies on the camera aperture.

4. **Image Generation** After the propagation, the wavefield is given as discrete complex samples of the wavefront, at the camera aperture. In our pipeline, we use a thin lens model in order to focus the wavefront by means of a phase modulation. The phase modulation function, also referred to as the lens function $L(P')$, is given per point P' of the aperture. The application of the lens function focuses the wavefront according to the focal length chosen for the camera. We refer to the focused wavefield as $\check{U}(P')$ (cf. Eq.(4.5)). More details about the lens is provided in Sect. 4.4.1.

$$\check{U}(P') = U(P') \cdot L(P') \tag{4.5}$$

The focal length of the lens function is chosen such that the image plane lies at infinity. This has the advantage, that the amplitude of a plane wave, can be interpreted as the intensity of a ray parallel to the propagation direction of the wave going through the center of the lens. Since the Fourier Transform of the wavefield leads to the decomposition into plane waves propagating in all directions, we already obtain the intensities of the rays required for the image generation. Finally, we can distort the sampling relative to the angular cosines into a regular grid of pixels. More specific descriptions about the required lens function, handling of boundary problems of the aperture, the distortion parameters, and image quality improvements are described in Sect. 4.4.

4.3 Recording of Specific Objects

Every object can be represented by a number of elementary waves, also called *primitives*. Besides the plane wave, our pipeline supports two different primitives, namely a spherical wave and a wavefield given on an aperture. These two primitives allow to record the wavefront of the two most common object representations in computer graphics, which are point-based objects and triangular objects. We show the characteristics, advantages, and disadvantages of representing an object either by a number of spherical waves, or apertures in the following two sections.

4.3.1 Point-based Objects

This representation approximates a surface of an object by an arbitrary number of points. The points are represented by point sources implemented as spherical waves, with the center at the position of the point. The waves can efficiently be evaluated using shaders on graphics hardware. Traditional point-based objects can, therefore, be represented by a number of point sources. The straight forward evaluation of every point source of an object onto the holographic plane, would make the object appear transparent, since occlusion is not being taken into account this way. Furthermore, the sampling density is a critical issue, when evaluating point-sampled objects, since the intensity of the finally rendered image of the hologram depends on it.

Occlusion

A point in space does not have a physical extent and can, therefore, not occlude propagating waves. We use a point-based representation, where every point is represented by a small oriented disk. This representation is called surfel representation and has been presented by Pfister et al. [PZvBG00].

In a first step, we reject all the backfacing surfels by considering the normal of the element. An example of this is shown in Fig. 4.4a. In a second step, we render the surfel object based on ray-based optics using a technique called splatting [ZPvBG01]. The viewpoint is chosen to lie on the hologram, while the viewing direction is perpendicular to the hologram facing the scene. A splatted surfel covers a certain number of pixels. If the distance of the surfel lies further away than all the surfels being splatted on any of those pixels, it is being considered as occluded. To avoid, that neighboring surfels are occluding each other, we add a bias to the depth of the closest surfel before checking for occlusion. Fig. 4.4b shows an example of this occlusion handling. Of course, the visibility of a surfel often depends on the parallax. Since the hologram stores the full parallax of the scene, the occlusion would have to be evaluated per pixel of the hologram. This is, however, too time consuming to compute. By evaluating the visibility only for every n^{th} pixel of the hologram and blend the values in between, allows to reduce the computational load without too much loss in quality. We present a method for wave-based occlusion handling in Chapter 5, which allows to handle the visibility for all possible viewing direction at once. Currently, this method is only presented for planar occluders and has not been implemented for surfel objects.

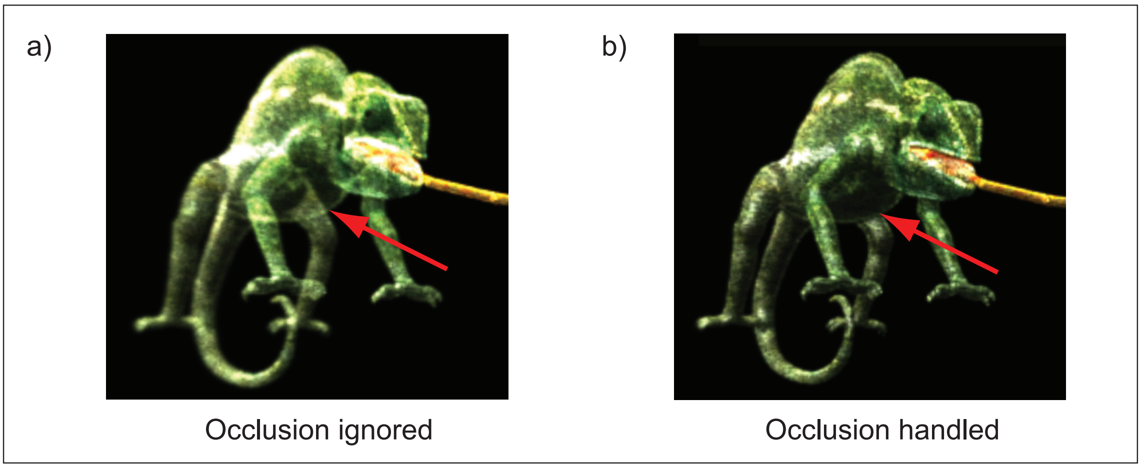

Figure 4.4: *a) This model was rendered by only rejecting backfacing surfels. We can clearly see the transparent appearance at parts of the object. b) Using a splatting step to determine the visibility allows rendering of opaque objects.*

Sampling Density

An arbitrary sampling of the point-sampled object can lead to different visible artifacts, which can vary depending on the viewing direction. If the view dependent sampling is not equally distributed over the entire object, we can observe intensity variations despite of homogeneous ambient color regions (cf. Fig. 4.5a). Higher sampling densities will appear brighter, since

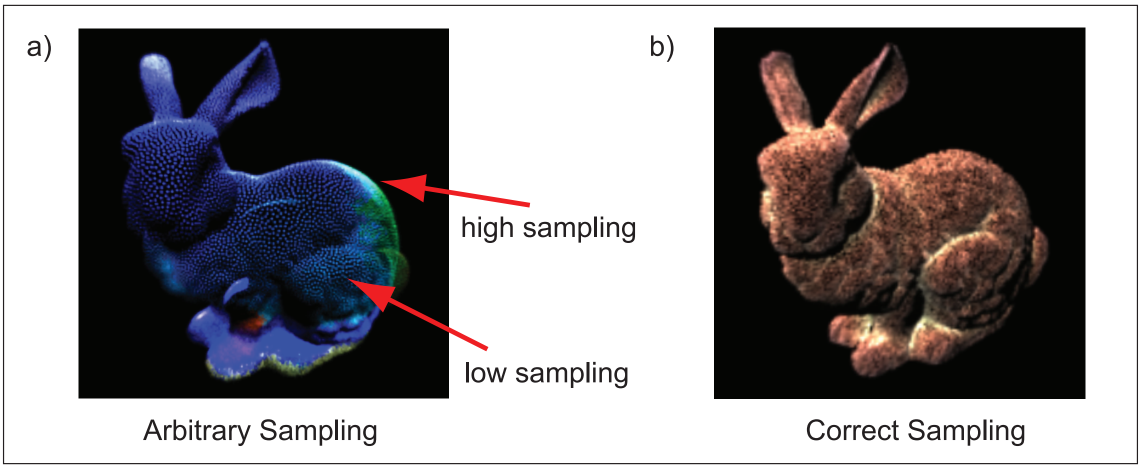

Figure 4.5: *a) This object features an arbitrary sampling of the point sources leading to too bright areas and holes. b) Correcting the sampling using a view dependent sampling approach allows to minimize these artifacts. The black dots in this rendering are not due to a low sampling, but due to speckle noise.*

more energy is being integrated over one pixel. In our pipeline, we scale the intensities of the point sources depending on the number of visible surfels contributing to one pixel. We evaluate one surfel per point source of the point-sampled object. Since we simulate each surfel as a spherical wave, we loose the geometric extent of the surfel, which can lead to holes in the final rendering. An undersampled area looks like a dotted surface (cf. Fig. 4.5a). This artifact can be minimized by creating additional spherical waves on the entire surfel. Fig. 4.5b shows an object, where the sampling density has been adapted according to the described method. The intensity variation is due to the applied directional lighting. The small black dots which can be observed in Fig. 4.5b do not stem from too low sampling, but from speckle noise. Speckle noise is described in Sect. 4.4.5.

A change of the viewing direction leads to different view dependent sampling requirements. An accurate recording would require a reevaluation of the sampling density per hologram pixel. However, since the sampling of the hologram is very dense, the required sampling density of the scene remains almost the same for neighboring hologram pixels. Therefore, we evaluate the sampling density only at every n^{th} pixel of the hologram, and blend the values in between. This is minimizing the artifacts while reducing the computational load. Although we minimize the artifacts, we cannot guarantee a hole free rendering for any zoomed in position of the camera unless we considerably increase the number of samples.

4.3.2 Triangular Objects

Triangular meshes are one of the most common representations in computer graphics. In this thesis, we present two different ways of evaluating the wavefront of such an object. One way of evaluation is similar to the point-based approach in the sense, that the triangles are subsampled into point sources, which are then being evaluated similarly to the point-based objects. The second way of evaluation is based on the propagation of discrete wavefields, which are given on a planar surface. The following two subsections elaborate the two kinds of evaluations in more detail.

Point Sampling

An equidistant sampling of every triangle into point sources creates good conditions for the regular view dependent sampling, which is very important for reasons elaborated in Sect. 4.3. To avoid the entire splatting step, which is required for an optimal sampling and weighting of the point sources, we can simply scale the intensity values with the cosine of the angle between the viewing direction and the normal of the triangle. This approximation works well if the depth extent of the object is not too big. A big depth extent would have a direct influence on the view dependent sampling and therefore on the evaluated wavefront.

Occlusion can be handled in a ray-based approach by evaluating the visibility of every triangle. However, it would have to be evaluated for every pixel of the hologram, which can be very time consuming. In Chapter 5, we present a way to handle the occlusion in a wave-based manner, which handles the occlusion for all pixels of the hologram at once.

Triangular meshes are usually provided with a texture or a color component. In order to avoid aliasing, we use graphics hardware to prefilter and rescale the texture according to the chosen sampling of the point sources.If computational performance is not an issue, the point sampling could be adapted to the highest texture frequency. This would avoid the filtering of the texture.

One of the problems, when using point sources to evaluate a triangular mesh, is the complexity of evaluation. Every point source has to be evaluated onto every pixel of the hologram. A more efficient evaluation is presented in Sect. 4.3.2.

Triangle Propagation

In this approach, we do not subsample the textured triangle into point samples, but directly propagate every triangle onto the hologram. In Sect. 3.3.6, we showed a way to efficiently propagate a wavefield given on one plane, to the next plane. We can use the same approach to propagate the triangles to the hologram. The propagation described in Sect. 3.3.6 relies on the

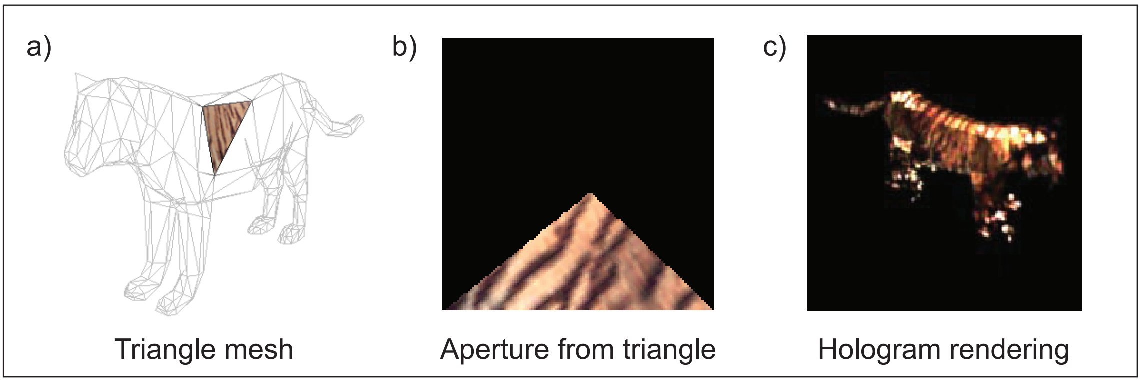

Figure 4.6: *a) A triangle of the object is picked and b) mapped to an aperture. c) A hologram rendering of an object, which is based on triangular primitives.*

Fourier Transform. In order to apply a fourier transform on a triangle, we have to map the triangle onto a rectangular aperture before propagation (cf. Fig. 4.6b). The texture filtering and mapping can be handled in one step by rendering the triangle using an orthographic camera into a viewport of the desired size. Of course, this assumes, that the texture filtering during rendering is handled correctly.

Occlusion has to be handled in a separate step. We evaluate a view dependent visibility mask, which is applied to the mapped triangle. Fig. 4.7a shows two triangles, where the occlusion has not been handled. We clearly see the transparent appearance of the blue triangle. In Fig. 4.7b the red tri-

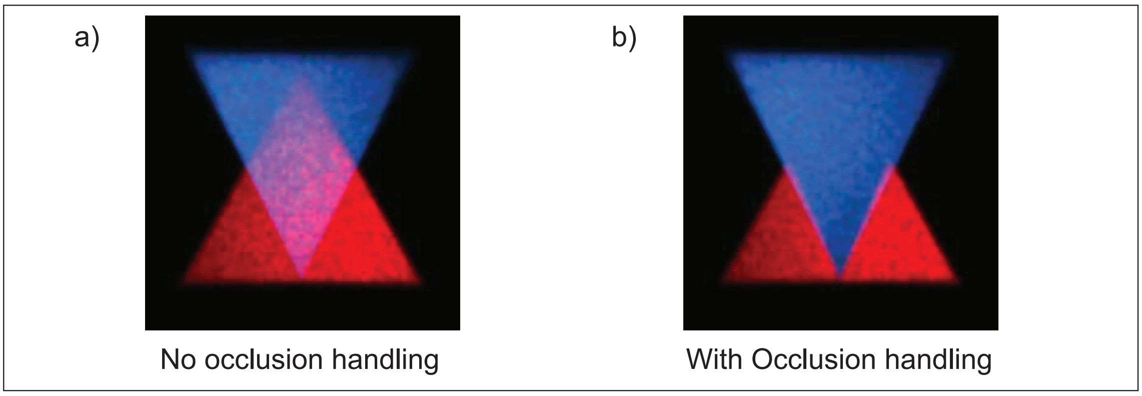

Figure 4.7: *a) Two occluding triangles without occlusion handling. b) The red triangle is multiplied by a visibility mask, before being evaluated on the hologram. This ray-based occlusion handling allows to simulate opaque objects, but requires a lot of computation.*

angle was first multiplied by its visibility mask. Therefore, the red part, which is occluded by the blue triangle is not emitting any waves anymore, such that the blue triangle appears opaque. An object rendered with such an approach is shown in Fig. 4.6c. It is important to note, that the visibility mask is view dependent and has to be evaluated for every hologram pixel in the worst case. Therefore, we propose a wave-based occlusion method in Chapter 5, which is handling the occlusion for all the direction at once.

4.4 Image Generation

Based on the propagation described in Sect. 5.2.2, any wave field given on a synthetic object or on a reconstructed hologram can be evaluated on an aperture of an arbitrarily placed camera in space. The image generation consists of four steps, namely the simulation of an optical camera model (Sect. 4.4.1), a multi-wavelength rendering for colored holograms (Sect. 4.4.2), a depth evaluation for compositing of multiple objects (Sect. 4.4.3) and a speckle noise reduction (Sect. 4.4.5).

4.4.1 Camera Model

In a first step the simplest architecture of a camera, the pinhole camera, is studied for image generation. Artifacts occurring because of limited aperture size are described. Furthermore a more complex camera model including a lens is introduced.

Pinhole Camera As shown in Fig. 4.8a) a wave is propagating through a pinhole and producing an image on a parallel plane at distance z. Modeling

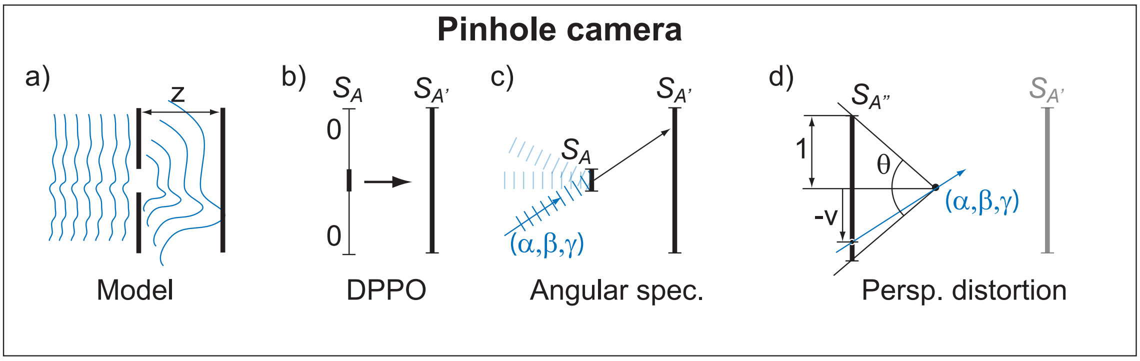

Figure 4.8: *Image generation simulating a pinhole camera as in a) can be achieved by DPPO b) or by making use of the directional cosines c). A perspective distortion leads to the final image d).*

the slit by an aperture S_A and applying DPPO to propagate the wave to the image plane $S_{A'}$ would be possible, but would require zero padding of S_A (cf. Fig. 4.8b). Assuming an indefinitely small slit and knowing the directional components (α,β,γ) of the wave front (cf.Sect. 3.3.5) is sufficient to generate the image at a distance z assuming geometrical propagation from the slit (cf. Fig. 4.8c). The directional components can be calculated from the frequency components ν_x and ν_y of the angular spectrum as $(\alpha,\beta,\gamma) = (\lambda\nu_x,\lambda\nu_y,\sqrt{1-(\lambda\nu_x)^2-(\lambda\nu_y)^2})$. Finally a mapping between the directional cosines and the image pixels given by coordinates u and v as in Fig. 4.8d can be found by a projection matrix P as $(\alpha,\beta,\gamma,1)P \Rightarrow (u,v)$ leading to

$$\left.\begin{aligned} \alpha &= -\tfrac{u}{s_x}\sqrt{(\tfrac{u}{s_x})^2+(\tfrac{v}{s_y})^2+1}^{\,-1} \\ \beta &= -\tfrac{v}{s_y}\sqrt{(\tfrac{u}{s_x})^2+(\tfrac{v}{s_y})^2+1}^{\,-1} \end{aligned}\right\} s_y = \frac{1}{tan(\frac{\theta}{2})}, s_x = r_a s_y \tag{4.6}$$

θ is the field of view in y direction with the aspect ratio $r_a = \frac{w}{h}$ where w is the image width and h the image height. The pixels of the image are computed by a weighted sum (e.g. bilinear interpolation) of the directional cosines.

Resolution The maximal resolution of the image is determined by the size of the pinhole W. This can be shown by the relation of the spacing δ_f in the frequency and the spacing δ in the spacial domain given as $\delta\delta_f = \frac{1}{N}$ where N is the number of samples. The smallest non-zero frequency is given as $\delta_f = \frac{1}{W}$, since $W = N\delta$. This corresponds to the direction cosine $\alpha_{min} = \frac{\lambda}{W}$ and gets close to the Rayleigh resolution criterion $\alpha = 1.22\frac{\lambda}{D}$ describing the critical angle at which two point sources can still be distinguished when diffracting at a circular aperture of diameter D [Tip91].

Apodization The effect of a rectangular aperture as used in our pipeline can be simulated as a multiplication by a rect-function in the spatial domain and as a convolution with a sinc-function in the frequency domain (cf. Fig. 4.9a). The convolution with the sinc-function can lead to visible artifacts called ringing, appearing as blur in the axial directions. This blur can be minimized by applying a tapering function as proposed in [Gbu05] and [GC98]. In our implementation we use a cosine window function being 1 in the middle of the aperture and decreasing to 0 at the border of the aperture, resulting in an apodization as shown in Fig. 4.9b).

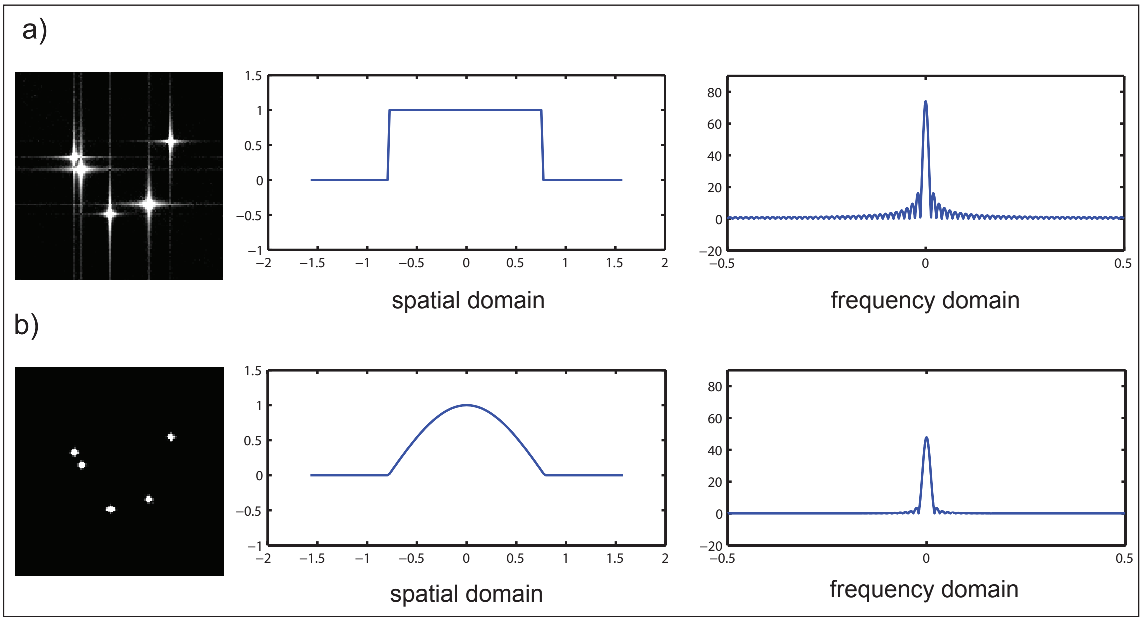

Figure 4.9: *a) Applying a rect-function for the aperture leads to ringing artifacts. The Fourier Transform of the aperture is shown on the right. b) Multiplying an apodization function to the aperture in the spatial domain results in an image free of ringing.*

Thin Lens Placing the camera at arbitrary positions while choosing the object appearing in focus requires a lens. We use a thin lens model:

$$\frac{1}{d_O} + \frac{1}{d_i} = \frac{1}{f} \tag{4.7}$$

with d_O being the distance from the lens to the object, d_i being the distance from the lens to the image and f being the focal length.

The effect of a thin lens can be simulated by multiplying a complex-valued function $L(x,y)$

$$L(x,y) = e^{-\mathrm{i}kr}, \text{ with } r = \sqrt{x^2 + y^2 + f^2} \tag{4.8}$$

inducing a phase shift to the wavefront [Goo68]. The phase shift transforms the spherical wave with origin at focal distance f into an almost planar wave:

$$(A_0 \frac{e^{\mathrm{i}kr}}{r}) L(x,y) = \frac{A_0}{r}. \tag{4.9}$$

Even though it does not lead to a perfect planar wave it results in a good approximation for a point source being far away compared to the size of the lens. Figure 4.10 shows a scene consisting of five points being placed at different distances from the lens taken with three different focal lengths. Off-axis points at the edge of the image can produce a comet-shaped image,

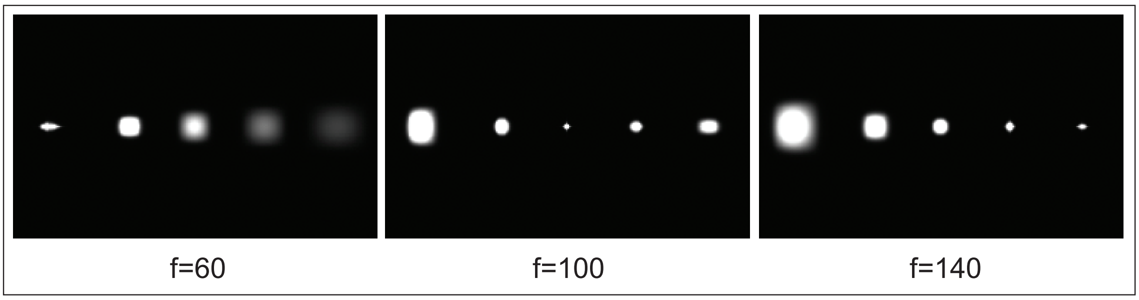

Figure 4.10: *Scene with five points placed at different depth and rendered with three different focal lengths f.*

an effect known as *coma aberration*. For more general applications one would need to implement a lens simulation, which distributes aberration and artifacts over the whole image.

Due to a physically based model of the holography pipeline effects like *depth of field* or *refocusing* are automatically generated (cf.Sect. 4.5).

4.4.2 Color holograms

A non-monochromatic wave could be split into all spectral frequencies, which could be propagated separately in the wave framework. However, instead of simulating all frequencies separately we chose three primary colors (λ_r,λ_g,λ_b, having approximately the wavelength of red, green and blue) from which other visible colors can be combined linearly. The amplitudes A_0 at the origin of each wave are set as $A_0 = \sqrt{I}$, where I is the intensity of one of the primary colors. The final image can thus be rendered by calculating the scene for each monochromatic primary color once.

Simulating non-monochromatic colors introduces another artifact called *chromatic aberration* when using lenses in the optical path. This means that when choosing a focal length f_r for a lens, the effective focal length for the image generated with wavelength λ_g is $f_r(\frac{\lambda_g}{\lambda_r})$. Applying a Fourier based image generation and perspective distortion does not introduce further chromatic artifacts.

When choosing the wavelength for the primary colors we are not restricted to take realistic values since the intensity I of the color channel only influences the amplitude. The wavelength λ_r,λ_g and λ_b could be exactly the same avoiding all aberrations and artifacts related to wavelength differences. In physical lens systems, multiple lenses and lens materials are minimizing the effect of chromatic aberration.

4.4.3 Evaluate Depth

So far we have shown the generation of an image of a single object. However, if we have scenes composed of several objects occlusions have to be taken care of. We present a depth reconstruction from a hologram in order to compose objects using the depth buffer.

Depth From Phase Calculating depth from phase is a straightforward approach also used in digital holographic interferometry as shown in [SJ05]. Knowing the phase of the source and the image, the relative depth can be calculated out of this phase difference modulo λ. In order to unambiguously reconstruct the surface the depth difference of two neighboring pixels has to be smaller than $\frac{\lambda}{2}$. However, a relative surface is not sufficient for depth composition.

Depth From Phase Difference A more advanced approach consists in generating several images with slightly different wavelengths λ_i. In our case we generate two images with two wavelengths λ_1 and λ_2 being related by $\lambda_1 = \epsilon\lambda_2$ with $\epsilon < 1$. The phase φ_i can be calculated as $\varphi_i = \left(\frac{2\pi}{\lambda_i} * r + \varphi_0\right) \bmod \pm \pi$ where $\varphi_i \bmod \pm \pi = ((\varphi_i + \pi) \bmod 2\pi) - \pi$.

This leads to a phase difference $\Delta\varphi$ in $[0,2\pi)$ as $\Delta\varphi = 2\pi r(\frac{1}{\lambda_1} - \frac{1}{\lambda_2})$. Considering the maximal phase difference 2π we get the maximal radius r_{max} as

$$r_{max} = \frac{2\pi}{2\pi\left(\frac{1}{\epsilon\lambda_1} - \frac{1}{\lambda_1}\right)}, \text{ with} \tag{4.10}$$

$$\epsilon = \frac{1}{1 + \frac{\lambda_1}{r_{max}}} \quad . \tag{4.11}$$

Therefore, the bigger r_{max} is, the closer ϵ has to be to 1 leading to a small difference between λ_1 and λ_2. Applying this depth reconstruction results in an absolute depth needed for depth composition. Before having a closer look at the influence of interference problems in Sect. 4.4.3, we will analyze the impact of a lens on depth reconstruction.

Depth Reconstruction Using A Lens Applying a lens for image generation imposes a consideration of influence on depth reconstruction as well. Considering the mathematical representation of a lens with focal length f as described in Eq.(4.8) we can find images $U_{\lambda_1}(P)$, $U_{\lambda_2}(P)$ at distance R from the origin as follows:

$$U_{\lambda_j} = A_0 \frac{e^{\mathbf{i}k_j R}}{R} e^{-\mathbf{i}k_j f} \quad \text{, where } k_j = \frac{2\pi}{\lambda_j}, j = 1,2\,. \tag{4.12}$$

The phase difference $\Delta\varphi$ evaluates to

$$\Delta\varphi = 2\pi(\frac{1}{\lambda_2} - \frac{1}{\lambda_1})(r_f - f) \tag{4.13}$$

for the phases given as $\varphi_1 = k_1(r_f - f)$ and $\varphi_2 = k_2(r_f - f)$. The distance of the source can thus be found by solving Eq.(4.13) for r_f. Comparing the reconstruction based on a camera with a lens leading to depth r_f with the reconstruction without a lens leading to depth r we find the simple relation to be $r_f = r + f$. In order to find the desired distance values d_{max} around the focal point $[f - d_{max}, f + d_{max}]$ we have to evaluate a new factor ϵ_l such that $\lambda_2 = \epsilon_l \lambda_1$ leading to:

$$\epsilon_l = \frac{1}{1 + \frac{\lambda_1}{2d_{max}}} \quad . \tag{4.14}$$

Influence Of Interference Despite of having better results using a depth reconstruction approach based on phase difference, the problem of interference still remains. We explain the case of two interfering point sources PS_1

and PS_2 with origins P_1 and P_2 at distance r_1 and r_2 from the point of evaluation P' (cf. Fig. 4.11a) and finally infer the general case of possible multiple point sources. Since r_1 and r_2 cannot be reconstructed separately we would like to find a depth r' which lies in between the two distance values such that $r_1 \leq r' \leq r_2$. This is equal to guaranteeing $\Delta\varphi_1 \leq \Delta\varphi' \leq \Delta\varphi_2$ for $\Delta\varphi_1$ and $\Delta\varphi_2$ being the individual phase differences of PS_1 and PS_2 respectively and $\Delta\varphi'$ being the phase difference of the superposition of PS_1 and PS_2 at P'.

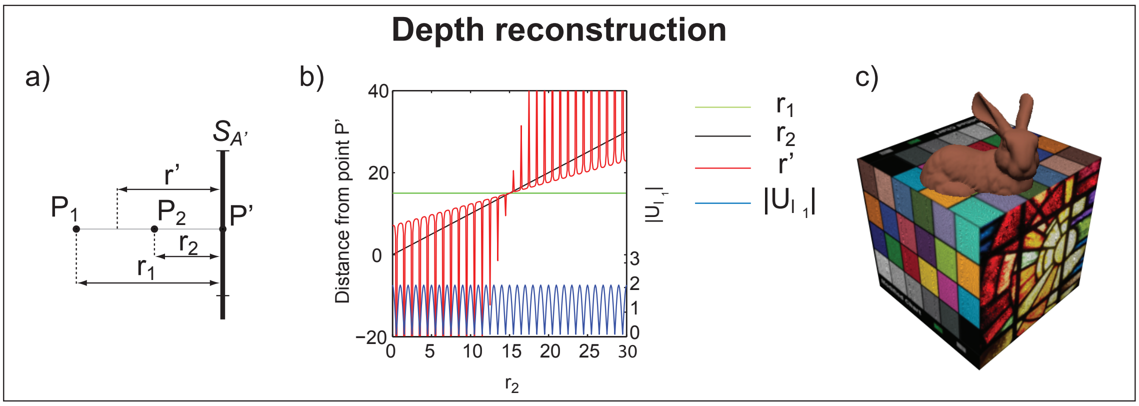

Figure 4.11: *Moving P_2 in a) from P' to $r_2 = 30$ is leading to a reconstruction of r' as in b). Improving results in b) by multiple iterations for depth refinement leads to results as in c).*

The complex valued image U_{λ_j} describes the interference pattern of PS_1 with amplitude A_{1_0} and PS_2 with amplitude A_{2_0} with wavelength λ_j as follows:

$$U_{\lambda_j} = A_{1_0}\frac{e^{\mathbf{i}k_j r_1}}{r_1} + A_{2_0}\frac{e^{\mathbf{i}k_j r_2}}{r_2} \quad \text{, with } k_j = \frac{2\pi}{\lambda_j} \quad . \tag{4.15}$$

Ignoring interference would lead to a depth reconstruction as seen in Fig. 4.11b). The desirable values of r' lying in between r_1 and r_2 always coincide with the maxima of $|U_{\lambda_1}|$. This can be shown since $|U_{\lambda_1}| = max(|U_{\lambda_1}|)$ implies $min(\Delta\varphi_1, \Delta\varphi_2) \leq \Delta\hat{\varphi} \leq max(\Delta\varphi_1, \Delta\varphi_2)$. The associative property of vector addition allows the extension of two point sources to multiple point sources, so the proof holds for the addition of a new point source with the sum of the remaining point sources as well. In order to improve depth quality in case of $|U_{\lambda_1}| \neq max(|U_{\lambda_1}|)$ the image is rendered several times while randomizing the phase of the point sources for every iteration. Depending on the number of iterations the depth reconstruction can be arbitrarily close to an optimal r'.

Finally, the depth buffer is multiplied by a mask created by a thresholding of an introduced opacity channel, verifying that a pixel belongs to a certain

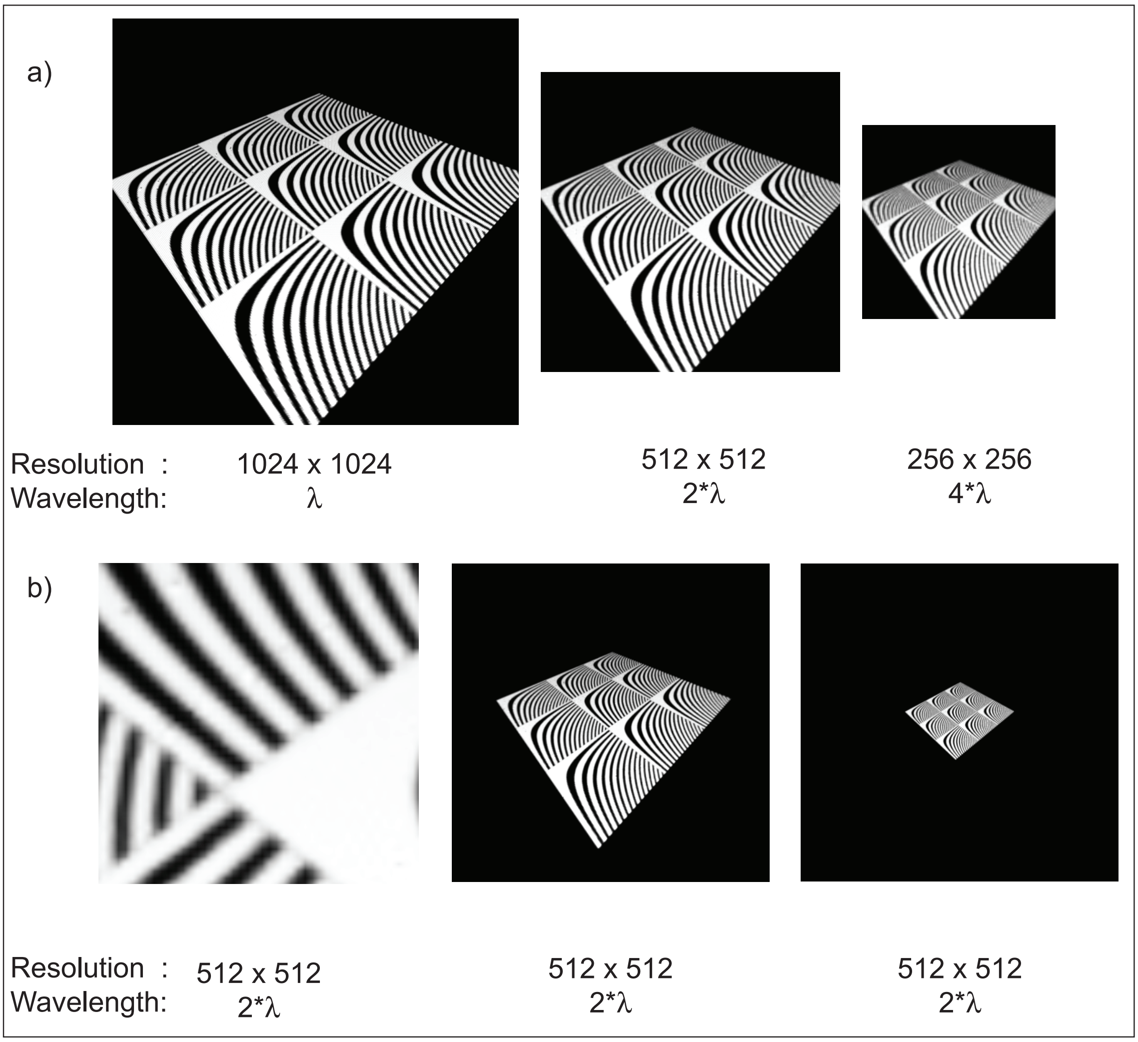

Figure 4.12: *a) The three images show the same scene evaluated at different image resolutions and therefore different wavelengths. b) shows different magnifications of the same textured plane without having to apply any filtering. (Please look at the pdf for correct anti aliasing appearance.)*

object. Note, that once the wave based object has been evaluated the compositing with rasterized objects as depicted in Fig. 4.11c) can be performed in real time.

4.4.4 Aliasing free images

A rasterized image is essentially a discrete function defined on a 2D regular grid. To generate an aliasing free image the sampling process has to satisfy the Nyquist criterion. Since the Nyquist criterion in traditional computer graphics depends on the frequency of the projected textures, appropriate filtering has to be applied to obtain anti aliased images as has been shown by Heckbert [Hec89][GH86]. However, in our pipeline generated images are inherently free of aliasing.

The scenes used in our pipeline can consist of point primitives or textured planar primitives emitting waves of equal wavelength λ. The color of the point sources as well as the planar primitives is encoded in the amplitude of the propagating waves. The interference generated by the superposition of all these primary waves create a wavefront containing the complete information of the whole scene. Since the wavefront does not have a higher frequency than the interference pattern, the sampling rate will not have to be higher either. In order to avoid any aliasing the sampling rate of the interference pattern has to guarantee the Nyquist frequency and therefore a sampling distance of $\frac{\lambda}{2}$, which has been shown to be the case in Sect. 3.5.1. By simulating monochromatic waves we implicitly guarantee a distinct lower bound for the wavelength λ and therefore a distinct upper bound for the required sampling rate $\frac{2}{\lambda}$.

Finally, by Fourier transforming the sampled interference pattern we get the aliasing free directional components of the wavefront. After an according mapping described in Sect. 4.4.1 we get the aliasing free image.

Changing image resolution implies a modified frequency in the sampling process of the image function. In traditional computer graphics this requires a pre- or post-filtering of the texture in order to comply with the Nyquist frequency. In our pipeline, however, a lower image resolution does not require any filtering, but simply an increased wavelength of the monochromatic source. In order to avoid a scaling of the object wave the hologram is regenerated using the new wavelength. Fig. 4.12a shows three renderings with different resolutions and therefore three different wavelengths. We did not apply any filtering on the original texture and still obtain aliasing free images. Furthermore, magnification or minification does not require any filtering or wavelength adjustments as can be seen in Fig. 4.12b.

4.4.5 Image improvement

Throughout our pipeline, coherence is a fundamental property to generate a standing wave which again is necessary to record and reconstruct a hologram. However, generating images based on coherent light can cause visible artifacts due to the static interference pattern called speckle noise, which can even be visible by the naked eye in physical setups as shown in [DEF+75].

Speckle Noise Reduction Through Higher Resolution The speckle size δ_S depends on the distance x of two interfering light waves with wavelength λ

being at distance z of the interference measurement as follows (see [SJ05]):

$$\delta_S = \frac{\lambda z}{x} \tag{4.16}$$

Therefore the minimal speckle size $\delta_{S_{min}}$ can be calculated dependent on the aperture size W as $\delta_{S_{min}} = \frac{\lambda z}{W}$.

These properties imply a higher frequency of speckle noise with an increasing image resolution meaning a bigger W (cf. Sect. 4.4.1), whereas the frequency of the actual image content stays the same. Hence, filtering out high frequency speckle noise does not affect the actual image content of the rendered image. Additional reduction can be achieved by down sampling the image using a low pass filter.

Speckle Noise Reduction By Temporal Averaging Instead of filtering in space as mentioned above, it might be faster to filter in time by rendering multiple images from a diffuse object with randomized phases. Computing the final image is achieved by either blending all the images or choosing the maximal value per pixel. The signal to noise ratio (SNR) is better for the maximal composition of N images if only two sources with amplitude A_1 and A_2 are contributing to an image point P of image k with intensity $I_k(P)$, since the expectation $E[\lim_{N\to\infty} max_k(I_k(P))] = (A_1 + A_2)^2$ is bigger than the expectation for averaging $E[\frac{1}{N}\sum_{k=1}^{N} I_k(P)] = A_1^2 + A_2^2$. However, the more interfering sources are involved, the smaller the probability that the maximal value is found after N iterations. The necessary and sufficient condition for a maximal composition of interfering sources consists in all sources having equal phases as shown in Sect. 4.4.3. The SNR is improving very quickly up to roughly 20 iterations. As stated in Tab. 4.1 the SNR for most objects lead to good results around 70 iterations. In our pipeline we choose an identical phase distribution of the entire scene for all three color channels. By doing so, we get an identical speckle noise for all three channels, if all of them are recorded using the same λ. Producing the same speckle pattern for all three color channels removes possible color shifts in the final hologram rendering.

4.5 Results

Using our framework we are capable of rendering high quality images, while taking into account depth of field and refocusing of the scene by simply adjusting the aperture size or focal length as if a real camera would have been used (cf. Fig. 4.13). Even big objects with primitives count up to 300K (cf. Fig. 4.14f) can be evaluated in a reasonable amount of time (cf. Tab. 4.1), due to usage of fragment shaders and angular spectrum propagation. By using propagation in frequency space, time can be reduced considerably. A 1024^2 hologram as used in Fig. 4.14e) can be propagated in 50s instead of 5min 30s. Creating a hologram of 1024^2 with double precision per primary color channel leads to a total size of 24MB. Undersampling of a high-frequency texture during the rasterization step in ray-based image generation can lead to severe aliasing artifacts whereas in a wave-based framework anti-alising comes for free Fig. 4.15c). All the images were generated on a Pentium 4 3.2GHz containing a NVidia GeForce 7800GT. Currently we make use of fragment shaders for point evaluations.

Table 4.1: *Statistics*

Figure	# Of Prim.	Holo. Size	Pict. Size	Time per It.	# Of It.
Fig. 4.13	100K	1024^2	1024^2	185s	70
Fig. 4.14a)-d)	50K	512^2	512^2	43s	24
Fig. 4.14e)	50K	1024^2	1024^2	155s	74
Fig. 4.14f)	300K	1024^2	1024^2	325s	70
Fig. 4.15a)/d)	137K	1024^2	1024^2	197s	90
Fig. 4.15b)	173K	1024^2	1024^2	241s	90

4.6 Holography vs. Light fields

Introducing holograms as a data representation also suggests a comparison to light fields. Independent of the scene complexity, light fields can be acquired quite easily and can be rendered in an efficient way, leading to high quality images. Simulation of depth of field and refocusing is possible, although inadequately sampled camera planes can lead to aliasing. The lack of geometry makes compositing and occlusion difficult and may lead to defocus artifacts. Furthermore, seamless full parallax light fields require a considerable amount of data and proper reconstruction filters for an anti-aliased image generation.

Holograms on the other hand are inherently free of aliasing. Providing information of phase and amplitude allows to reconstruct depth, which enables scene compositing, occlusion, future reshading or morphing. Further effects, such as depth of field, refocusing or diffraction can be intuitively adjusted by changes of the aperture, focal length, wave length or shape of the aperture, without requiring knowledge of the scene. Although holograms of reasonable size can lead to high quality images, capturing of real objects requires a big amount of data. Additionally, rendering time is significantly more extensive than rasterized rendering, while the image quality can still suffers from speckle noise.

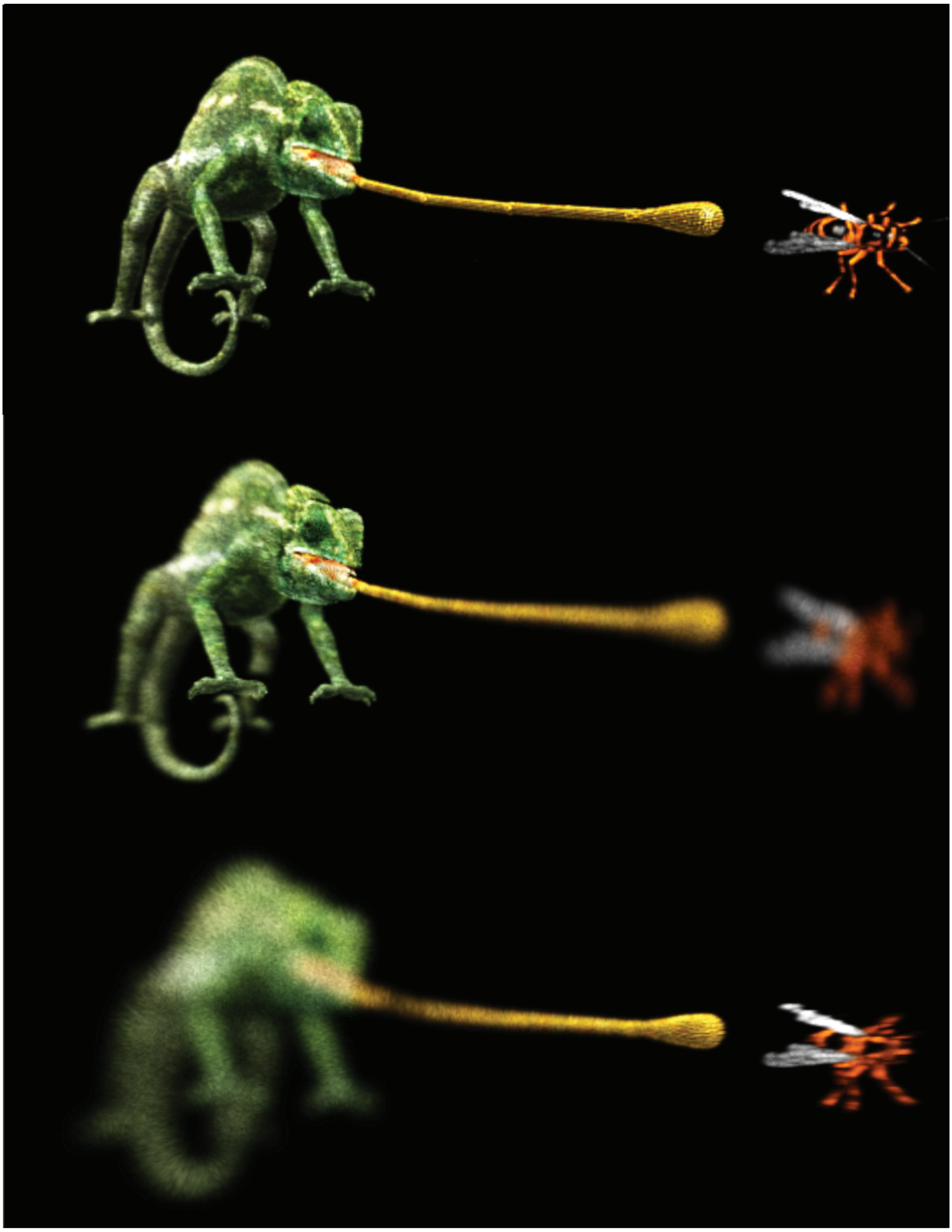

Figure 4.13: *The top scene has been calculated with a small aperture leading to a large depth of field. The two images below were computed using a bigger aperture and two different focal lengths.*

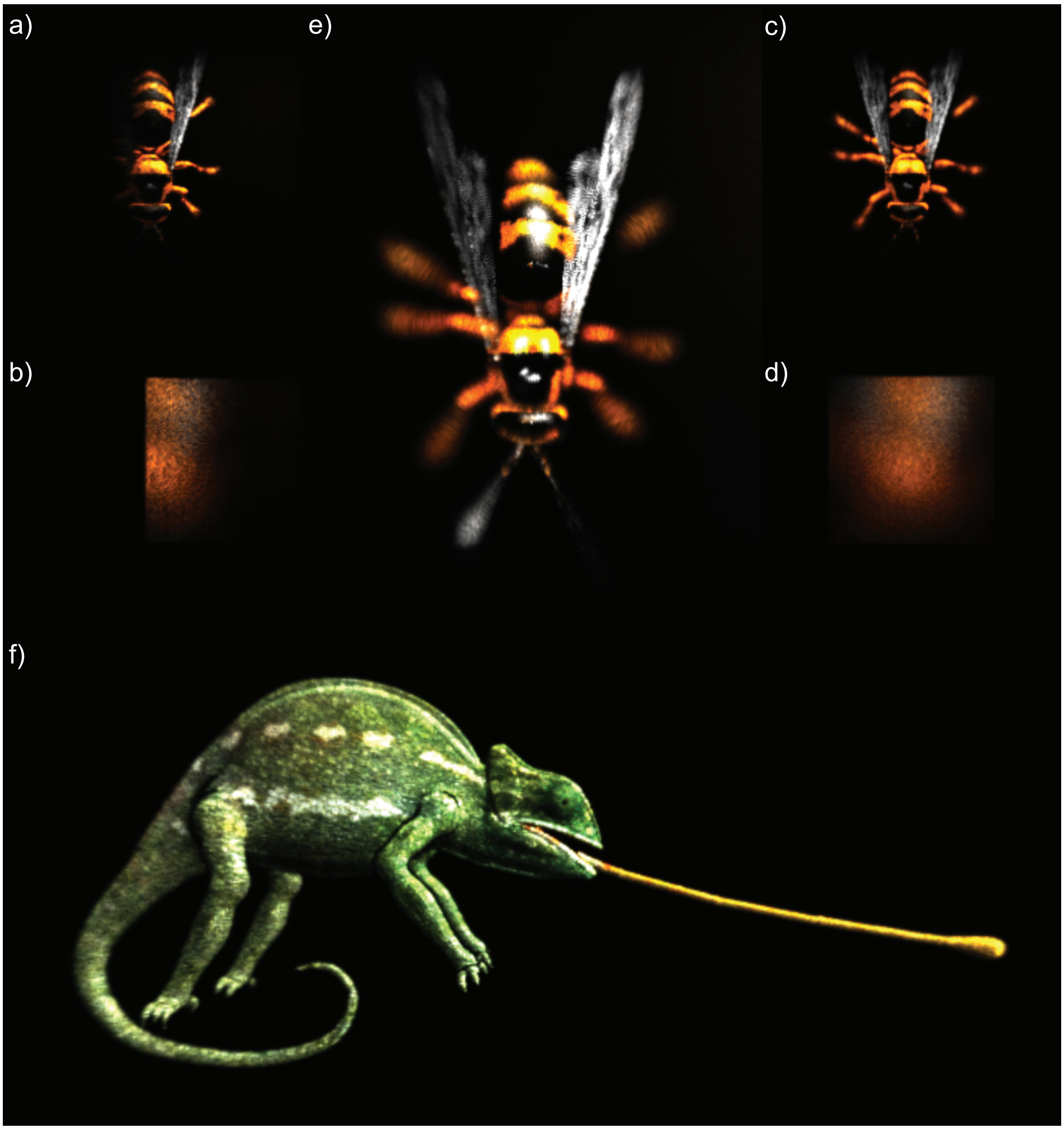

Figure 4.14: *a)b) and c)d) are pairs of images of the same scene rendered with different foci. b) and d) are images focused on the holographic plate revealing its rectangular shape. Using a bigger hologram in e) avoids evanescence of the object at the border of the hologram. The chameleon in f) is rendered with a bigger depth of field than the wasp in e).*

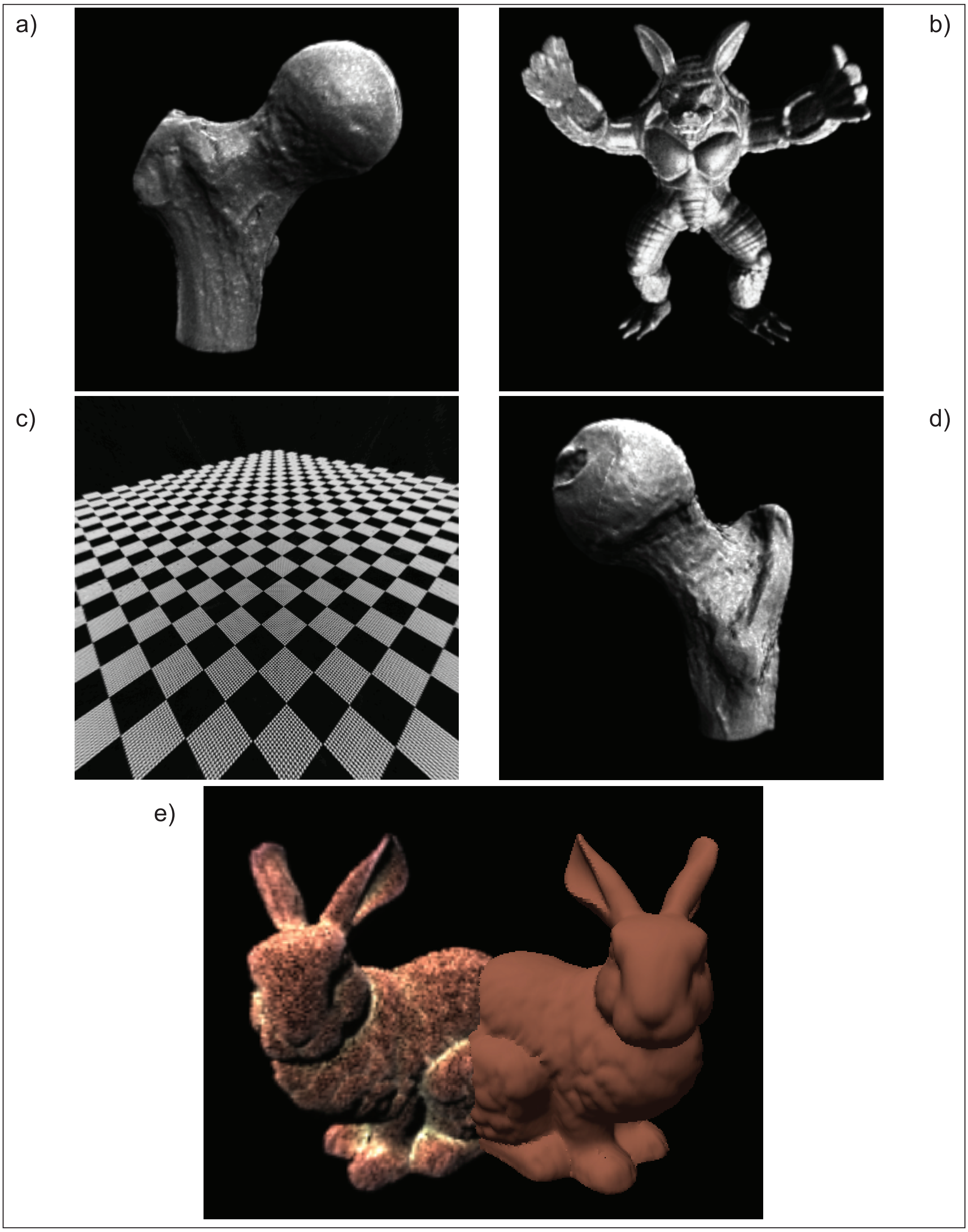

Figure 4.15: *Wave-based evaluation of point-based objects can represent surfaces of high geometric detail, as long as the sampling of the point-based object is dense enough in the first place. On the lower left c) an aliasing-free checkerboard generated using our pipeline is shown. e) is showing a scene composed of a holographic and a mesh based representation of the same object.*

Table 4.2: *These parameters were used to create the provided figures.*

Figure	Wavelength	Aperture	Object position	Camera position	"Look at" position	Focal distance
Fig. 4.13a	0.00390625	2	(0,0,-450) (0,30,-150)	(-250,20,-4)	(0,0,-320)	230
Fig. 4.13b	0.03125	16	(0,0,-450) (0,30,-150)	(-250,20,-4)	(0,0,-320)	480
Fig. 4.13c	0.03125	16	(0,0,-450) (0,30,-150)	(-250,20,-4)	(0,0,-320)	270
Fig. 4.14a	0.0625	16	(0,0,-60)	(-40,0,40)	(0,0,-450)	92
Fig. 4.14b	0.0625	16	(0,0,-60)	(-40,0,40)	(0,0,-450)	40
Fig. 4.14c	0.0625	16	(0,0,-60)	(0,0,40)	(0,0,-450)	92
Fig. 4.14d	0.0625	16	(0,0,-60)	(0,0,40)	(0,0,-450)	40
Fig. 4.14e	0.0625	32	(0,0,-60)	(0,0,40)	(0,0,-450)	94
Fig. 4.14f	0.015625	8	(0,0,-450)	(-450,20,-450)	(0,0,-450)	350
Fig. 4.15a	0.03125	8	(0,0,-450)	(0,0,0)	(0,0,-450)	350
Fig. 4.15b	0.015625	4	(0,0,-450)	(0,0,0)	(0,0,-450)	310
Fig. 4.15c	0.015625	8	(0,0,0)	(0,100,50)	(0,-64,0)	200
Fig. 4.15d	0.03125	8	(0,0,-450)	(0,0,-900)	(0,0,-450)	350
Fig. 4.15e	0.03125	8	(0,0,-450)	(-150,220,00)	(100,0,-450)	474

Although we can find quite some differences between these two representations, we observe a lot of similarities. In Chapter 6 we describe a bidirectional light-field to hologram transform, which shows their properties and possibilities of complementing their strength.

4.7 Conclusions

We presented a novel framework for holographic scene representation and rendering that integrates smoothly into conventional graphics architectures. The wave nature of holographic image generation provides a variety of interesting features including depth of field and high quality antialiasing. The phase information inherently embedded into the representation allows us to reliably reconstruct view-dependent depth and to compose digital holograms with standard 2D framebuffer graphics. To the best of our knowledge this is the first pipeline which allows the rendering of holographic representations as well as traditional geometric objects in the same scene. While holograms require a very high sampling rate in theory, our practical implementation computes high quality renderings from reasonably sized digital holograms.

In this chapter, we handled lighting evaluation and occlusion handling in a ray-based and straight forward way. In Chapter 5, we present a more efficient evaluation of the lighting and occlusion for wave-based scene evaluation. Furthermore, we present a way to light a hologram rendering in real time, providing a better possibility to analyze geometric details of the captured scene.

CHAPTER 5

Lighting and Occlusion

The holographic pipeline with the four different stages has been described in Chapter 4. The evaluation of the wavefront and the wave-based propagation build the biggest computational load. Increasing computational power, and the fast development of a multifunctional highly parallel GPGPU makes the huge, but also highly parallel computation efforts more tractable.

In this section, we introduce two fundamental operations of computer graphics to holography and wave-based rendering, namely occlusion and lighting. Most of the previous methods [ABMW06, JHS06, FLB86, KDPS01, DH98, Mat05b, ZKG07b] assume unobstructed wave propagation or simulate visibility in a ray-based way. The wavefront evaluation during the recording described in Sect. 4.3, is based on ray-based occlusion handling also. Besides being a time consuming evaluation, ray-based occlusion has to handle diffraction at object boundaries or refraction due to phase manipulating occlusions as special cases. In this section, we introduce a novel complex-valued wave-based occlusion simulating opaque, semi-transparent, and refractive objects based on scalar diffraction theory. This representation is useful not only for scene evaluation on a hologram, but also for holographic rendering, where different aperture dependent glare effects can easily be rendered.

Another way of enhancing the appearance of a scene is the introduction of lighting. The scene can be rendered for a specific view under a static illumination setting, by altering the wavefield of the primitives. This results

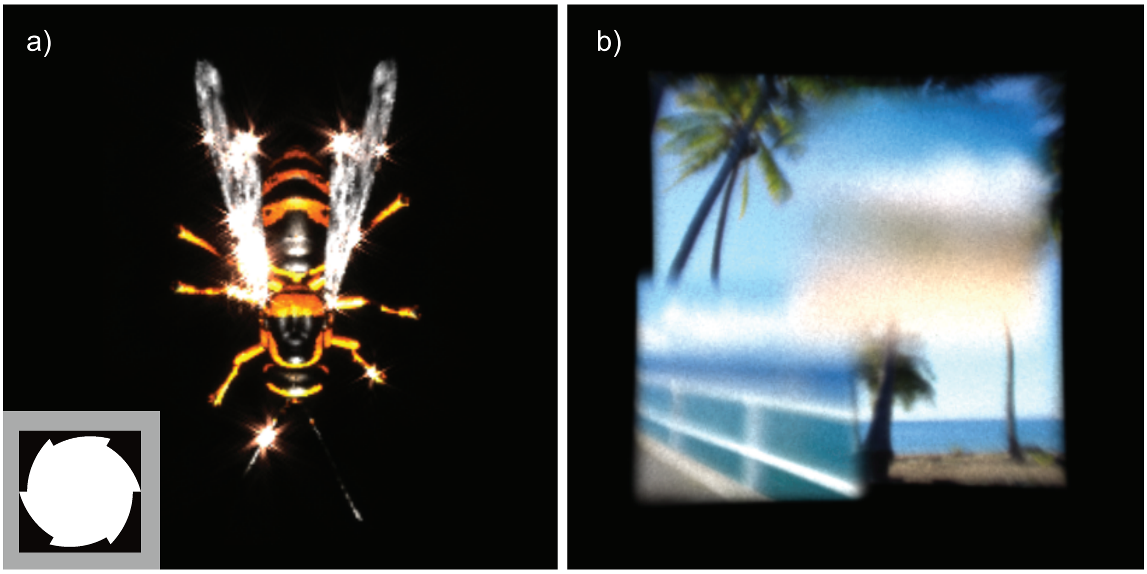

Figure 5.1: *a) A point-based object rendered using the aperture shape shown on the bottom left. b) A scene consisting of three overlapping planes placed at different depths rendered using wave-based occlusion.*

in an object, which seems to have been painted using the intensities from the light evaluation when seen from different paraxial views. We propose a novel method adapting the reflections according to different parallax views by evaluating the illumination model for every point in the scene on a subset of holographic pixel positions. Furthermore, we propose a novel lighting approach for holographic renderings that allows changing the illumination settings in real-time by evaluations on the GPU.

5.1 Overview

In this chapter, we present different ways to improve CGH, hologram rendering, and wave-based rendering (see Fig. 5.2). CGH describes the wave-based evaluation of a scene on a hologram from which different images can be created from different viewpoints. Hologram rendering and wave-based rendering evaluate the complex-valued wavefront on the camera aperture, creating an image from a single viewpoint. The former takes only one hologram as an input, while the latter creates an image of an entire scene, which could also contain a hologram.

The two areas of improvements are occlusions and lighting in a wave-based framework. First, we introduce a complex-valued occluder that can handle non-paraxial wave propagation to tilted planes, allowing arbitrary scene creation. With this wave-based approach we are able to simulate defocusing and diffraction according to scalar diffraction theory. Note that wave-based

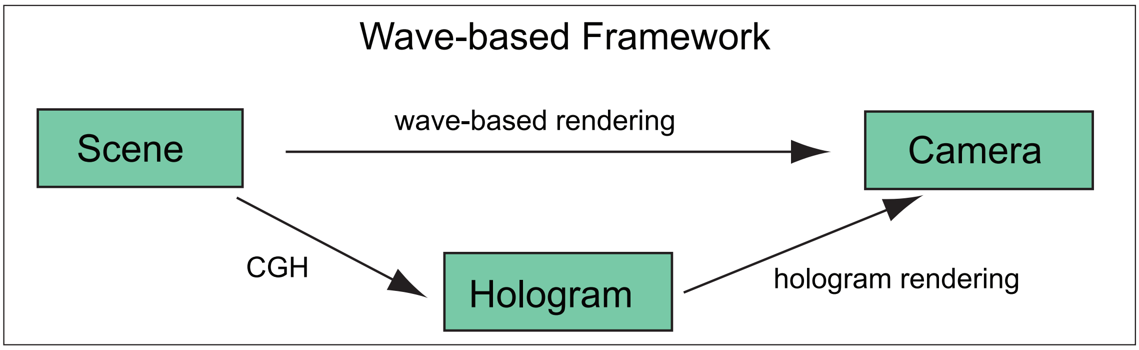

Figure 5.2: *In our wave-based framework we distinguish between three different wave propagation steps, namely CGH, wave-based rendering, and hologram rendering.*

occlusion handles all the occlusion from all possible viewing directions in a single step, as opposed to view dependent ray-based occlusion. Second, we present a lighting approach for wave-based rendering as well as a way of lighting holographic renderings in real-time. The real-time lighting allows the user to enhance the appearance of geometric details by moving a point light over the holographic rendering.

5.2 Wave-based Occlusion

Handling occlusion in a wave-based manner allows correct simulation of diffraction at object boundaries and provides view-independent occlusion handling. Building on work presented by Matsushima [Mat05a], we approximate the scene by piecewise planar objects acting as occluders of propagated waves as well as emitters of new waves. In contrast to [Mat05a], we extend the binary occluder to a complex-valued occluder in order to simulate semi-transparent objects as well as optical elements such as thin lenses.

5.2.1 Complex-Valued Occluder

In our framework an occluder is always a plane. This limitation is due to the applied propagation function based on Rayleigh-Sommerfeld, which requires the wavefield to be defined on a planar surface.

Every planar surface is split into two parts, an occluding and an emissive part, both of which are complex-valued. The **occluding part** modifies every complex value of the incoming wavefield U, including the amplitude and phase. We model the occluding part by a complex-valued texture o, which is multiplied by U. There are three classifications of occlusion: binary mask, semi-transparent object, and a thin optical element altering phase and amplitude. As shown in [Mat05a], o can be a binary mask, which simulates arbitrary planar shapes, and therefore includes any polygonal shape. Thus an occluding triangle can be represented by a binary mask occluding the incoming wave. An example is shown in Fig. 5.3a. Reducing o to real

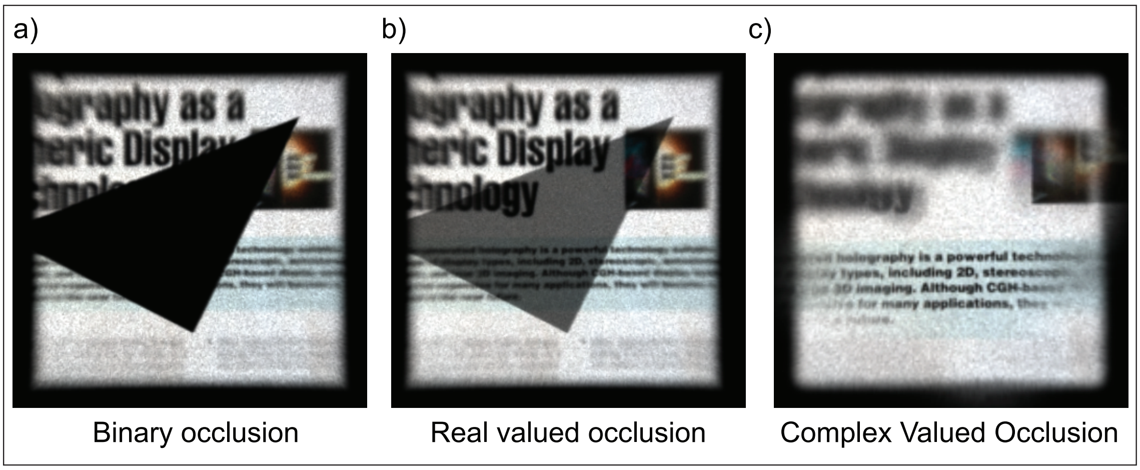

Figure 5.3: *Depending on the values of the complex-valued occluder, three different material properties, opaque (a), transparent (b), and refractive (c), can be simulated. c) has a slightly shifted occluder as well as a different focal length of the observing camera as compared to a) and b) in order to best visualize the refraction effect.*

values leads to a function that allows a modulation of the incoming amplitudes and permits semi-transparent objects, as depicted in Fig. 5.3b. By finally applying the full complex function we can alter the phase and therefore the direction of the propagating waves (see Fig. 5.3c). Choosing the values appropriately lets us simulate arbitrary thin optical elements. They have the characteristic that the incoming and outgoing positions of light are the same. One possible element is the thin lens, where the phase shift is defined by $o(x,y) = e^{\mathbf{i}\frac{2\pi}{k}\sqrt{x^2+y^2+f^2}}$, with f being the focal length of the lens, x and y the coordinates on the occluder, and k the wave number defined as $k = \frac{2\pi}{\lambda}$ with λ being the wavelength. In our implementation, we provide an interface to create a simple thin lens model, as described in [Goo68], by specifying the two radii R_1 and R_2. The output of varying thin lenses is shown in Fig. 5.4. Our system also allows complex-valued textures to be used as a complex-valued occluder input, such that the user can specify its own optical elements.

The **emissive part** is also a complex-valued function able to simulate an arbitrary wavefront propagating in any arbitrary direction in the positive hemisphere. This wavefront is added to the occluded part leading to the outgoing wavefront U_B.

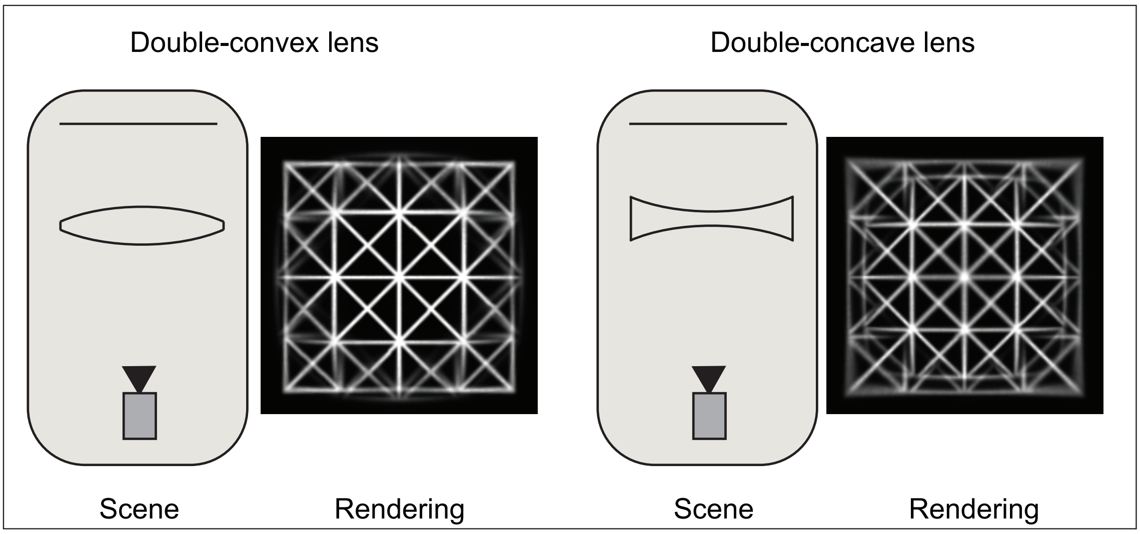

Figure 5.4: *We provide an intuitive interface for thin lens models that creates the complex-valued texture automatically.*

5.2.2 Plane Propagation

In this section, we describe the wave propagation method we apply in our implementation. This propagation is the same for any arbitrary plane pair.

The propagation $\mathcal{P}_{A\rightarrow\hat{B}}(U_A)$ from A to B consists of four steps. In the first step, the propagation cone of every plane is evaluated as shown in Fig. 5.5a. In the second step, the wavefield u_A is transformed into the angular spectrum $\mathcal{U}_A = \mathcal{F}\{U_A\}$ using the Fourier Transform. As shown in Sect. 3.3.5, the angular spectrum can be rotated by applying a Matrix M such that the rotated spectrum is given by $\mathcal{U}_A^\circ = M \cdot \mathcal{U}_A$. The new plane A° lies parallel to plane B as depicted in Fig. 5.5b, but might not be paraxial to B. In the third step, we apply an off-axis propagation from A° to $\hat{B}$. $\hat{B}$ corresponds to the extended plane B limited by the union of B and the propagation cone PC_A shaded in grey. Different ways of off-axis propagation have been presented in previous work. The most frequently used one is the Fourier-Shift approach as presented in [DH98, ZKG07b]. The Fourier-Shift propagation has the disadvantage requiring a lot of padding when the wavefield has to be propagated over big distances. Therefore, we use the propagation method presented in [SJ05], which is based on a convolution of the wavefield with a propagation function $G(\cdot)$ as shown in Eq.(5.2).

$$U_b' = U_a * G(\cdot) \tag{5.1}$$
$$= \mathcal{F}^{-1}\{\mathcal{F}\{U_a\} \cdot \mathcal{G}(\cdot)\} \tag{5.2}$$

$\mathcal{G}(\cdot)$ denotes the Fourier Transform of $G(\cdot)$. This approach evaluates

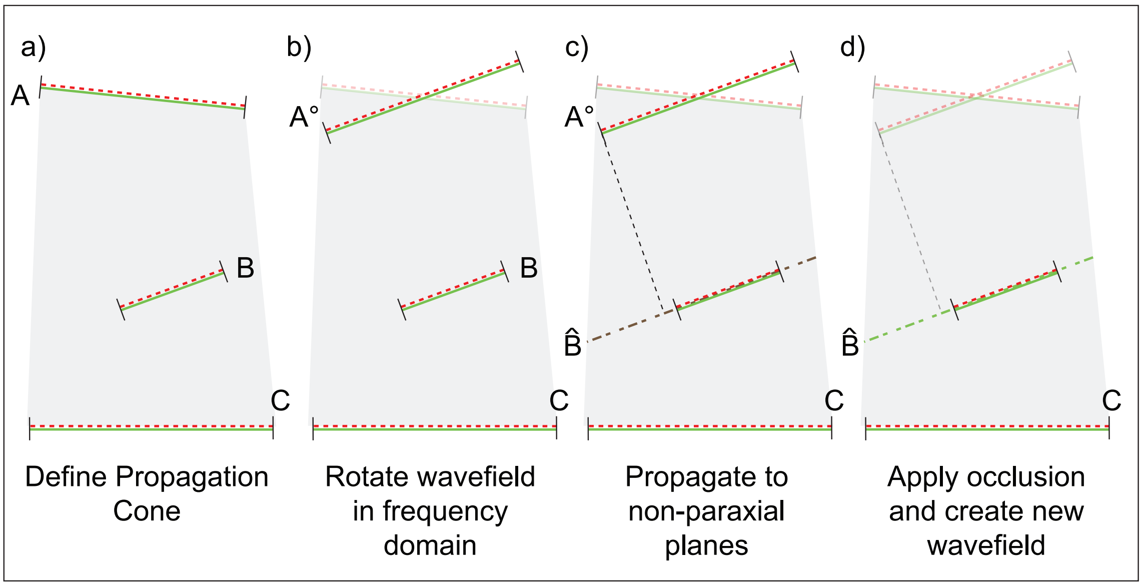

Figure 5.5: *a-d shows the required steps to propagate the wavefield from one plane to the next. Refer to Fig. 5.6 for the legend of the symbols.*

the exact Rayleigh-Sommerfeld formula without any approximations in $O(N^2 log N)$ time, where N is the number of samples of a square wavefield. Since there is an analytical form of the Fourier Transform $\mathcal{G}(\cdot)$ of the propagation function $G(\cdot)$, we only require one Fourier transform of the wavefield, a multiplication with $\mathcal{G}(\cdot)$ and an inverse Fourier Transform to propagate wavefield U_a. The convolution approach only requires a padding which is as big as the propagation function $\mathcal{G}(\cdot)$, independent of the propagation distance. All the incoming wavefields to $\hat{B}$ are named U_b'. The occlusion function o_B can now be multiplied by U_B' to apply the occlusion to the complete incoming wavefield.

5.2.3 Scene Evaluation

The scene is subdivided into planar objects, which are evaluated from back to front. Starting at the furthest plane A, as depicted in Fig. 5.6, the wavefront U_A is propagated to the next closer plane B. In order to capture parts of the wavefront U_A which are not occluded by B, we have to evaluate the wave U_B' on a bigger plane $\hat{B}$ leading to $U_B' = \mathcal{P}_{A \to \hat{B}}(U_A)$. The extent of $\hat{B}$ is limited by the propagation cone PC_A as well as the extend of B. The propagation cone PC_A is defined as the convex hull containing the hologram plane, plane C, as well as plane A. In our implementation, we choose a conservative propagation cone. It is defined as an orthographic cylinder that tightly bounds all the planes while its ground plane is coplanar with the final evaluation plane, e.g. C.

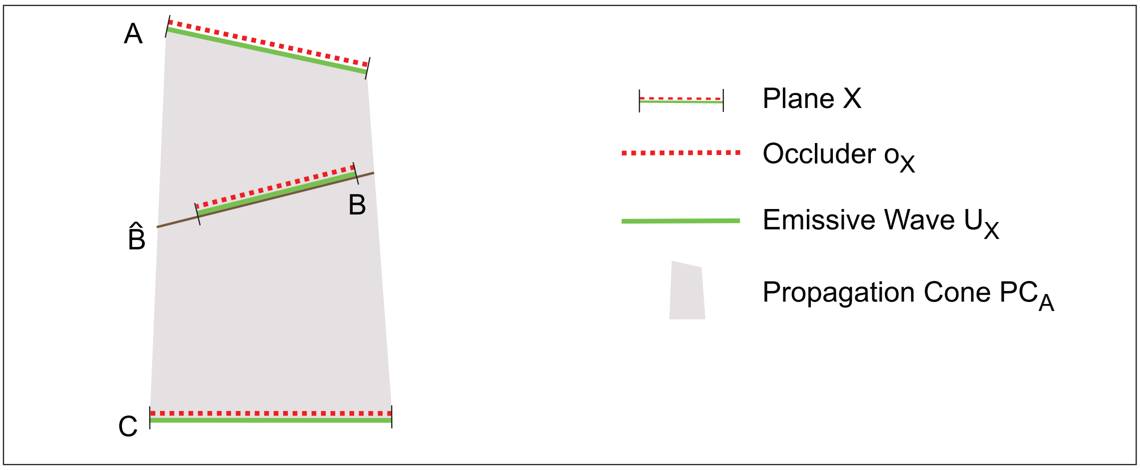

Figure 5.6: *A simple scene with three occluding planes. The wavefront is propagated from A over B to C.*

The propagated wavefront U'_B is multiplied by the complex-valued occlusion function o_B of plane B leading to $\overline{U}'_B = U'_B \cdot o_B$. The wavefield U_B emitted by B is added to $\overline{U}'_B$ and propagated to C. So, the final propagated wavefront on plane C is defined as $U'_C = \mathcal{P}_{B\rightarrow C}(\overline{U}'_B + U_B)$. Iterating over all the planes from back to front leads to the wavefront evaluation on the hologram.

Applying wave-based occlusion allows one to evaluate the visibility for all different viewing directions at once. This is saving a lot of computational time. Fig.5.7c illustrates the multiple viewpoint occlusion in a small example consisting of a point source P occluded by a plane C. In a ray-based approach, the visibility would have to be evaluated for all possible rays, such as the sample rays shown in Fig. 5.7c. In a wave-based approach, the wavefront evaluated on W (cf. Fig. 5.7a) is set to zero at the occluder C as depicted in Fig. 5.7b. Since the frequencies lost due to the occlusion correspond to directional components, the occlusion can be considered as a loss of rays in different directions. This efficiency does not necessarily pay off for single viewpoint renderings, since the ray-based occlusion has only to be evaluated once. For hologram evaluation, however, the visibility has to be reevaluated for every hologram pixel for a ray-based approach, whereas only one wave-based occlusion has to be considered to handle the visibility for all hologram-pixels.

5.2.4 Occlusion for Point-based Objects

If a hologram or a wave-based rendering of a point-based object has to be evaluated, it is important to take occlusions into account. Not doing so results in transparent looking objects [ABMW06], where it is hard to make out

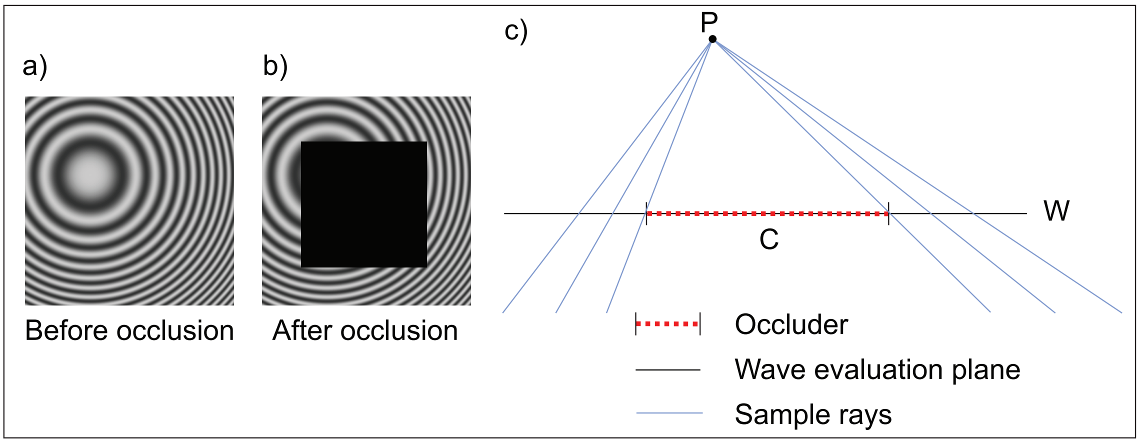

Figure 5.7: *a) shows the wavefront from a point source P evaluated on plane W. b) depicts the same wavefront from a) occluded by C. c) illustrates the setup of this small example scene.*

any shape. In Sect. 4.3.1, we have shown a way of handling occlusions of a point-based object using a ray-based approach. If we want to include point-based objects into a wave-based occlusion framework, we have to evaluate the wavefield on at least one plane right in front of the object based on Sect. 4.3.1. If the evaluation of the propagated wave has to be evaluated from every direction around the object, the wavefield has to be evaluated on six planes building a cube around the object. These planes can then be treated as any wave-based occlusion. This has not been implemented explicitly, but our implementation would allow a straight forward extension using existing components.

5.2.5 Glare

Glare is directly dependent on the obstruction of the wavefield entering the camera or the eye. The aperture shape of the camera, or eye lashes, and the pupil of the eye lead to characteristic glare effects. The aperture of a camera can easily be simulated by a complex-valued occluder, allowing the modeling of an arbitrary aperture shape combined with a thin lens model. The eye lashes can either be modeled on a separate plane than the pupil function for more accuracy, or be combined by multiplication to one occluder.

Glare occurs for all the pixels of the final image. The reason that glare effects are not seen for the entire scene observed by the human eye is due to the high dynamic range of the input. Only the very brightest points lead to visible glare. To illustrate the glare effects we scaled the low dynamic range input data by multiplying it with a user defined transformation function. Our pipeline does also handle high dynamic range input data, which is automatically leading to visible glare effects for high intensities. Before visualizing the rendered images, we are applying a tone-mapping step as presented in [RSSF02a].

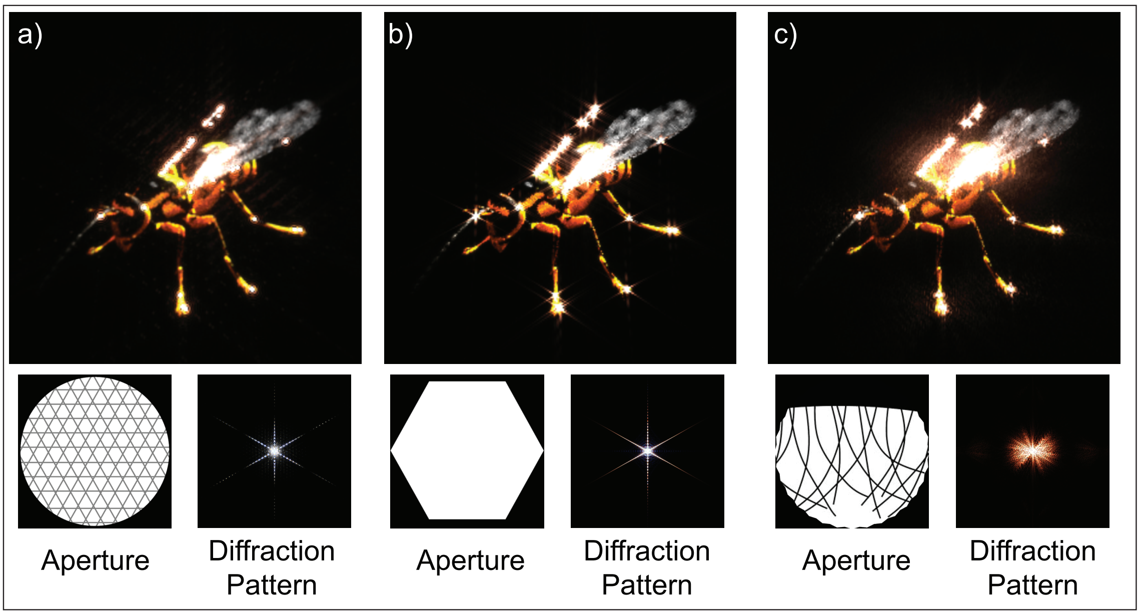

Figure 5.8: *Every aperture shape leads to a very characteristic glare pattern. a) and b) show geometrical apertures, whereas c) simulates the vision of an eye when looking through the eyelashes. On the bottom right the diffraction pattern of one point source is depicted. To fully see the diffraction pattern we recommend viewing the electronic version of the pictures. (Resolution: 512x512, Time per iteration: 8s, # of iterations: 60, more parameters: Table 5.1)*

Various aperture dependent glare functions can be simulated easily by changing the complex-valued occluder function. In Fig. 5.8, we show different aperture shapes and occlusion functions, partly taken from [KMN$^+$04], together with the resulting glare when applying the complex-valued occluder and wave propagation. Unlike in [KMN$^+$04], we generate the final image by applying the occlusion due to the aperture shape or eye lashes to the entire wavefield of the scene, and do not require a compositing step using billboards.

5.3 Lighting

Rendering full-parallax images from a hologram reveals varying reflection properties depending on the BRDF of an object. To generate the complete spectrum of reflections, the illumination model of the scene has to be evaluated per hologram pixel [JHS06, IMV07]. Depending on the evaluation of

Table 5.1: *These parameters were used to create Fig. 5.8a-c and Fig. 5.9a-c.*

Figure	Wavelength	Aperture	Object position	Camera position	"Look at" position	Focal distance	Apodize	Filter
Fig. 5.8a	0.03125	8	(0,0,0)	(0,0,500)	(0,0,0)	450	false	texture
Fig. 5.8b	0.03125	8	(0,0,0)	(0,0,500)	(0,0,0)	450	false	texture
Fig. 5.8c	0.03125	8	(0,0,0)	(0,0,500)	(0,0,0)	450	false	texture
Fig. 5.9a	0.03125	16	(0,0,0)	(-190,170,200)	(0,0,0)	280	true	none
Fig. 5.9b	0.03125	16	(0,0,0)	(-190,170,200)	(0,0,0)	280	true	none
Fig. 5.9c	0.03125	16	(0,0,0)	(-190,170,200)	(0,0,0)	280	true	none

the illumination model, the amplitude of the corresponding object point is scaled. In our implementation, we focus on two approximations for this evaluation. One evaluates the lighting only for the central position of the image aperture, adapts the amplitude of the scene, and renders the scene on the entire image aperture. This leads to the same lighting as would be obtained using a rasterizing pipeline for image generation. Fig. 5.9a shows a scene under ambient illumination, while Fig. 5.9b and Fig. 5.9c show spotlights as well as directional lights.

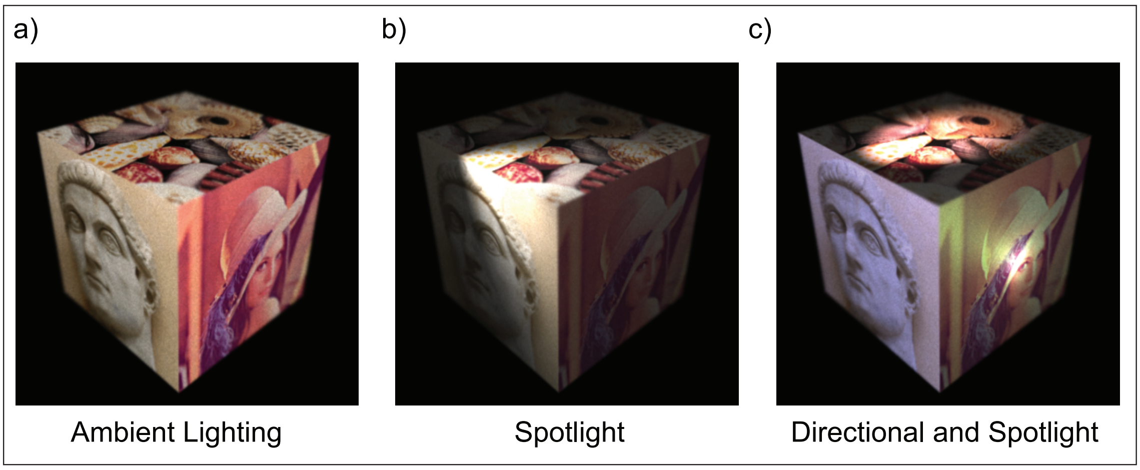

Figure 5.9: *A scene is rendered with ambient illumination a), with one spot light in b), and using colored spotlights and directional lights in c). The Lena dataset is courtesy of the Signal and Image Processing Institute at the University of Southern California. (Resolution: 512x512, Time per iteration: 100s, # of iterations: 800, more parameters: Table 5.1)*

The second approach evaluates the lighting model on a subset S of hologram pixels and interpolates the amplitudes for the hologram pixels in between. We implemented a lighting evaluation where the number of samples of S can be chosen by the user. The samples $s_i \in S$ are placed on a regular grid over the entire aperture or hologram, as depicted in Fig. 5.10a, while using a bilinear interpolation for the illuminations corresponding to hologram pixels in between. This way, the lighting of the scene changes with varying viewpoint, while limiting computational costs.

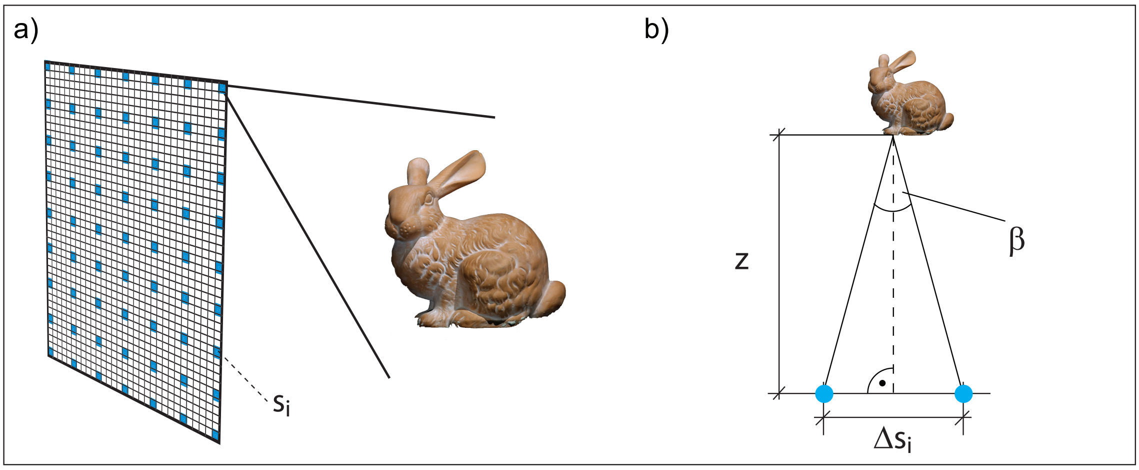

Figure 5.10: *a) the lighting is only evaluated for a subset (marked in blue) of the hologram pixels. The maximum frequency that can be captured is directly dependent on the angle β depicted in b).*

The maximal measurable frequency of the BRDF is limited by the maximal sampling distance Δs_i between all positions s_i and the distance z of the closest point of the scene to s_i as depicted in Fig. 5.10b. The angle β is defined as

$$\beta = 2 \cdot arctan \frac{\Delta s_i}{2 \cdot z}. \tag{5.3}$$

A very similar limitation applies to the employed camera. For points in focus, all the rays passing through the aperture are integrated to a final pixel intensity. This implies that the maximum sampling angle limiting the maximal frequency is determined by the aperture size and the distance to the object. This is also true for a physical camera capturing a scene. In Fig. 5.11, we show a rendering of the same hologram using a camera with different sized apertures. The size of the specular reflection in red shrinks with a smaller aperture, and therefore, the maximal measurable frequency increases.

5.3.1 Lighting of Hologram Renderings

Holograms are recorded using a static illumination setting, consisting either of a laser source or an arbitrary illumination for Computer Generated Holog-

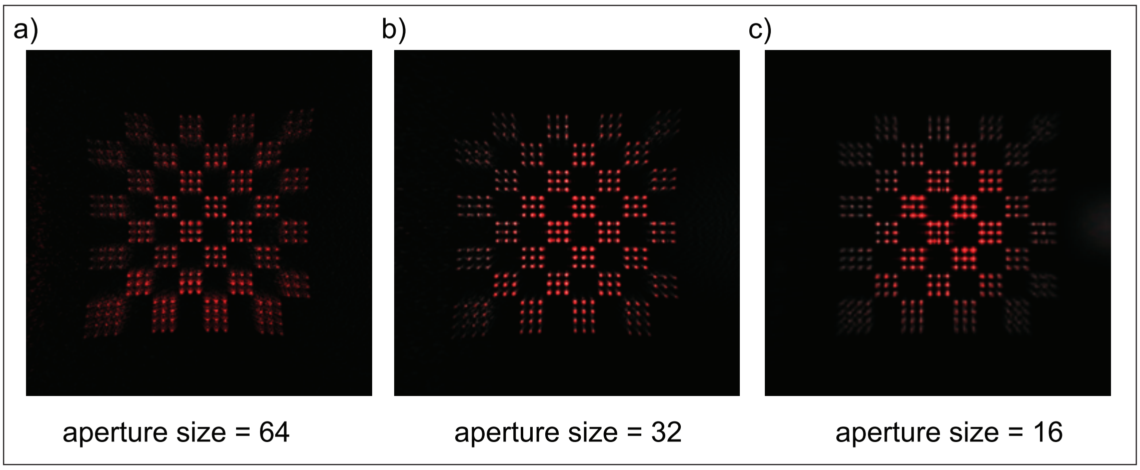

Figure 5.11: *By decreasing the aperture size from a) to c), higher frequencies of the BRDF become visible, since the area of integration per pixel is decreasing. (Resolution Hologram: 2048, Resolution Image: 256x256 for aperture size 16, Time per iteration: 20s, # of iterations: 1)*

raphy (CGH). However, when rendering holograms, the observer might want to choose a different illumination for reasons such as revealing more information about geometrical properties of the holographic scene. Therefore, we propose a novel real-time lighting approach for hologram renderings.

In a first step, the image of a hologram is rendered by propagating the wavefront of the hologram to the image aperture. Using the approach presented in Sect. 4.4.3, we are able to extract a color image as well as its corresponding depth map. Based on these values, a local lighting model requiring the normal and reflectance properties at every point of the object surface, is evaluated.

The points in the scene P_i can be computed from the corresponding points in the depth map d_i by applying a back projection matrix M_P^{-1} such that $P_i = M_P^{-1} d_i$.

We consider two different approaches for the evaluation of the normal vector. The first approach requires three points P_1, P_2 and P_3 on the surface and computes the normal $\mathbf{n}_0$ by evaluating the cross product of the two vectors defined by the chosen points as $\mathbf{n}_0 = (P_1 - P_0) \times (P_2 - P_0)$. Although this method is very fast, it is also very sensitive to noise. Therefore, a filter may need to be applied to avoid a speckled depth map. The second approach fits a tangential plane defined by

$$z - z_0 = a \cdot (x - x_0) + b \cdot (y - y_0) \tag{5.4}$$

through a given set of points, such that the normal of this plane can be used as the surface normal. The parameters a and b can be evaluated by

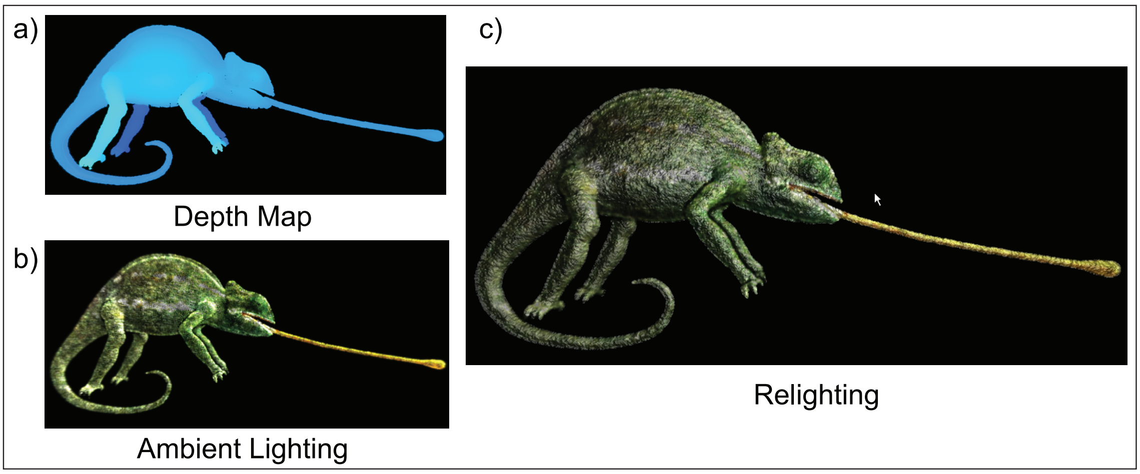

Figure 5.12: *Using the reconstructed depth map a) and the color map b) the lighting can be performed in real-time c). This scene used the same parameters as Fig. 4.14f.*

solving a least squares minimization problem leading to a surface normal $\mathbf{n}_0 = (a, b, -1)^T$. In our implementation, we provide a plane fitting for five, nine or thirteen neighboring points. A possible alternative for normal estimation would be to use a MLS-based reconstruction [GG07].

Since this kind of scene illumination is based on the rendered image stored in the color buffer, and on the depth map stored in the depth buffer, the evaluation of the lighting model can be performed entirely on the GPU using pixel shaders. The user can interact with the scene by changing the position of the point light source or the direction of the rays of a directional light source in real-time.

5.4 Results

We integrated the occlusion handling and lighting into a bigger wave-based propagation framework. An easy setup of the scene consisting of different lighting possibilities as well as different object representations can be defined in an XML-format file. The application is written in C++ and makes frequent use of the Graphics Hardware programmed using HLSL. Performing computations on the Graphics Card increases the speed for point source evaluation, occlusion handling, and lighting considerably. The renderings are generated using an Intel Core 2 CPU 6700 at 2.66GHz with a NVidia 7950GT graphics card.

5.4.1 Occlusion

In Fig. 5.13 we created a scene consisting of three textured planes occluding each other. The focal length is set to each plane in turn in Fig. 5.13a to Fig. 5.13c. Note the correct handling of the occlusion boundary for the defocused front plane. This means that points laying directly behind the edge of the front plane will have a partial influence on the final rendering, which corresponds to the physical image creation using a limited aperture size. Although only one propagation has been applied per plane, we obtain all possible occlusion directions for viewpoints placed on the target aperture, resulting in view-independent occlusion handling. This is true independent of the complex-valued function of the occluder. Fig.5.14 shows renderings for different complex occlusion functions leading to refraction effects. Being able to model any planar occluder leads to a very intuitive and easy simulation of arbitrary aperture shapes allowing visible glare effects for bright scene points (Fig. 5.8).

5.4.2 Lighting

Most of our scenes are direct wave-based renderings, meaning no hologram is created for the image generation. When creating an image for a specific viewpoint, we approximate the lighting of the scene by only evaluating the illumination for this viewpoint (Fig. 5.9). We have shown that this is an approximation which only holds for a theoretical pinhole camera. A realistic camera simulation requires the lighting evaluation for every point on the aperture. The variation over the aperture of the BRDF of a surface point is integrated to one pixel value for points in focus. This effect can be seen as a low-pass filtering of the BRDF depending on the aperture size. The change in the specular reflection is clearly visible for a scene rendered from a hologram in Fig. 5.11, which has been evaluated using our lighting evaluation described in Sect. 5.3.

We present several examples of real-time lighting of hologram renderings in Fig. 5.15. Altering the lighting can be used to reveal more information about the surface shape as well as to generate more pleasing images.

5.5 Limitations

For highly complex scenes, the wave-based occlusion has several limitations. In order to place the planes in a back to front order, some planes have to be split into smaller ones, leading to an even bigger number of required propagations. Furthermore, the scene evaluation can lead to propagations of very

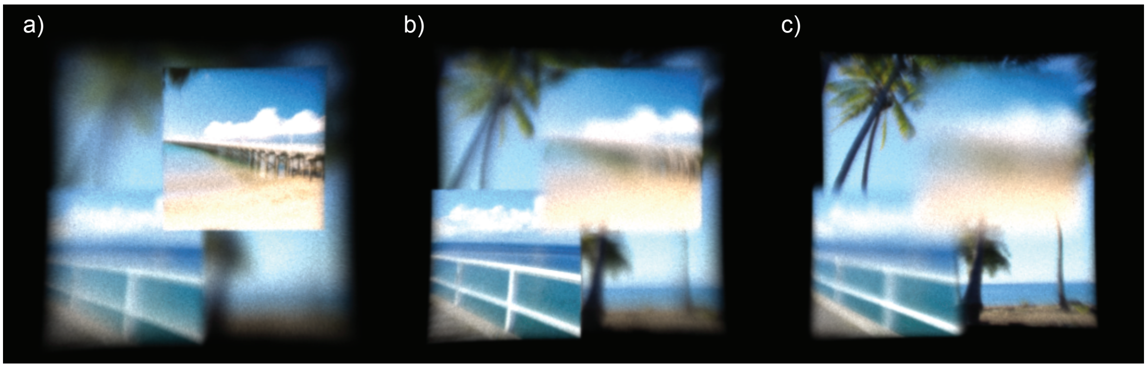

Figure 5.13: *Wave-based occlusion treats the plane boundaries correctly, even for defocused objects. The three images depict the same scene rendered with three different focal lengths. (Resolution: 512x512, Time per iteration: 248s, # of iterations: 700, more parameters: Table 5.2)*

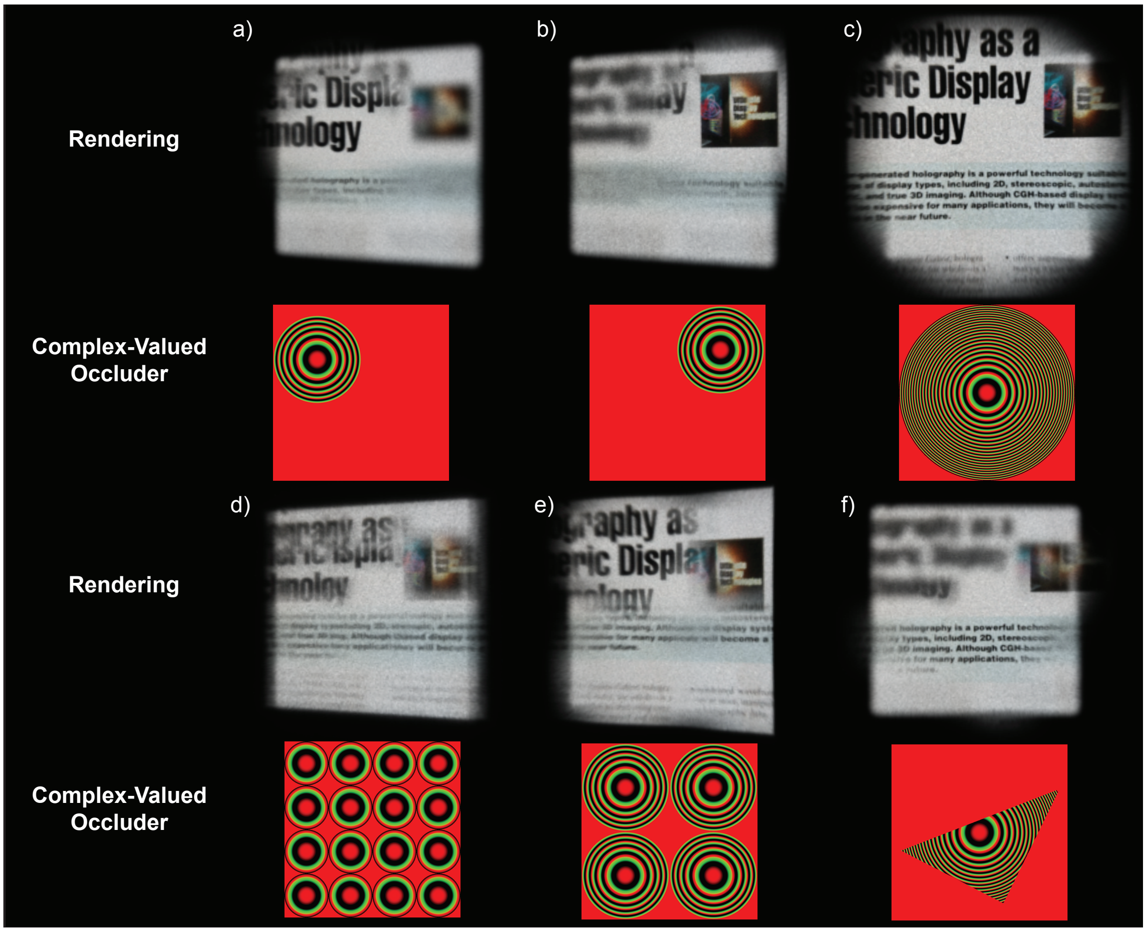

Figure 5.14: *The complex-valued occluders are illustrated with the red and green textures, where the red channel is used to store the real value and the green channel the imaginary value. The renderings are based on a tilted textured aperture behind an occlusion plane. (Resolution: 512x512, Time per iteration: 195s, # of iterations: 500, more parameters: Table 5.2)*

Table 5.2: *These parameters were used to create Fig. 5.8a-c and Fig. 5.9a-c.*

Figure	Wavelength	Aperture	Object position	Camera position	"Look at" position	Focal distance	Apodize
Fig. 5.13a	0.0625	32	(0,0,-20) (0,0,-15) (0,0,-10)	(0,0,40)	(0,0,0)	53	true
Fig. 5.13b	0.0625	32	(0,0,-20) (0,0,-15) (0,0,-10)	(0,0,40)	(0,0,0)	57	true
Fig. 5.13c	0.0625	32	(0,0,-20) (0,0,-15) (0,0,-10)	(0,0,40)	(0,0,0)	63	true
Fig. 5.14a-f	0.0625	16	(0,0,70) (0,0,0)	(0,0,100)	(0,0,0)	260	true

big wavefields. Although we tried out a subtractive approach, where the evaluation of the planes never got bigger than the planes already existing in the scene, it is limited to real-valued occluders not altering the direction of light propagation. A different way of scene evaluation as well as the use of a hardware implementation of the FFT, such as the one presented by NVIDIA® CUDA™, would reduce propagation time considerably.

5.6 Conclusion

We presented several novel methods to improve wave-based image generation and pointed out their usefulness for CGH. Handling occlusion based on the propagation of waves allows view independence and requires only one propagation step. Furthermore, defining the properties of an occluder becomes very intuitive, since the transmission function can be described in the form of a complex-valued texture. In this way, opaque, transparent, and refractive objects can be integrated as new primitives when creating a wave-based rendering. Additionally, we showed that the occluder interface can be employed to generate diffractive effects of the camera aperture causing glare in the final rendering. The arbitrary aperture shape can also be defined by a texture.

Furthermore, we presented a way of evaluating the lighting of the scene for a full-parallax hologram, and demonstrated the influence of the aperture size

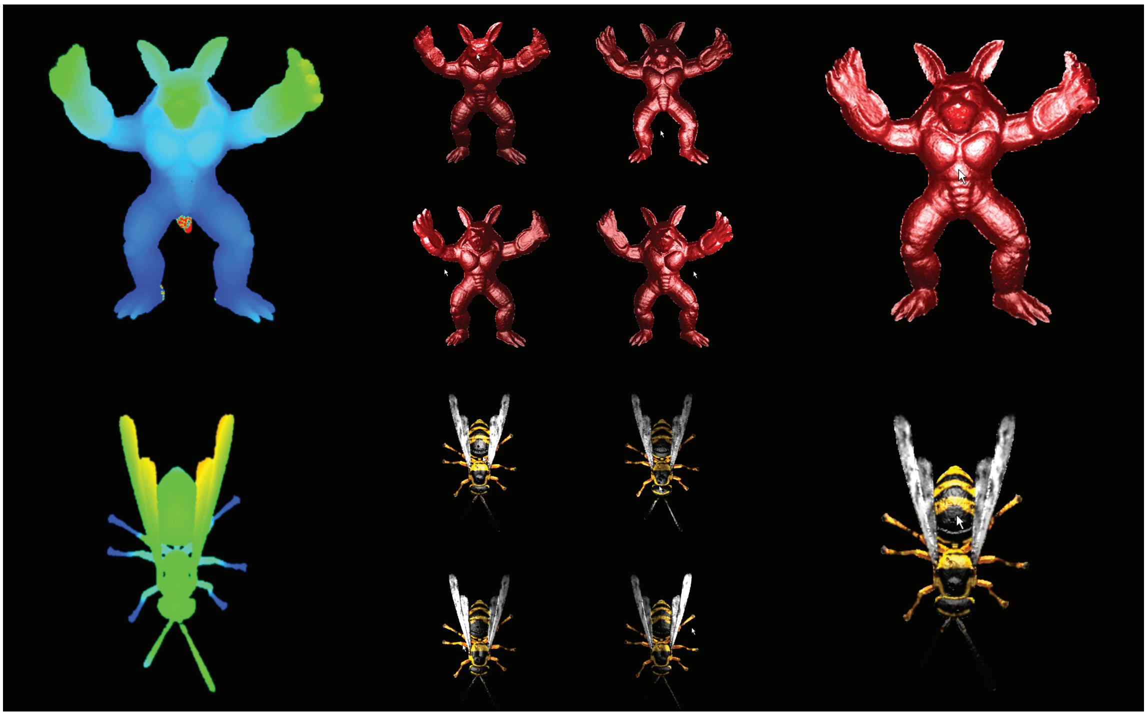

Figure 5.15: *On the left, the color-coded depth map of the scene is shown, which is used as a base for the normal estimation. The scene can then be rendered for different light positions in real-time. The relighting of the armadillo is done based on the same parameters as Fig. 4.15b, whereas the parameters for the wasp were taken from the scene displayed in Fig. 4.14e.*

of the camera on the highest measurable frequency of the BRDF. Computing the lighting only for a subset of hologram pixels yields a big speed up in hologram evaluation while approximating the BRDF only for very high frequencies.

Rendered scenes from holograms suffer from a fixed illumination scheme. Our novel real-time hologram lighting of holograms allows interactive user-defined lighting based on a normal map created from the reconstructed depth field of the hologram. By altering the lighting, geometric details become more visible, which could be used for analyzing object surfaces as well as creating appealing renderings. Furthermore, the hologram could be integrated in a new scene and rendered under arbitrary illumination.

Although, our approach is leading to more appealing and realistic looking wave-based renderings, it is still not capable of simulating real scenes with all the details. In Chapter 6, we present novel ways of capturing a hologram under white light conditions. This can be achieved by introducing a bidirectional light field - hologram transform. Using a light field camera, a scene can be captured in a single shot. This allows to capture none static scenes as well.

CHAPTER 6

Light Field - Hologram Transform

So far, we have presented different ways to create computer generated holograms. CGH allows to generate holograms from virtual scenes, which can further be displayed on a holographic screen. Virtual scenes have the advantage to show objects, which do not exist in reality, or which could not be seen, due to their size, their surroundings, or their accessibility. However, the representation of realistic looking objects using virtual scenes is more difficult. Different models for light transport have been presented. Throughout computer graphics rendering, geometric (ray) optics is frequently being adopted as a physical model of the image formation process, for some very good reasons: geometric optics is a mathematically simple and yet surprisingly powerful model that is able to explain and also quantitatively describe most optical effects that we can perceive with our eyes. Given all necessary information about a scene, geometric optics is regularly sufficient to achieve fast as well as realistic rendering performance. Nevertheless, geometric optics also possesses a number of limitations. Most prominently, any scene to be rendered must be represented rather inelegantly in terms of 3D geometry, texture, and local reflectance characteristics. Obtaining these separate descriptions of real-world scenes proves tedious, time-consuming, and expensive.

To overcome this drawback, image-based rendering techniques, and specifically light field rendering [LH96] have been proposed. In light field rendering, a (large) set of photographs taken from various different positions all around the scene are used to represent the visual appearance of the scene. Unfortunately, very large numbers of photos are needed to guarantee aliasing-free light field rendering results [CCST00, IMG00], which is why subsequent image-based rendering techniques again resort to additional (approximate) geometry information to interpolate views from much reduced numbers of photographs.

With holograms, an alternative representation of visual scene appearance is known. Based on wave optics, holography is mathematically considerably

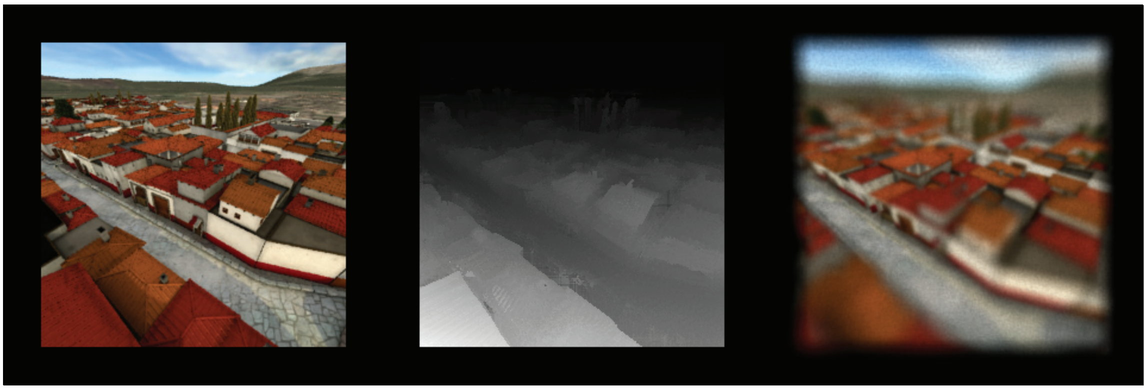

Figure 6.1: *The left most picture shows the light field input from which the depth field in the middle is created. The light field can further be transformed into a hologram, as shown on the right, by using the reconstructed depth field.*

more demanding than geometric optics. The need for monochromatic, coherent illumination during acquisition and speckle patterns during display additionally seem to argue against considering holograms in the context of computer graphics rendering. On the other hand, holograms represent visual scene appearance in the most elegant way, containing any possible view from a continuous viewport region without aliasing. In many ways, holograms are complementary to light fields, see Tab.6.1. Geometric optics turns out to be simply the approximation of wave optics in the limit of infinitesimally small wavelength [BW59].

One of the major deficiencies of holograms, is the requirement of monochromatic light during recording. Although it can be simulated for all kinds of objects during CGH, it is a very hard constraint when capturing a real scene. And exactly this property, to be able to capture a real scene instead of modeling it, has proven to be very valuable in creating realistic looking images in computer graphics. As already mentioned, one of the most prominent image-based techniques is the light field representation.

In this thesis, we propose to use both the light field and the hologram representation of a scene's visual appearance in tandem. Our goal is to perform processing steps on that respective representation for which the processing step can be done easier, faster, or more accurately. To switch between ei-

Table 6.1: *Advantages and Disadvantages*

Compare	Hologram	Light Field
Function dimension	2D	4D
Light representation	wave	ray
Single-shot acquisition	Yes	Yes
Refocusing	Yes	Yes
Natural light recording	No	Yes
Speckle free	No	Yes
Real time rendering	No	Yes
Aliasing free	Yes	No
Scene independent sampling	Yes	No
Phase information for depth encoding	Yes	No
Recording without optical elements	Yes	No
Compression	Yes	Yes
Combination with geometrical representations	Yes	No

ther representation, we describe how to map from the hologram to the light field representation, and vice versa. These mappings give us the freedom to exploit the advantages of either representation. The advantages and limitations of light fields and holograms are summarized in Table 6.1.

The core technical contributions are two functions to transform between holograms and light fields. A key ingredient of the forward transform includes a novel algorithm to reconstruct depth from input light fields featuring almost arbitrary BRDF by exploiting 4D epipolar volume representations. Our mapping functions provide a solid theoretical basis to record full-parallax holograms by means of light field cameras, and they enable us to convert any input light field into a hologram for output on future holographic displays. In addition, the wave optics representation of the hologram allows for a variety of sophisticated processing algorithms. Furthermore, since the correct depth gets encoded into the hologram, the images created from the hologram do not show any ghosting artifacts, and operations such as refocussing and varying depth of field are still possible. The highest frequency of the BRDF reconstructed from the hologram will however, not be higher than the one captured by the light field.

In Sect. 6.1, we discuss the properties inherent to the light field and the hologram representation. An overview of our framework is presented in Sect. 6.2, followed by the description of the forward transform from the light field to the hologram in Sect. 6.3, operations on the hologram in Sect. 6.4, and the inverse transformation elaborating the essential physical characteristics in Sect. 6.5. To demonstrate the advantages of our proposed dual light field-hologram representation, we present results for real and synthetic light fields as well as digitally recorded holograms (DRH) in Sect. 6.7 before we conclude with an outlook on future work.

6.1 Representation

Holograms and light fields have been parameterized in numerous ways. In Sect. 6.1.1 and Sect. 6.1.2, we describe the specific representations for the hologram and the light field, which are used as a basis for the transform.

6.1.1 Parametrization of Light Fields

There exist different parameterizations of light fields. Two of them are best suited to describe the transform [Lüt04]. We either use the popular two-plane parametrization $LF(u,v,s,t)$ as presented by Levoy [LH96] (see Fig. 6.2b) or consider the light field as angular parametrization $LF(u,v,\theta,\phi)$ dependent of position on the uv-plane and direction dependent on θ and ϕ as in Fig. 6.2c.

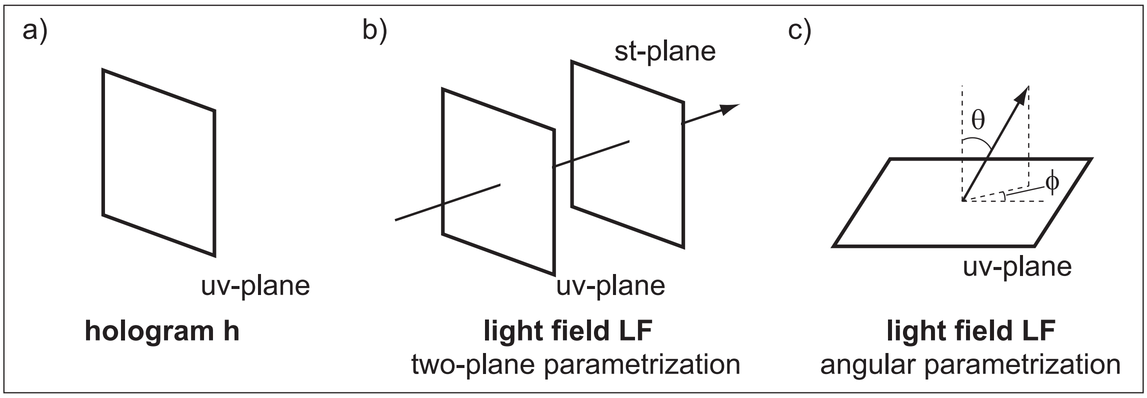

Figure 6.2: *a) depicts the representation of a hologram. b) and c) show two different representations of a light field.*

6.1.2 Parametrization of Holograms

In general, a hologram is a real-valued function describing the intensity of the static interference pattern of a complex-valued object wave with a complex-valued reference wave. The original object wave can be reconstructed from the hologram. In the following we will use the term "hologram" in the spirit of the *wavefield*, which is a complex-valued wave function $U(u,v)$, instead of a real-valued intensity field. This simplification does not have an influence on the transformation from a hologram to a light field, since the complex-valued wave function can be reconstructed from a real-valued hologram.

6.2 Light Field Mapping Pipeline

We describe a pipeline based on a novel mapping $\mathbf{M}$ and its inverse $\mathbf{M}^{-1}$giving the possibility to transform a light field into a holographic representation and vice versa. The holographic data representation is similar to a light field in that the hologram as well as the light field measure the light propagated through a plane in space into all directions. The input to

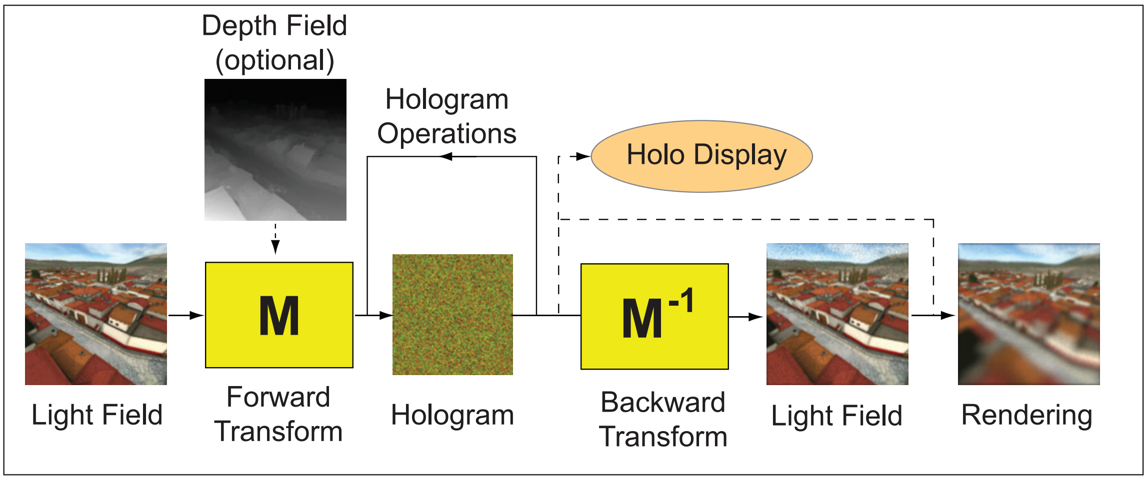

Figure 6.3: *The input to the pipeline is a pure light field, with an option of providing an accurate depth map for every view. Using* $\mathbf{M}$ *the light field is transformed to a holographic representation, where functions such as compression and progressive transmission can be applied directly on the hologram. A mapping function* $\mathbf{M}^{-1}$*allows an inverse transformation into a light field, from which images from different view points can be rendered in real time.*

our pipeline depicted in Fig. 6.3 is a pure light field without any depth information. $\mathbf{M}$ (cf. Sect. 6.3) transforms the light field into a holographic representation. A core ingredient of $\mathbf{M}$ and a core contribution of this thesis, is a method to extract depth from the input light field (cf. Sect. 6.3.1). If accurate depth information is available for the light field it can optionally be added to the input of $\mathbf{M}$, increasing the quality of the holographic representation as described in [Hal94, CCST00]. Different algorithms can be applied to the manipulation of the hologram, such as compression, progressive transmission, wavefront propagation simulating diffractive effects and others. In Sect. 6.4, we present a rendering technique, a compression algorithm and study effects of loss of data. Arbitrary parallax images can be rendered efficiently from the holographic representation as long as the COP of the virtual camera lies on the holographic plane. For arbitrary viewpoints we present an inverse mapping $\mathbf{M}^{-1}$(cf. Sect. 6.5), transforming the holographic representation back into a light field representation, from which it can be rendered to arbitrary viewpoints in real time.

6.3 Forward Mapping

The forward mapping $\mathbf{M}$ takes a pure light field and maps it to a hologram. $\mathbf{M}$ consists of two main steps, namely a depth reconstruction from light fields with almost arbitrary BRDF, and hologram evaluation based on the reconstructed depth proxy and the light field if available. An optional accurate depth field can be added to $\mathbf{M}$, making a depth reconstruction of the light field obsolete, speeding up the mapping and slightly enhancing the accuracy of the forward mapping in case of inaccurate automatic depth reconstruction. Intermediate steps of the forward mapping are shown in Fig. 6.4.

Since such a depth map is usually not at hand, we present a novel depth reconstruction method from light field data based on a 4D Epipolar Volume representation.

6.3.1 Depth Reconstruction from Light Fields

Our method takes advantage of the full light field information and redundancy captured by cameras aligned to a matrix instead of a line. We call the resulting per view depth map the *depth field*. Throughout this chapter, depicted depth maps are in fact always disparity maps.

Representation The key advantage of the Epipolar-Plane Image (EPI) representation $EPI(u,s)$ (cf. Fig. 6.5a) is the collocation of corresponding pixels from different views on one line l_c. In case of Lambertian scenes, such lines are consisting of a constant color in absence of occlusions. Furthermore, the inclination of l_c is dependent on the depth. Line l_c corresponds to a plane

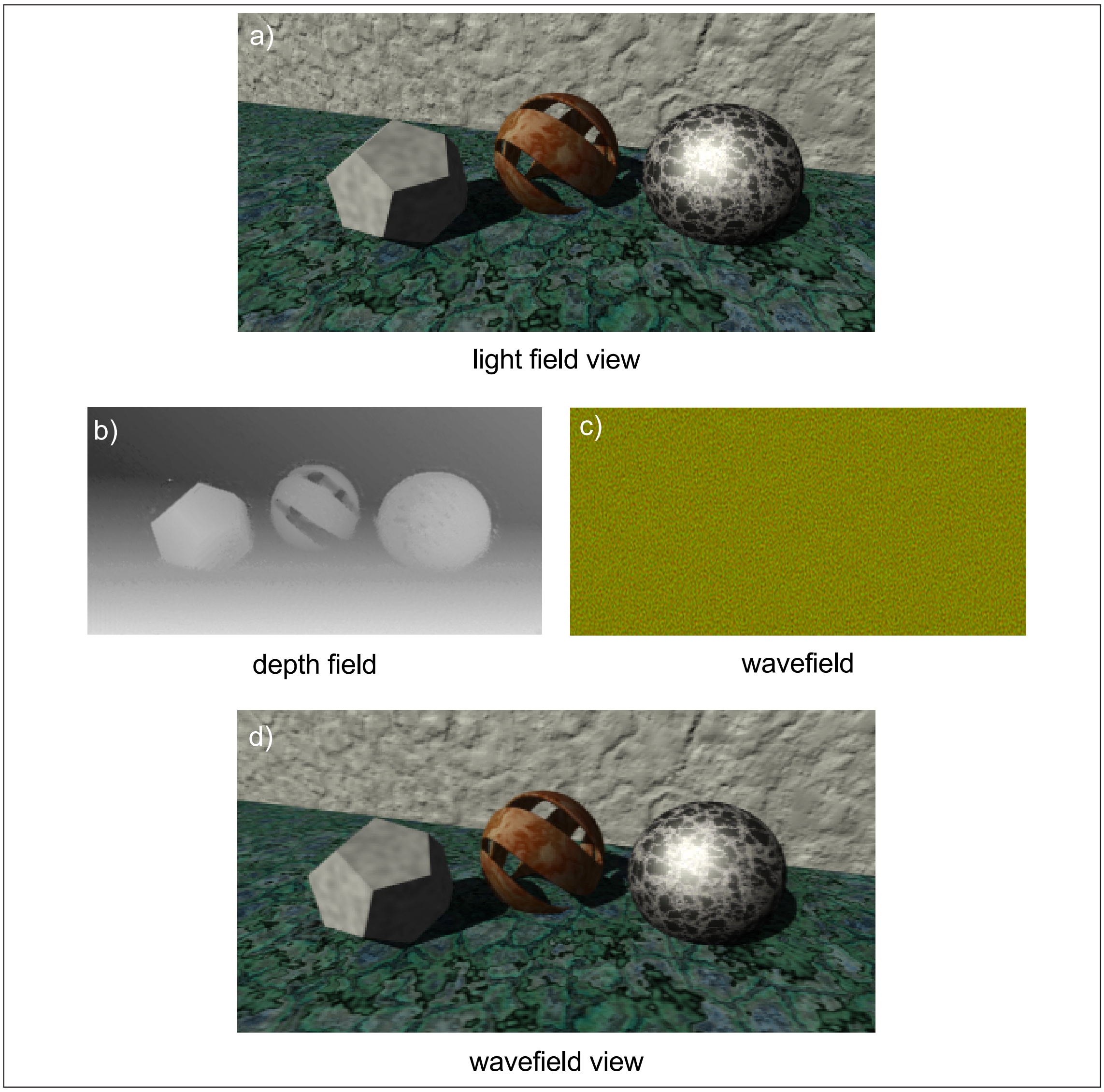

Figure 6.4: *For every input view of the light field a) a depth map b) is reconstructed, which is used to evaluate the wavefield c). A rendering of the generated wavefield is depicted in d).*

p_c in our 4D Epipolar Volume $\widetilde{EV}(u,v,s,t)$. In all our examples $\Delta v = \Delta u$ the inclination in s and t is the same. The plane p_c can also be interpreted as the set of all the samples of the plenoptic function of one point P sampled by the light field.

Discretization Assuming a continuous light field, every point in space leads to a continuous line in the EPI as long as occlusion is ignored. However, the rasterization of l_c (cf. Fig. 6.5b) at an inclination smaller than 45° will lead to dotted lines, which are hard to be recognized using any filter or edge detector. The same problem arises when trying to fit the inclination of planes p_s in $\widetilde{EV}(u,v,s,t)$. Therefore, we compute a sequence of sheared spaces by progressively changing the shear factor s_{xy} corresponding to an inclination, such that $\widetilde{EV}'(u',v',s',t') = \widetilde{EV}(u,v,s + s_{xy} \cdot u, t + s_{xy} \cdot v)$ and check only shear planes $p_s = \widetilde{EV_s}'(u',v',s',t')$ orthogonal to the s and t direction. The reconstructed depth precision can be improved by increasing the number of shears.

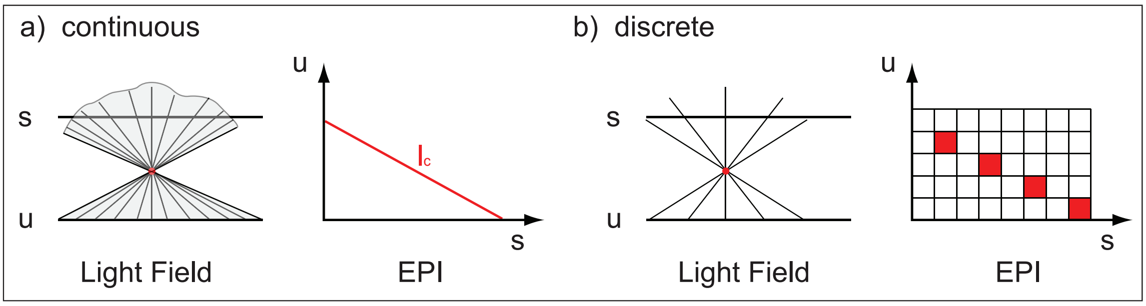

Figure 6.5: *a) shows the continuous light field and its corresponding EPI. b) shows a discrete light field and its corresponding discrete EPI.*

Frequency Minimization Criteria In the case of a Lambertian scene the color of the plane p_s is constant, as long as its pixels stemming from every light field image $LF(u,v,\cdot,\cdot)$ corresponds to the same point in the scene. This consistency criterion can be evaluated by minimizing the variance over p_s. In case of arbitrary BRDFs, the variance will fail (cf. Fig. 6.7) in most of the depth reconstructions of specular objects. By comparing Fourier Transforms of different shear planes containing $\widetilde{EV}'(u,v,s,t)$, we observe predominantly low frequencies if the shear corresponds to the depth of the pixel at that position, albeit points showing specular highlights. This is based on the fact that the specular highlight becomes smaller for non-matching shear planes leading to higher frequencies in the spectrum. Additional texture magnifies the high frequencies for non-matching shears as well. Therefore, we introduce the following novel criteria based on the Fourier Transform $\mathcal{F}$, which

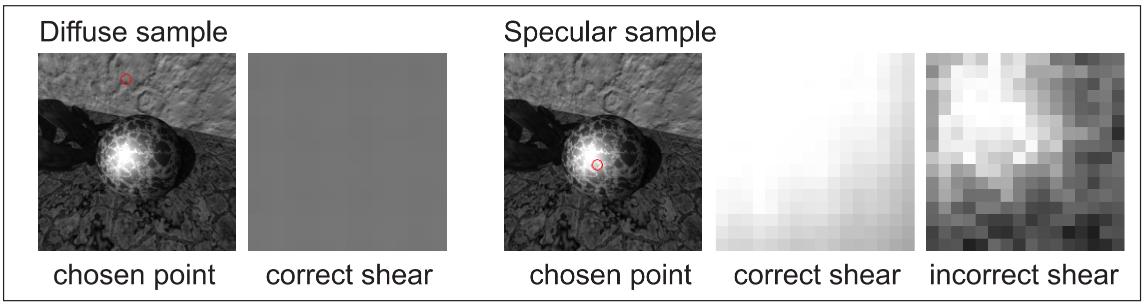

Figure 6.6: *Shear planes for a diffuse and a specular point.*

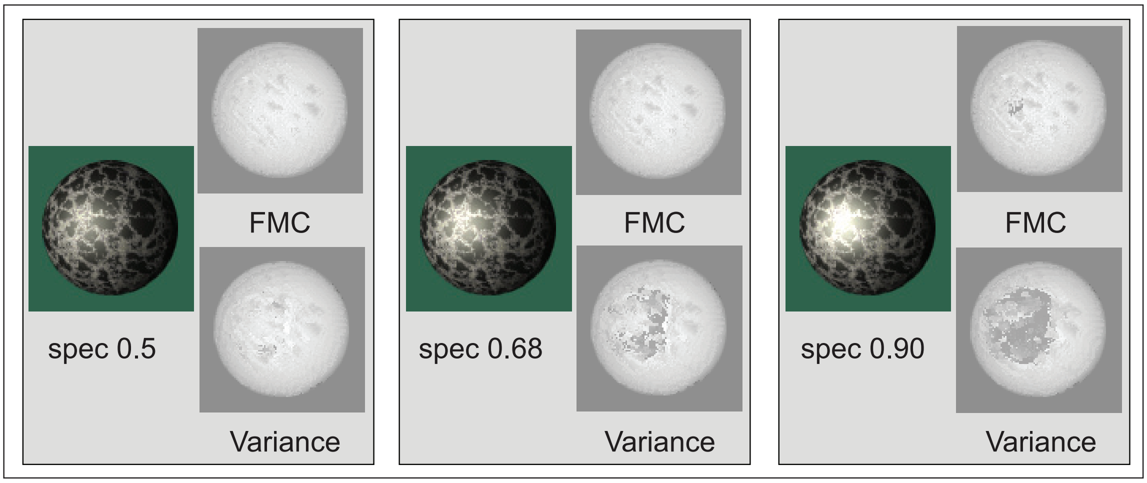

Figure 6.7: *Depth reconstructions based on FMC and variance of the same sphere with varying specular coefficients are compared after the first pass of the 2-pass algorithm of Sect. 6.3.1. The variance leads to holes at specular reflections.*

we will refer to as *Frequency Minimization Criteria (FMC)*:

$$FMC(u,v,s,t,s_{xy}) = \sum_{P \in \mathcal{F}\{\cdot\}} w(P) \cdot |\mathcal{F}\{(p_s - \overline{p_s}) \cdot f_{apod}\}(P)|^2 \tag{6.1}$$

$$\overline{p_s} = \text{mean}(\widetilde{EV_s}'(u',v',s',t')) \tag{6.2}$$

$$z_R = \frac{\Delta u}{\frac{d_{px}}{N} \cdot \tan(\frac{\vartheta}{2}) \cdot 2} \tag{6.3}$$

with f_{apod} being an apodization function and $w(P)$ a frequency dependent weighting function. We use a weighted sum of the power spectrum, penalizing high frequencies and disregarding lower frequencies. This approach gives a finer control of the shear plane analysis than the variance. The shear s_{xy} corresponding to $\min_{s_{xy} \in S}(FMC(u,v,s,t,s_{xy}))$, with S being the set of all possible shears bounded by the closest and farthest object, leads to depth z_R (cf. Eq.(6.3)) of the ray $LF(u,v,s,t)$. d_{px} is the disparity in pixels, N the number of pixels in one view and ϑ the field of view (FoV).

Gray Channel Evaluation Since FMC can only be evaluated on monochrome images, we transform our color images into gray scale images using a technique presented in [GD05]. In order to transfer the gray levels to all the other views of the light field, we create either a big image containing all of the images for small light fields, or evaluate the first principal component of the first image and use it as an axis to which all the colors of all the images are projected to. More elaborate but slower versions of color transformations such as presented in [GOTG05, KRJ05], did not seem to achieve better results according to our requirements.

Occlusion and Specular Highlights Occluded points always lie further away than the occluding points and hence lie on a shear plane with a larger shear factor s_{xy}. As soon as a foreground point is detected, all pixels corresponding to it are ignored for further FMC calculations. In order to avoid high frequencies due to missing pixels, a 3×3 Gauss filter is applied on the shear plane, leading to pixels usable to fill the holes. Furthermore, sharp features at occlusions as well as thin structures are preserved, since our algorithm ignores any kind of neighborhood for correspondence matching. Therefore, occlusions are modeled correctly and included in our reconstruction method.

Since the FMC finds the optimal shear despite specular highlights, we do not have to handle them in any particular way.

Algorithm We evaluate the FMC from the smallest shear, corresponding to the closest point, to the largest shear, corresponding to the farthest point, in order to find the global minimum per sample of the light field. We suggest two methods which focus on speed and accuracy respectively. In a one-pass algorithm, the global minimum is chosen once a local minimum has not been changed after the last w steps and the minimum frequency is below a certain threshold, where w is any number of shear steps (cf. Algo. 1). If the variance over p_s is smaller than a certain threshold, we do not evaluate the FMC, but choose the current s_{xy} as the optimal fit.

Algorithm 1: *Depth reconstruction of the one-pass algorithm.* $prev_w(\cdot)$ *takes the minimum of the last w steps.*

Input : Light Field
Output: Depth Field

% initialize FMC
$FMC(u,v,s,t,s_{xy}) = \infty$ **for** s_{xy}=*smallShear* **to** *largeShear* **do**
  for *all (s,t)* **do**
    eval $FMC(\frac{u_{max}}{2}, \frac{v_{max}}{2}, s, t, s_{xy})$ for p_s **if** $Var(\widetilde{EV_s}'(u',v',s',t')) < Threshold$ **then**
      choose shear remove pixels from p_s
    else if $FMC(u,v,s,t,s_{xy}) > prev_w(FMC(u,v,s,t,s_{xy}))$ **then**
      assign prev. s_{xy} to pixels of p_s remove pixels from p_s
    end
  end
end

A more robust but slower two-pass algorithm removes the points for which a global minimum below a certain threshold has been found after completing the first pass. In a second pass the remaining points are detected by evaluating the minimum FMC for the remaining samples. For a very complex scene the number of passes could be adapted.

Scenes showing reflections of surrounding objects will not be reconstructed properly, since altering colors caused by other objects will lead to high frequencies and are leading to an arbitrary FMC. In the case of big homogenous patches, no unique solution exists due to the lack of information in the light field. In this case we select the first minimal FMC to evaluate the depth.

6.3.2 Hologram Evaluation

The holography pipeline presented in Chapter 4 can be extended in order to handle the evaluation of a light field with a corresponding depth field. Instead of merging the depth maps from all the views to one sampled scene, we set one point source P_{uv} along each ray $LF(u,v,\theta,\phi)$ corresponding to a frequency component of the hologram at a depth corresponding to the depth field in order to minimize speckles in the reconstruction (cf. Sect. 6.5). Each P_{uv} is evaluated over the entire tile T_{uv} of size $\Delta u \times \Delta v$ as depicted in Fig. 6.8a. By evaluating one point source per ray, we implicitly include knowledge of view dependent occlusion and reflection properties of the scene captured by the light field. Efficient per point wavefield evaluation is enhanced by a hardware based implementation of point source evaluation. The contribution of every point source can be added up to obtain the wavefield.

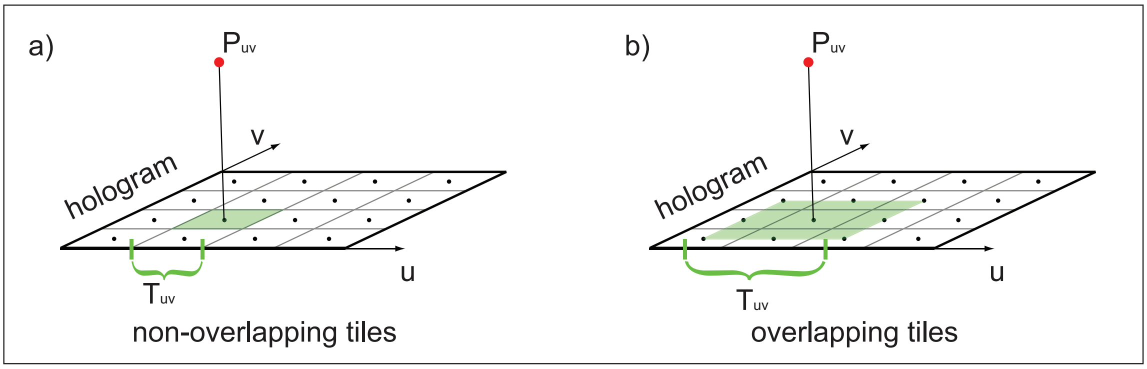

Figure 6.8: *A point source P_{uv} lying on the ray $LF(u,v,s,t)$ will only be evaluated on the tile T_{uv} conserving knowledge of occlusion and reflection captured by the light field.*

Overlapping Tiles Since the captured light field leads to a discrete sampling of the BRDF of points on the scene surface, discontinuities between two neighboring tiles can occur. The maximal discontinuity is dependent on the maximum frequency of the BRDF. In order to avoid ringing artifacts during the inverse mapping $\mathbf{M}^{-1}$, we overlapp and blend the tiles as shown in Fig. 6.8b. The overlap can be freely chosen between 0 to maximally Δu or Δv. The overlapping parts are linearly blended. Note that the blending does not have an influence on the interpolation of intensities between different BRDF samples for novel viewpoints, but solely avoids ringing artifacts when evaluating novel viewpoints.

Choice of Wavelength If the hologram is only used as an intermediate representation and not as an interference pattern for holographic screens, we are able to choose a wave length with more flexibility. On the one hand, the wave length should be as big as possible in order to keep the required sampling fulfilling the Nyquist criteria as low as possible. On the other hand, the wavelength has to be short enough in order to guarantee the required resolution for the back transformation $\mathbf{M}^{-1}$into the light field leading to Sect. 6.5.

Speckle elimination Speckles occur if multiple point sources create the same frequency and the phases are distributed in a way canceling each other out. This is a known physical phenomenon inherent to coherent light modeling. Point sources create the same frequency if they lie in the same frequency band and therefore, in approximately the same direction from the center of the aperture (cf. Sect. 6.5). In order to reduce speckle noise in the final views, we only evaluate every n^{th} point source to create a hologram, leading to a number of n holograms.

The final image resulting from the holographic rendering or inverse mapping is a sum of the images of the n holograms. Using this technique we are able to improve the image quality from a straight forward evaluation Fig. 6.9a with $n = 1$ showing speckle to Fig. 6.9b and Fig. 6.9c being almost speckle free. To avert speckles created from corresponding points, we set identical phases for all of them. By increasing the aperture size over several tiles, speckles become visible at the straight forward transformation. However, since the bigger aperture size is leading to a higher resolution with high-frequency speckle, we low-pass filter and down sample the image to create an almost speckle free smaller image. The resized smaller image still has the same resolution as the corresponding light field view would have had.

6.4 Hologram Operations

A practical application of the transform $\mathbf{M}$ is to create unique input for a holographic screen. Moreover, a holographic representation has various ad-

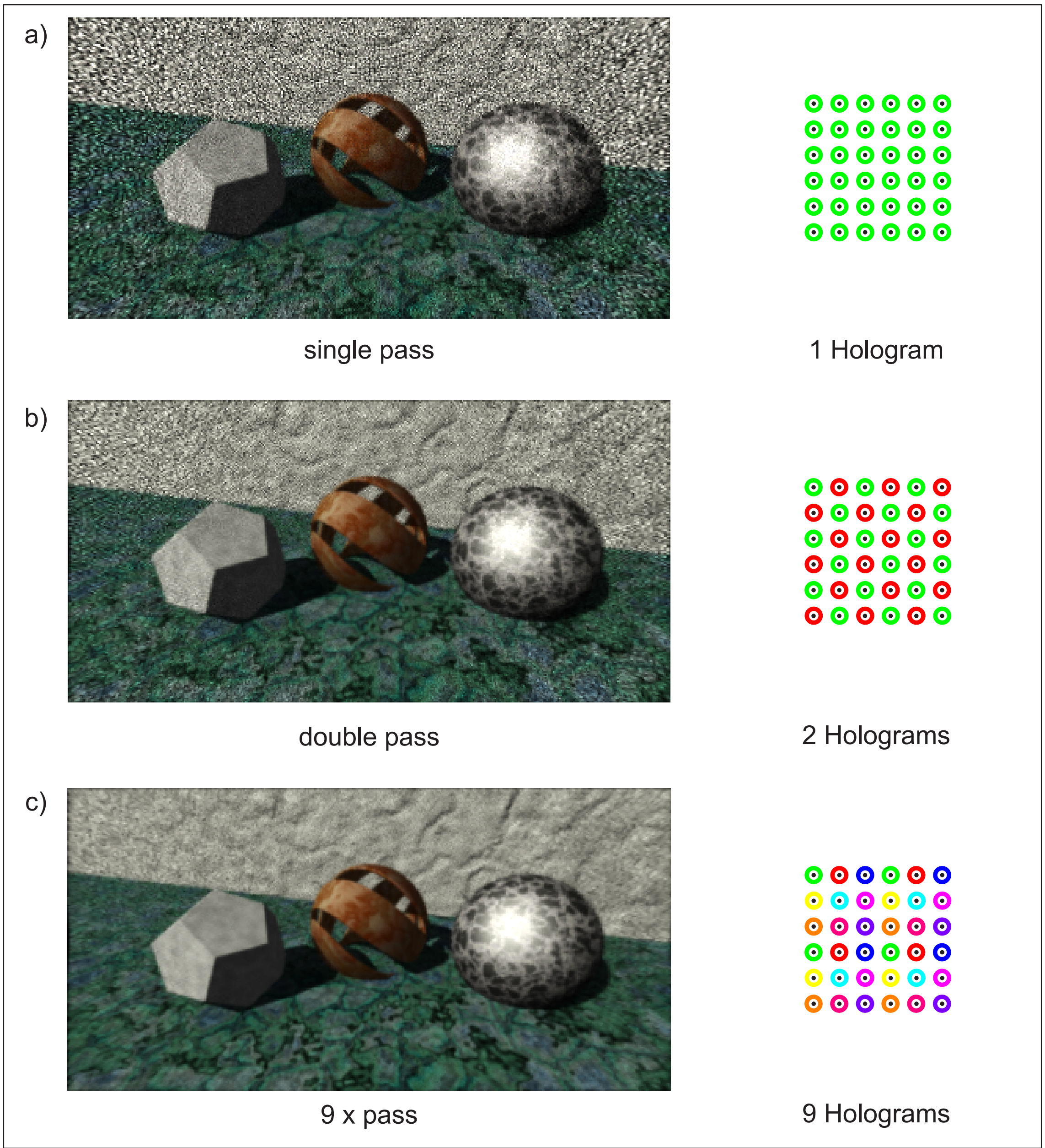

Figure 6.9: *Speckle reduction through selective point source rendering. Each point source surrounded by the same colored ring will be evaluated on the same hologram. Mostly two holograms are sufficient to achieve speckle suppression.*

vantages, such as smooth parallax rendering without ghosting artifacts, robustness regarding data loss, and diffraction simulation.

Smooth parallax rendering The hologram rendering is based on the holographic pipeline presented in Chapter 4. Setting the aperture and viewpoint for a hologram rendering will handle interpolation of intensity information from the light field implicitly. Since depth information of the scene is encoded in the phase of the hologram no ghosting artifacts are visible in novel viewpoints as shown in Fig. 6.10b and Fig. 6.10d. Although light fields do show ghosting as in Fig. 6.10c if no depth is known, they are not prone to speckles as holograms are. By choosing the optimal focal plane, the light field images can be improved considerably as depicted in Fig. 6.10a. Ghosting reduction has been studied intensively in [CCST00, IMG00, SYGM03].

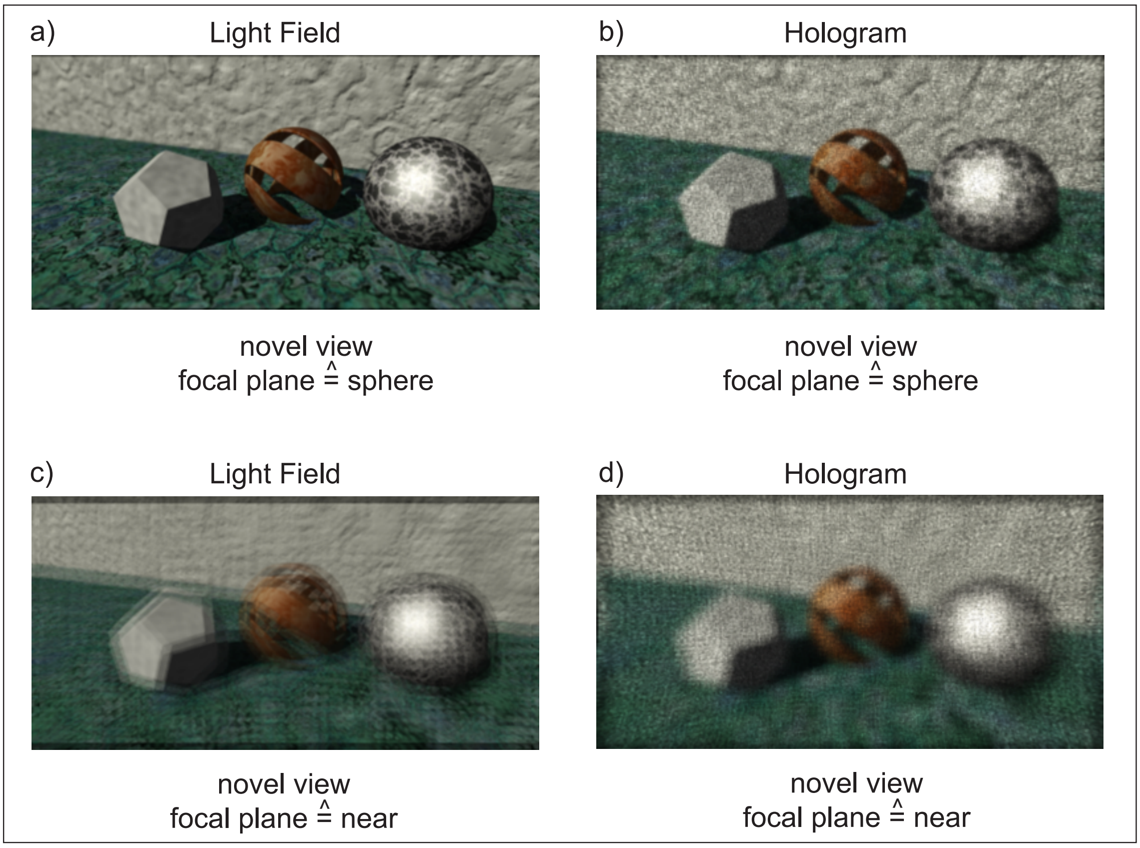

Figure 6.10: *The same novel viewpoint is rendered for the light field and the hologram using various focal plane distances.*

Effects of Loss of Data The hologram stores the information of the scene in the form of a wavefront. Therefore, every point of the scene has an influence on every pixel of the hologram as long as it is not occluded. This means that we can cut out parts of the hologram and still retain information about all the points in the scene as long as at least one part of the evaluation of all the points is still visible. If the aperture is chosen large enough in order to never fully lay over the cut out part, images for all viewpoints can still be reconstructed. Artifacts can occur if the cut out parts are not apodized, since high intensity differences can lead to ringing artifacts.

Compression Light field compression was already addressed in the pioneering work of Levoy and Hanrahan [LH96] and Gortler et al. [GGSC96]. Since then many compression strategies have been presented, which were most often inspired by standard image or video compression techniques. In contrast to this, hologram compression does not lend itself to standard image compression since the reconstruction quality can depend on the entire frequency range. Naughton et al. [NFM+01] showed the limited use of lossless compression for holograms. A better strategy for hologram compression is non linear quantization [SNJ06]. This preserves the spatial detail while still requiring relatively few bits per complex number.

6.5 Inverse Mapping

The inverse mapping $\mathbf{M}^{-1}$cannot be implemented as a straight forward inverse of $\mathbf{M}$, since the point sources are combined to one hologram. A wave based inverse propagation would lead to a lot of problems due to the limited aperture size. Furthermore, the complex valued spatially dependent point spread functions would have to be deconvolved in order to reconstruct each point source independent of one another.

Instead we render images at positions (u,v) on the holographic plane leading to directional components θ and ϕ (cf. Sect. 6.6), which can be interpreted as samples of an angular light field $LF(u,v,\theta,\phi)$. By applying a perspective distortion as described in Sect. 4.4.1 the angular components can be transformed into a two-plane parametrization $LF(u,v,s,t)$. The transformation from the hologram to a light field can be done for any hologram, for which the original wavefront of the scene can be restored.

Desired light field resolution In the first step we have to determine the desired resolution $\Delta\theta$ and $\Delta\phi$ and the spacing Δu and Δv defining the light field. According to Sect. 6.6.2 the centers of the artificial apertures are set on every sample position (u,v) of the hologram. The size of the aperture a has to be chosen such that the minimal angular resolution $\Delta\alpha = \arcsin\left(\frac{\lambda}{d}\right)$ corresponds to $\min(\Delta\theta,\Delta\phi)$. For every aperture, a lens function and apodization function is multiplied with the wavefield before getting the directional components through the Short-Term Fourier Transform (STFT) described in Sect. 6.6.1. The resolution of u is limited by the number of samples on the hologram in this dimension.

Upper boundary of angular resolution The best possible resolution, which can theoretically be achieved for the transformation depends on the maximal depth extension of the visible part of the scene $\Delta\hat{z}_{visScene}$. The depth of field Δz (cf. App. A.2) has to be congruent with $\Delta\hat{z}_{visScene}$ allowing a maximal aperture size a leading to the highest resolution $\Delta\alpha$ for a given wavelength λ. Different techniques to elaborate $\Delta\hat{z}_{visScene}$ are found in numerous papers. We are using a technique proposed in Sect. 4.4.3 for depth reconstruction from the hologram.

6.6 Transformation

The angular spectrum $\mathcal{A}(\cdot)$ as presented in Sect. 3.3.5 is a Fourier Transform $\mathcal{F}(\cdot)$ decomposing the complex valued wavefront $U(u,v)$ into a collection of plane wave components propagating in different directions $\mathbf{k}$ dependent on the spatial frequencies ν_u and ν_v (see Eq.(6.4)). The vector $\mathbf{k}$ is defined as $\mathbf{k} = k \cdot (\alpha, \beta, \gamma)$ with $k = 2 \cdot \pi/\lambda$ being the wave number, λ the wavelength, and (α, β, γ) being the unit vector pointing into the direction of the wave propagation. The components of the vector are called *directional cosines* and are related to the spatial frequencies by $\alpha = \nu_u \cdot \lambda$, $\beta = \nu_v \cdot \lambda$ and $\gamma = \sqrt{1 - \alpha^2 - \beta^2}$.

$$\begin{aligned} \mathcal{A}(\nu_u, \nu_v) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} U(u,v) e^{-2\pi\mathbf{i}(\nu_u u + \nu_v v)} du dv \\ &= \mathcal{F}\{U(u,v)\} \end{aligned} \tag{6.4}$$

Every spatial frequency extends over the entire uv-domain and can, therefore, not be spatially localized. Nevertheless, [Goo68] shows that *local spatial frequencies νl_u* and νl_v can be obtained by a spatially limited Fourier transform as long as the phase $\varphi(u,v)$ does not change too rapidly (see Sect. 6.6.1).

6.6.1 Local Spatial Frequencies

We employ the Short-Term Fourier Transform (STFT) also known as the Sliding-Window Transform, where the wavefront $U(u,v)$ to be Fourier transformed, is multiplied by a window function $h(u,v)$, which is nonzero for a limited area around the origin. The resulting spectrum $S(\cdot)$ is called the *local frequency spectrum* and is defined as follows:

$$S(\nu_u, \nu_v, x, y; h) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{U}(u,v) e^{-2\pi\mathbf{i}(\nu_u u + \nu_v v)} du dv \tag{6.5}$$

$$\hat{U}(u,v) = U(u,v) h(u - x, v - y). \tag{6.6}$$

The multiplication by $h(\cdot)$ suppresses $U(\cdot)$ outside the window and causes a localization. However, since this transformation is governed by the Heisenberg-Gabor inequality, as shown in [Fla99] we cannot get a perfect localization in the spatial domain as well as in the frequency domain.

Considering the analysis of a wavefront, we can say that the better the localization of the directional components, the less directions can be specified. Nevertheless, applying a lens localizes the frequencies for points at specific depths. In the following section we use the principle of STFT, but improve the quality of localization for certain depths.

6.6.2 Aperture

The window $h(\cdot)$ can be regarded as an aperture S_A which blocks the incoming wavefront outside of it. By evaluating the wavefield $U(u,v)$ from a point source P on S_A and transforming it using the STFT we obtain the directional components of the planar waves describing $U(u,v)$. For a point P of finite distance (cf. Fig. 6.11a) $U(u,v)$ leads to several planar waves, and therefore no localization in the frequency domain.

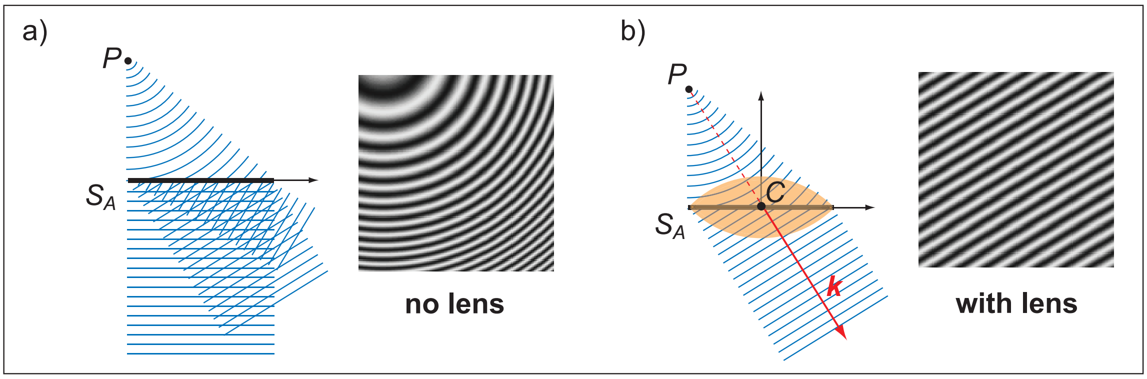

Figure 6.11: *a) shows the frequency distribution over the whole aperture caused by a point source. b) shows the frequency distribution when using a lens.*

By introducing a lens with focal length f as in Eq.(6.7), the incoming wavefront from P can be transformed into a single plane wave as shown in Fig. 6.11b. Introducing a lens does not only have the benefit of creating a single plane wave, but also gives information about the spatial location of the directional wave. Under the assumption of P being perfectly in focus, P lies on the line defined by the center of the aperture C and the directional vector $\mathbf{k}$.

$$L(u,v) = e^{-\mathbf{i}kr}, \text{ with } r = \sqrt{u^2 + v^2 + f^2} \tag{6.7}$$

A lens has theoretically exactly one focal plane, which lies at distance z from the lens and therefore, only planar surfaces parallel to the aperture could be transformed into rays. In practice however, every capturing system has a certain resolution, which determines a Circle of Confusion (CoC) (cf. App. A.1). Taking this CoC into account we can define a Depth of Field (DoF) (cf. App. A.2) in which all the point sources appear in focus and can be transformed into rays going through C. If the whole object is considered to lie in the depth of field of a lens, the frequency distribution of each point will not extend over more than one discrete frequency measure and the wavefield of the scene at the aperture can be transformed into a light field. To get the highest angular resolution (cf. App. A.3) and therefore the biggest aperture we have to achieve a tight fit of the DoF around the object including diffraction.

The lens function, aperture size, wavefront sampling and the wavelength define the resulting CoC and DoF yielding a light field with a specific maximal resolution for s and t as well as a maximal FoV for every position (u, v). The dependence between these characteristics are elaborated in the following appendices App. A.1, App. A.2, App. A.3 and App. A.4.

6.7 Results

All hologram renderings presented in this section are only computed to give an illustration of direct output of holographic content on future generation holographic displays. They do not compete with the light field renderings which are by far more efficient for conventional 2D framebuffer displays. We show the versatility and the power of $\mathbf{M}$ and $\mathbf{M}^{-1}$by applying it to several examples, such as synthetic light fields, real light fields and digitally recorded holograms. The rendered images can be evaluated directly from the holographic representation or through light field rendering.

We implemented a light field renderer using a spatial method capable of simulating different aperture sizes as well as focal length for viewpoints in the uv-plane. A more efficient implementation has been presented in [Ng05] and would have to be used if real-time performance was a requirement. Evaluations of a hologram from a light field and depth field as well as all the renderings from the holograms have been integrated into the pipeline presented in Chapter 4.

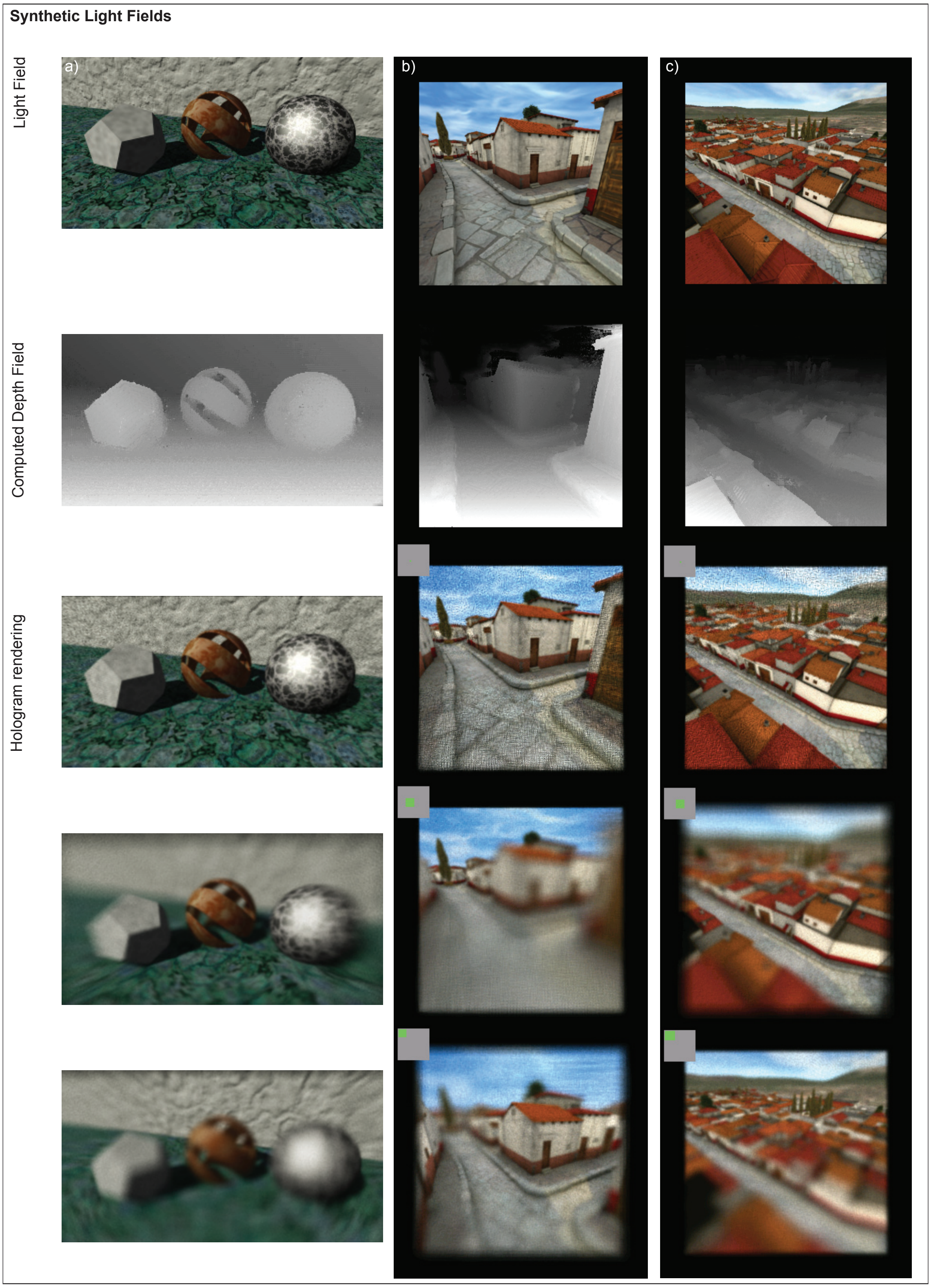

Figure 6.12: *All the image sequences show an original view of the light field input, its corresponding depth field, the rendered transformation into a hologram and two arbitrary views of the hologram with varying aperture and focal length. The grey square symbolizes the hologram, while the green square shows the position and aperture size of the camera. The Pompeii dataset is courtesy of Procedural Inc.*

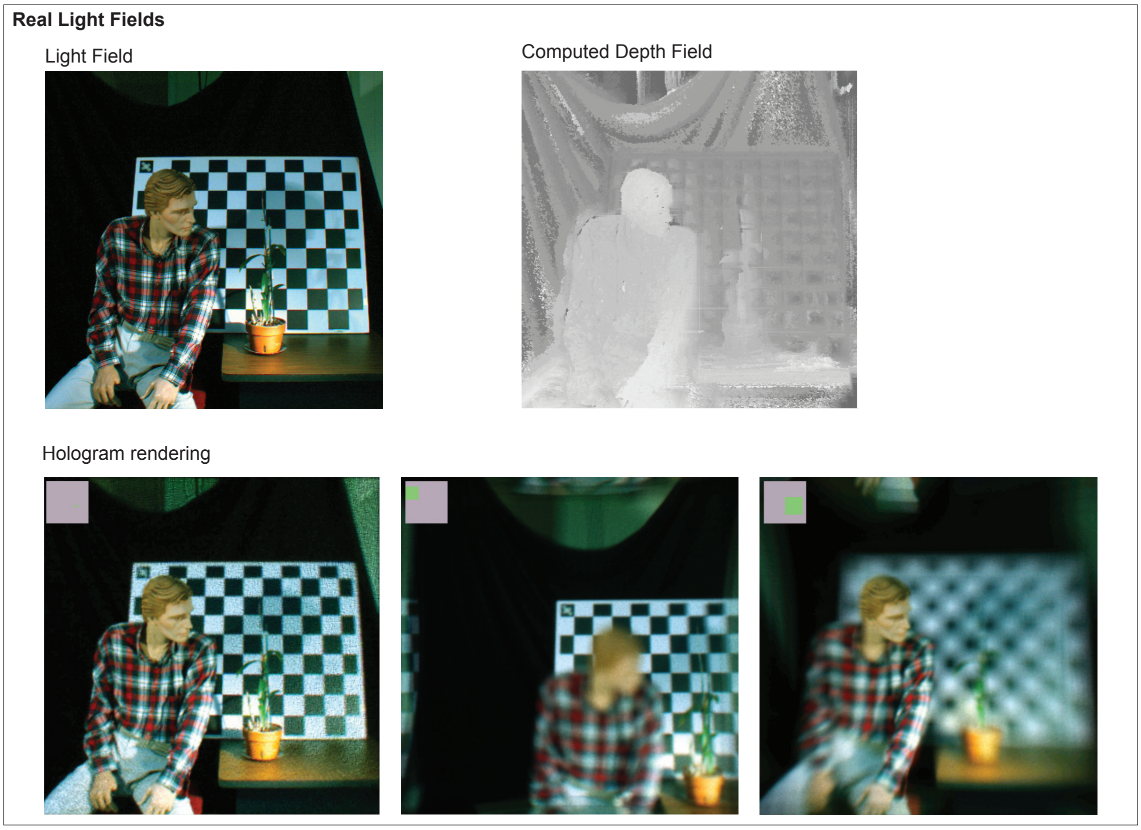

Figure 6.13: *All the image sequences show an original view of the light field input, its corresponding depth field, the rendered transformation into a hologram and two arbitrary views of the hologram with varying aperture and focal length. The grey square symbolizes the hologram, while the green square shows the position and aperture size of the camera. The Mannequin light field dataset was kindly provided by Leonard McMillan.*

6.7.1 Forward Mapping

We compute three **synthetic** scenes shown in Fig. 6.12. The dataset in Fig. 6.12a is a POV-Ray rendered $384 \times 192 \times 16 \times 16$ light field for which our depth map reconstruction algorithm requires 60 minutes on a Pentium 4 with 3.2GHz. The result is a depth map with 163 possible depth values per light field sample, while handling various difficulties such as occlusions, areas of low texture and specular highlights correctly. Fig.6.12b and Fig. 7.11c depict a light field rendered using RenderMan13 of the procedurally generated Pompeii scene presented in [MWH$^+$06]. Both scenes contain a very big depth range, which is bigger than the depth of field of the camera used for the holographic rendering. Regions being slightly out of focus are therefore spread over multiple frequencies leading to some speckle noise. According to Sect. 6.3.2 speckle diminishes for bigger apertures again. The renderings of the hologram are not primarily shown to demonstrate holographic rendering, but to show a possible view, which could be generated on a holographic screen. The human eye would transform the wavefront into an image and therefore determine aperture size and focal length.

Most importantly we transform a **real** light field (cf. Fig. 6.13) into a hologram in order to show that our method can be applied to capture holograms under white light illumination. The depth map reconstruction shows some artifacts since we had no camera calibration and the images suffered from lens distortion. However, the depth map is still precise enough in order not to show any ghosting.

6.7.2 Inverse Mapping

The transformation from the hologram to a light field can be done for any hologram, for which the original wavefront of the scene can be restored. The third row of every sequence of Fig. 6.12 represents a reconstructed light field view by applying $\mathbf{M}^{-1}$. Direct comparisons show some distortions at off-axis rays for cameras with a big FoV.

Furthermore, we transformed a digitally recorded hologram into a light field in order to show the versatility of our framework and transform Fig. 6.14. Our proposed speckle reduction cannot be applied to digitally recorded holograms, so the final renderings are speckle prone.

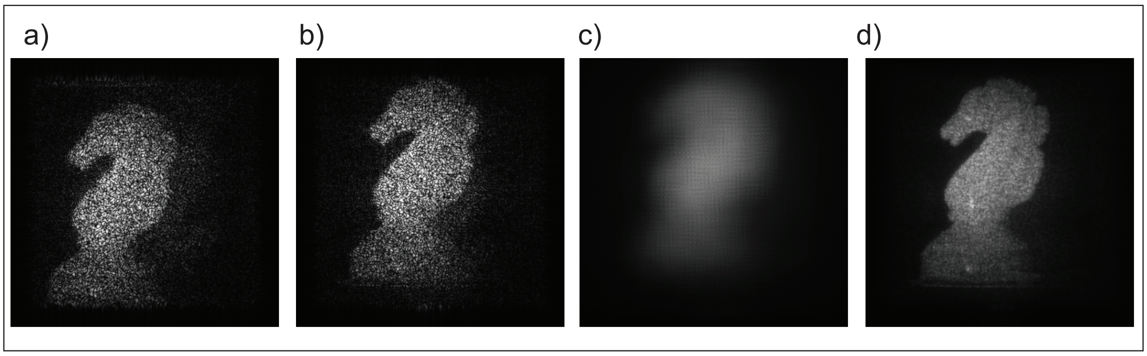

Figure 6.14: *a) and b) are light fields generated from a digitally captured hologram rendered by a small aperture at two positions. c) shows a big aperture with short focal length and d) a big aperture with focal length on object. The Horse dataset is courtesy of Ervin Kolenovic and Jan Mueller from BIAS.*

Limitations A limitation of this framework consists in transforming arbitrary light fields featuring strong reflections and transparent objects, since those regions can fail during depth reconstruction. Furthermore, the lack of visible rays of a point at the border of the light field might not provide enough information for a robust depth estimation in all the views. In our examples at least 20% of the rays have to be visible to reconstruct the depth of a point. The resulting holes are filled through interpolation of surrounding depth values. Inaccurate depth could lead to ghosting for novel viewpoints.

For scenes with a big depth extent speckles can be noticed. Therefore, the bigger the depth extent of the scene the more holograms have to be evaluated for a perfect image. Furthermore, the applied lens model leads to abberations for non-paraxial rays and can therefore lead to speckles.

6.8 Conclusions

We presented a fundamental mapping from light fields to holograms and demonstrated the versatility on multiple examples. For the forward mapping we introduced a novel 3D reconstruction technique based on frequency spectrum analysis capable of evaluating depth despite of occlusions, specular highlights, and low texture information. The created depth field provides the base for a forward transform into a hologram. Most importantly this gives the possibility to capture full parallax holograms under natural illumination, which has not been possible so far. This creates a big potential for future work in this field. Furthermore, the inverse mapping operation allows for digitally captured holograms to be rendered in real-time using the light field representation.

Based on the elaborated mapping operation, holograms can be captured using a light field camera as presented in [Ng05] and rendered on a holographic screen as presented by Qinetiq in [SCS05]. This technique can take advantage of the realism and detail preserving benefits of a real light field while giving the possibility of a 3D output on a holographic screen. Furthermore, the 3D reconstruction technique can be used for ghosting reduction in light field rendering without having to blur any part of the scene. Various lens effects can further be used to create realistic looking renderings for general graphics processing. Finally, digital holograms not requiring optical elements for acquisition, can be rendered in real-time after mapping them into light fields. Therefore, future work can benefit in numerous ways from the fundamental mapping by taking advantage of either representation, depending on the needs.

Throughout this chapter, we assumed sufficient lighting of the scene, in order to capture the light field in one step. As in conventional photography, light field capturing suffers also from from low light conditions. If the exposure time is kept short in order to avoid motion blur, the light field or photograph is featuring noise. If the exposure time is long enough to reduce noise, the light field or photograph is often showing motion blur due to camera shake or object movement. In Chapter 7, we present a novel spatially adaptive flash. This flash is able to homogeneously light a scene, such that a light field could even be captured under dark lighting conditions. Since the light field is often captured as a set of photographs, we describe our setup from the perspective of conventional photography.

CHAPTER 7

Spatially Adaptive Flash

In the previous chapter, we have shown that a hologram can be captured under white light conditions by transforming a captured light field into a hologram. A light field can be acquired by capturing a bunch of photographs, which are then being parameterized to the required representation. As in photography, one of the requirements of such an acquisition is - to have enough light. The presented method did, however, not take this requirement into account. In common photography, the lighting conditions are controlled artificially by either using studio lighting in a fixed setup, or by using a flash unit in more mobile setups. These ways of enhancing the illumination artificially can also be used for light field acquisition.

Although, flash lighting seems to be a plausible solution for our problem, common flash units also show deficiencies. In this chapter, we present a novel spatially adaptive flash unit, which alleviates problems of reduced illumination of distant scene parts, avoids overexposure due to spatially dependent light attenuation, and projects spatially dependent colored light for mood lighting or textured projection. Since we need to acquire photographs to acquire a light field, and therefore, run into the same problems, we focus on the usage of a spatially adaptive flash for common photography in this chapter. The results can, however, be applied for light field acquisition as well.

7.1 Flash in Common Photography

Digital cameras have become inexpensive, robust, and small. As a result, most consumers own one, carry it everywhere they go, and expect it to take good pictures under a variety of shooting conditions. To accommodate dark scenes, most digital cameras include electronic flash - either integrated with the camera or in a detachable unit. More expensive units offer control over brightness, duration, and angular spread. Regardless of how these parameters are chosen, flash illumination always falls off sharply with distance from the camera, making most flash pictures look unnatural. It is tempting to address this problem by capturing a flash-noflash image pair, estimating depth as the ratio of the flash and noflash images, then inversely scaling pixel intensities by this depth. However, separately scaling the intensities of individual pixels does not account for possible interreflections among the corresponding features in the scene. This can lead to physically invalid scenes, and hence to images that appear unnatural.

7.2 Spatially Adaptive Photographic Flash

We solve this problem by changing the physical illumination of the scene on a pixel-by-pixel basis. Our system employs an infrared time-of-flight rangefinder to determine the distance to each object and a modified video projector to modulate the illumination. While this approach is limited by the power of our light source, the illumination it produces is physically plausible, as are the interreflections it triggers in the scene, so our images look natural.

Using such a system it is possible to simulate a variety of virtual light sources. Unless the light source is a dramatic player in the scene, most photographers prefer sources that are diffuse, distant from the primary subject, and angled with respect to the direction of view. For interior shots, examples are ceiling lights or the sun passing through a side window. Since our spatially adaptive flash unit sits atop the camera, we cannot simulate lights that arrive from the side or above the subject. However, by modulating the illumination of each object as a function of its distance from the camera, we can simulate a light source located some distance behind the photographer. Illumination from such a source falls off quadratically with distance, but at a slower rate than from a flash unit mounted on the camera. In order to avoid unnatural shadows above the subject, we mount our physical light source above the camera, as for conventional flash.

Our method requires only one snapshot by the main camera, preceded by a single capture using the rangefinder. For scenes in which flash provides

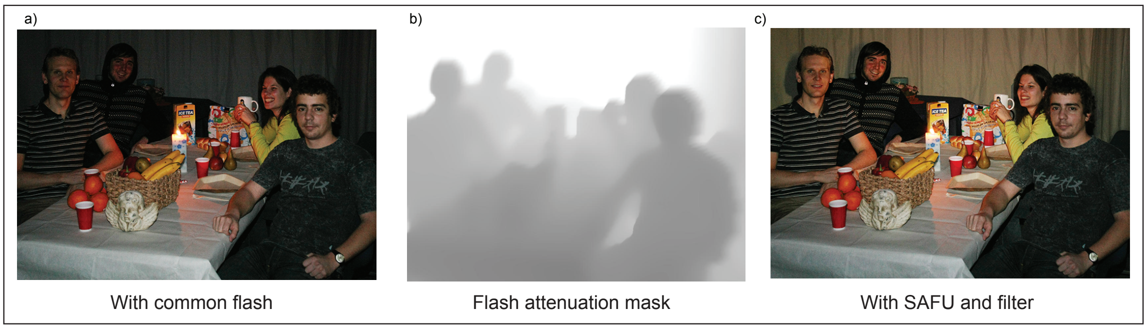

Figure 7.1: *a) Image captured with a conventional uniform flash, whose intensity has been attenuated to avoid saturating the faces of the frontmost people. Note that the people in the back are too dark. b) A flash adjustment mask, computed from depth values captured by a time-of-flight rangefinder. c) Using this mask, we can increase the flash intensity for distant parts of the scene, up to the maximum available flash power. This brightens up these regions, which improves signal-to-noise and looks more natural. In this case, we have simulated a virtual light source located two meters behind the photographer.*

the bulk of the illumination, a second problem is that highly reflective objects may saturate the sensor, especially if they lie close to the camera. By capturing an additional pilot image to estimate object reflectances, we can control the video projector to reduce the illumination of these objects during the final image capture. One can treat this additional modulation as a sort of "physical tone mapping" - reducing the dynamic range of the scene by adjusting its illumination locally. In this paper we describe the hardware and algorithms of our system, and we demonstrate its use in situations where flash typically performs poorly, such as candid shots of human faces.

7.3 System Overview

Our proposal can be partitioned into two variants: single-shot and multi-shot. In the single-shot variant, one depth image is acquired, followed by a color image that captures the subject under spatially varying illumination. Since the depth image can be acquired during the focusing of the color camera, we consider this a single-shot method. Fig. 7.1 was captured using this method. In the multi-shot variant, we capture one depth image and one or two color images. The color images serve as pilot images, from which we estimate the brightness of each feature in the scene. We use this information to further refine our spatially varying illumination during a final color image acquisition. We can also use this information to refine our depth image. Fig.7.11h was captured using this method.

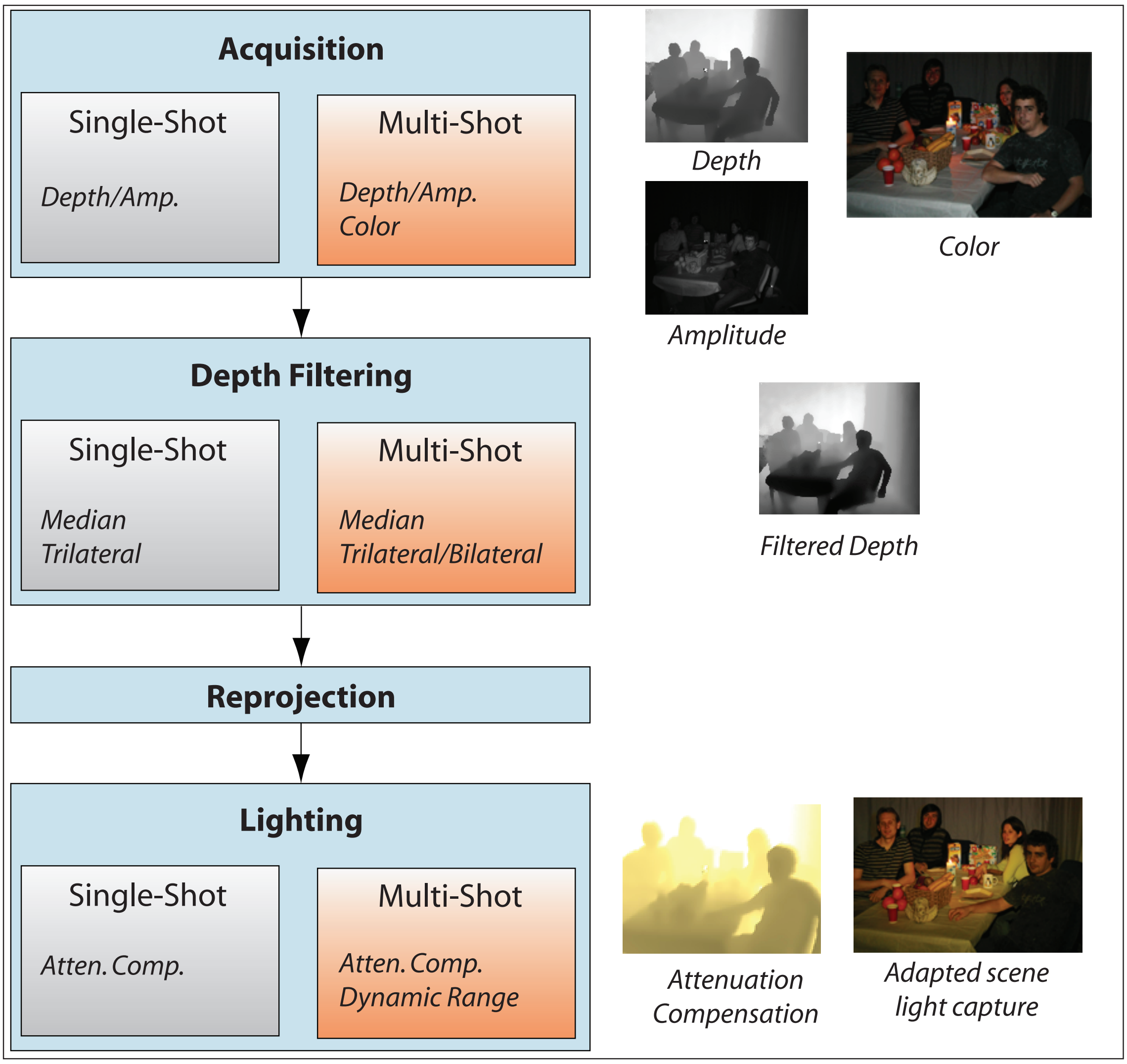

Figure 7.2: *The system can be split into two approaches, namely a single-shot and a multi-shot method. They consist of an acquisition, depth filtering, reprojection and lighting step.*

Fig.7.2 shows an overview of our system. The single-shot and multi-shot approaches can be split into four steps, acquisition, depth filtering, reprojection, and lighting. During the filter step, a median filter is applied to remove outliers in depths returned by the ToF-camera, and a novel joint adaptive trilateral filter is applied to reduce noise. The filtered depth data is then back-projected, triangulated, and rendered into the projector view using graphics hardware. This step performs warping and resolves visibility in real-time. To cope with inaccuracies in calibration or depth values, we apply a conservative depth discontinuity diffusion filter. This filter gives us depth values that are correct or slightly closer than the actual depth, thereby diminishing the visibility of possible misalignments. Next an attenuation mask is computed. This computation takes as input the filtered depths, the virtual light position, and a user defined artistic filter. This attenuation mask is loaded onto the LCD of the Spatially Adaptive Flash Unit (SAFU), and a final image is captured by triggering the flash and camera to fire.

In the multi-shot approach, depth filtering is carried out in two phases rather than one. In the first phase, the depth filtering of the single-shot approach is applied to the depth data, after which it is warped into the color camera using the techniques just described. In the second phase, we capture a color image, then apply a slightly modified joint bilateral filter [PSA$^+$04] to refine the depth edges according to the color data. The resulting high resolution depth information is then reprojected into the projector view. Finally, the attenuation mask is computed as in the single-shot method, but based on a high resolution mesh instead of low-resolution. Additionally, reflectance information from a flash-noflash pair of color images is used to refine the attenuation mask to reduce illumination of highly reflective objects.

The hardware setup and the required calibration are described in the remainder of this section. In the two following sections we describe, firstly, the entire system using a single-shot approach (cf. Sect. 7.4), and secondly, the additional benefits and applications using the multi-shot approach (cf. Sect. 7.5).

7.3.1 Hardware Setup

The setup consists of a Spatially Adaptive Flash Unit (SAFU) as well as a color camera. In our standard setup we position the SAFU, such that the location roughly corresponds to a conventional flash unit mounted on the hot shoe of the color camera (cf. Fig. 7.3). This allocation of the different components has been chosen to get as close as possible to a candid shot setting. However, our setup is not limited to this setting, since the SAFU could also be triggered wirelessly from the camera. However, if the SAFU is not attached rigidly to the camera, the multi-shot approach would require recalibration each time either is moved.

SAFU

The SAFU consists of a lamp, a LCD panel, a lens and a depth camera based on time-of-flight (ToF) technology. The conceptual setup is shown in Fig. 7.3a. In our current setup shown in Fig. 7.3b, we removed the light bulb from a projector (SANYO PLC-XW50) and replaced it by a Canon Speedlite 580EX II. In order to increase the intensity from the SAFU, we built a mirrored tube. This maximizes the light projected into the light path. The rest of the light path, including three LCD’s, a lens, and the depth camera, are not modified. The alignment of the optical axis of the depth camera and the optical axis of the lens of the light are chosen to be as close as possible, in order to minimize the disparity between their views of the seen.

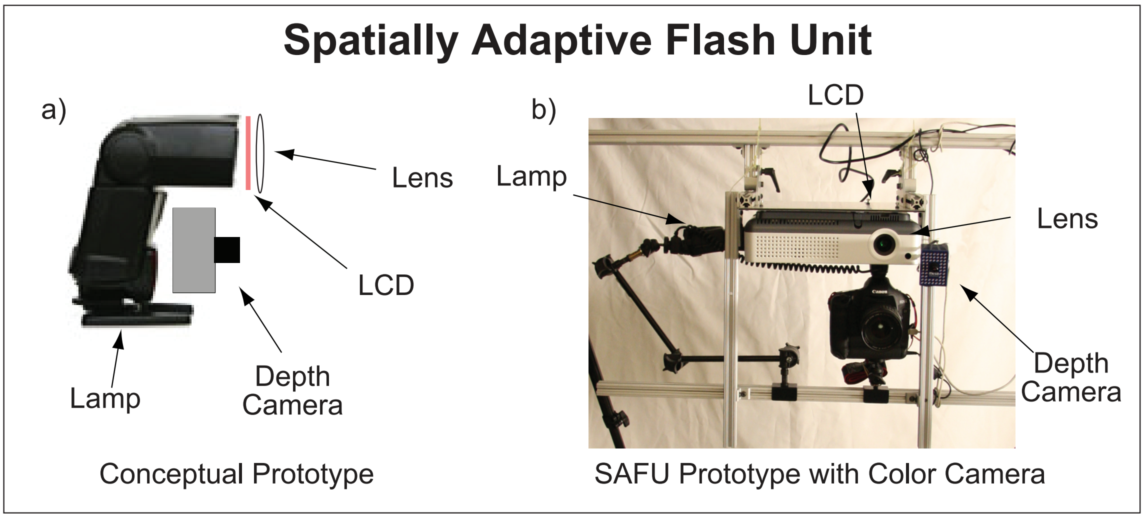

Figure 7.3: *a) shows the proposed SAFU with lamp, LCD, lens and ToF camera. b) Our prototype is implemented using a modified projector and a depth camera.*

Time of Flight Camera

We employ a Swissranger ToF-camera from Mesa-Imaging with a resolution of 176×144 pixels, providing the depth and amplitude at a modulation frequency of 20MHz. This allows for a depth range of 7.5m without ambiguities. Our ToF camera has a built in illumination source, which operates in the infrared spectrum of 850 nm. The accuracy is highly dependent on the reflected amplitude as well as on the background illumination. According to [BOL$^+$] the best depth accuracy lies in the mm range. Our captured depth maps partially suffer from substantial noise, which has to be filtered as described in Sect. 7.4.1.

Digital SLR Camera

Our color camera is a Canon EOS 1D Mark III camera with a Canon EF 16-35mm f/2.8L USM lens. The camera is connected by USB to the PC. Using the Canon SDK we are able to control settings of the camera.

7.3.2 Calibration

The distortion, as well as the extrinsic and intrinsic matrices of the depth and the color camera, are computed from six images of the printed checkerboard pattern. This method can be used for both types of cameras, since the checkers are also visible in the infrared spectrum of the depth camera.

The projector is calibrated relative to the color camera based on the method presented in [SM02]. This method requires four images of a plane in space. To increase the accuracy of the calibration, the two planes have to be as far apart as possible. The first two images are taken from the plane being as close as possible to the camera, while the third and fourth image are taken while the plane is moved as far away as possible from the camera. The image pairs of the plane at one position, always consist of an image of a printed checkerboard pattern placed on the plane and a projected checkerboard pattern on the same plane. The problem of this method is twofold. On the one hand, this acquisition is tedious, since the printed checkerboard pattern has to be covered with a white sheet to acquire the image with the projected pattern. On the other hand, the method could be inaccurate, since the plane could move between the first and the second image. To increase the accuracy of the projector calibration, we use novel checkerboard patterns of different colors, allowing for only one image per plane position. We use a blue colored and printed checkerboard pattern, while projecting a red checkerboard pattern on top of it (cf Fig. 7.4a). The calibration pattern can be extracted from the red channel (cf Fig. 7.4b) for the 'first' image, while it is extracted from the blue channel (cf Fig. 7.4c)for the 'second' image. This modification has the advantage, that the plane of reference will be exactly the same for both images, leading to a better accuracy, and more simplicity than the previous approach.

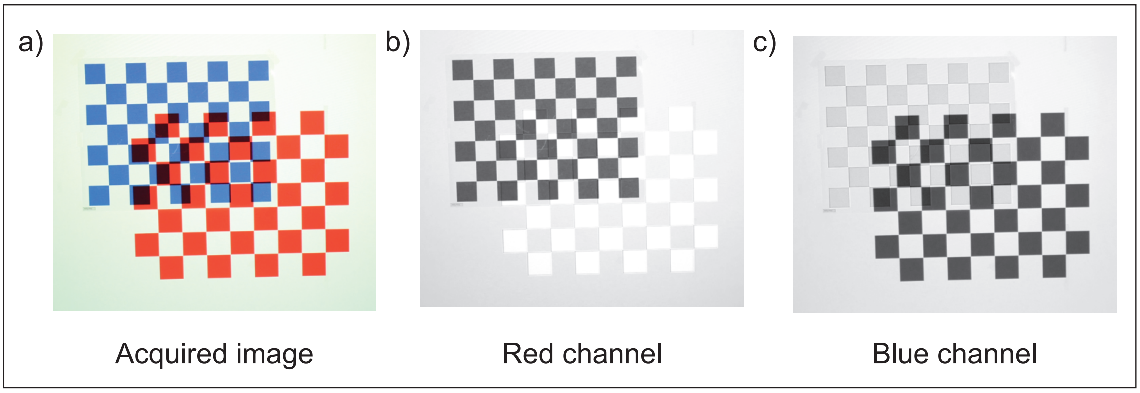

Figure 7.4: *a)*

It is important to note that this calibration has to be done only once, assuming the SAFU is mounted rigidly to the camera.

7.4 Single-Shot Flash Adjustment

Our single-shot method permits compensation for the unnatural falloff of a conventional flash by virtually repositioning the light source further back -

behind the photographer. To simulate the falloff of such a light source, we require per-pixel depths. Our time-of-flight camera captures these depths at the same time the color camera is being aimed and focused. We call this method, where flash intensity is modulated based solely on depth, as ”single-shot flash adjustment” (SSFA). SSFA has the advantage of minimizing motion artifacts relative to methods that require multiple color images. However, variations in object reflectance cannot be taken into account by this method.

Assuming an already calibrated setup, SSFA can be split into following steps:

1. Acquisition of depth and amplitude (infrared)
2. Depth filtering
3. Reprojection from the depth-camera to the flash unit
4. Per pixel, depth-dependent light adjustment
5. Color camera capture

The acquisition using the depth-camera is straight forward, and therefore, not elaborated any further.

7.4.1 Depth Filtering

The captured depth data has a direct influence on the projected intensity per pixel. Noise in the depth data is, therefore, also visible as noise in the evaluated illumination. Since the human visual system is very sensitive to relative intensity variations, noise in the flash intensity becomes very easily visible. Although increasing the exposure time of the depth camera improves the SNR, it also increases motion blur artifacts. Therefore, we reduce the noise by applying different filters on the depth data allowing to keep a short exposure time.

Bilateral filtering [TM98] has shown to perform well for filtering noise while preserving discontinuities in the signal. However, in order to remove the noise, the discontinuity of the signal always has to be higher than the noise level, which is not always the case when dealing with depth data from ToF-cameras. Eisemann and Durand [ED04] and Petschnigg et al. [PSA+04] introduced the cross bilateral filter (a.k.a. joint bilateral filter), as a variant of the classical bilateral filter. They use a low noise image, meaning a high confidence in the high frequency data, as input to the range weight. This avoids smoothing across discontinuities of the signal. We extend this approach to a Joint Adaptive Trilateral Filter, where the noise model of the depth values of the ToF-Camera is taken into account for the confidence of the range weights. According to [BOL+], the standard deviation σ_D of the depth data

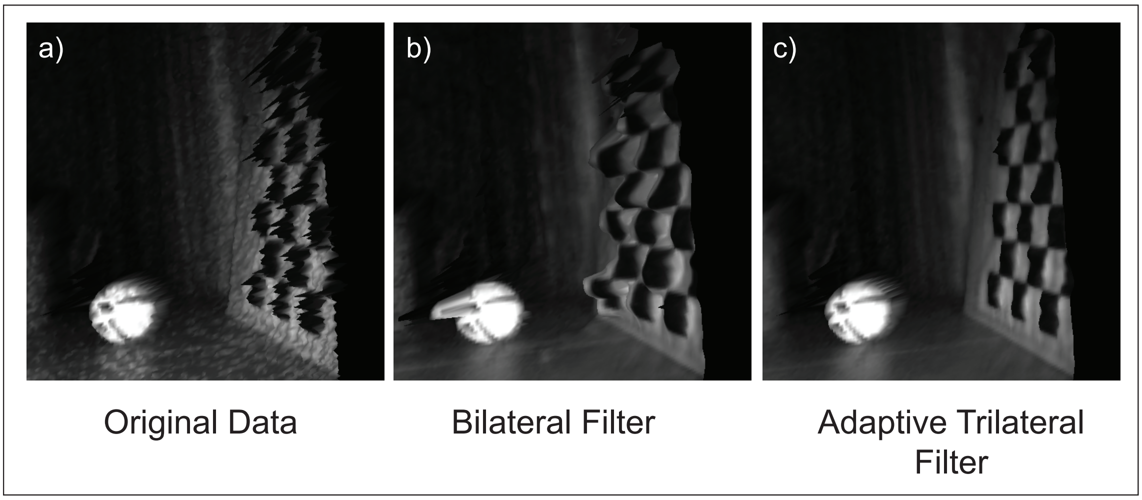

Figure 7.5: *a) Original depth data. b) Filtered data after applying twice a bilateral filter of size* 5×5*. c) Filtered data after applying twice a joint adaptive trilateral filter of size* 5×5*.*

can be described as

$$\sigma_D = \frac{c}{4\pi \cdot f_{mod}\sqrt{2}} \cdot \frac{\sqrt{B}}{c_{demod} \cdot A_{sig}}, \text{ with} \tag{7.1}$$

$$c_{demod} = \frac{A}{B}, \text{ and} \tag{7.2}$$

$$B = A_{sig} + BG, \tag{7.3}$$

where c is the speed of light, f_{mod} is the modulation frequency, BG is the background illumination, A_{sig} is the mean number of electrons generated by the signal, and A is the measured amplitude in electrons. Additional noise sources, such as thermal noise or dark current electrons are minimized by the camera and are neglected by our filter. Since c and f_{mod} are constant and c_{demod} can be approximated by 0.5 according to [BOL+], the standard deviation of the depth distribution is $\sigma_D \approx \frac{k}{\sqrt{A_{sig}}}$, with $k = \frac{c}{2\pi \cdot f_{mod}\sqrt{2}}$. This relationship shows that σ_D, and therefore, the noise in the depth image, mainly depends on the amplitude.

Joint Adaptive Trilateral Filter We introduce a novel joint adaptive trilateral filter (JATF), which takes the amplitude dependent standard deviation σ_D of the depth values into account. The JATF consists of a spatial proximity component $g(\cdot)$, a depth dependent component $f(\cdot)$ in the range domain and an amplitude dependent component $h(\cdot)$ in the range domain. The filtered depth value $d_{\mathbf{m}}$ at position $\mathbf{m}$ can be described as:

$$d_{\mathbf{m}} = \frac{1}{k_c} \sum_{\mathbf{i} \in \Omega} d_{\mathbf{i}} \cdot \underbrace{g(\|\mathbf{m} - \mathbf{i}\|)}_{spatial} \cdot \underbrace{f(d_{\mathbf{m}} - d_{\mathbf{i}}) \cdot h(a_{\mathbf{m}} - a_{\mathbf{i}})}_{range} \tag{7.4}$$

$$k_c = \sum_{\mathbf{i} \in \Omega} g(\|\mathbf{m} - \mathbf{i}\|) \cdot f(d_{\mathbf{m}} - d_{\mathbf{i}}) \cdot h(a_{\mathbf{m}} - a_{\mathbf{i}}) \tag{7.5}$$

where $a_{\mathbf{m}}$ and $a_{\mathbf{i}}$ are amplitude values at position $\mathbf{m}$ and $\mathbf{i}$ respectively. The three functions $g(\cdot)$,$f(\cdot)$ and $h(\cdot)$ are chosen to be Gaussian kernels, each with a kernel support of Ω, and the corresponding standard deviations σ_p, σ_d and σ_a respectively.

Intuitively, the filter is smoothing the depth values, while limiting the influence at depth or amplitude discontinuities. However, since the noise is not uniformly distributed, we adapt σ_d and σ_a, and therefore, the influence of $f(\cdot)$ and $h(\cdot)$. Having a good approximation of the uncertainty measure of the depth values Eq.(7.1), we can define two amplitude dependent functions $\sigma_d(\cdot)$ and $\sigma_a(\cdot)$ as follows:

$$\sigma_d(a_{\mathbf{m}}) = \sigma_{d_{init}} \cdot (\frac{\tau_a}{a_{\mathbf{m}}} + 1) \tag{7.6}$$

$$\sigma_a(a_{\mathbf{m}}) = \frac{\sigma_{a_{init}} - \sigma_{a_{min}}}{1 + e^{-a_{\mathbf{m}} + \tau_a}} + \sigma_{a_{min}}. \tag{7.7}$$

The initial standard deviation $\sigma_{d_{init}}$ is set to a low value, while the $\sigma_{a_{init}}$ is set to a high value, dropping to $\sigma_{a_{min}}$ at $a_{\mathbf{m}} = \tau_a$.

Median Filter For very low amplitudes, the provided depth is very often an outlier. Although the JATF would be able to remove such outliers, it would require several iterations to do so. Therefore, we apply a median filter, which removes the outliers in one step. To avoid unnecessarily blurring regions without outliers, we apply the median filter only to pixels with an amplitude below a threshold τ_m. A threshold of $\tau_m = 40$ has shown to be a good value.

7.4.2 Mesh-Based Reprojection

To attenuate the light according to the depth values, they have to be transformed from the depth camera view to the projector view. In order to solve this reprojection step efficiently, we form a trivial triangular mesh using all the back projected pixels from the depth camera and render them into the projector view. This provides a depth value per projector pixel and solves the visibility issue using the graphics card.

Artifacts can be introduced through resolution differences or remaining inaccuracies in the depth values. Since we do not have more information to

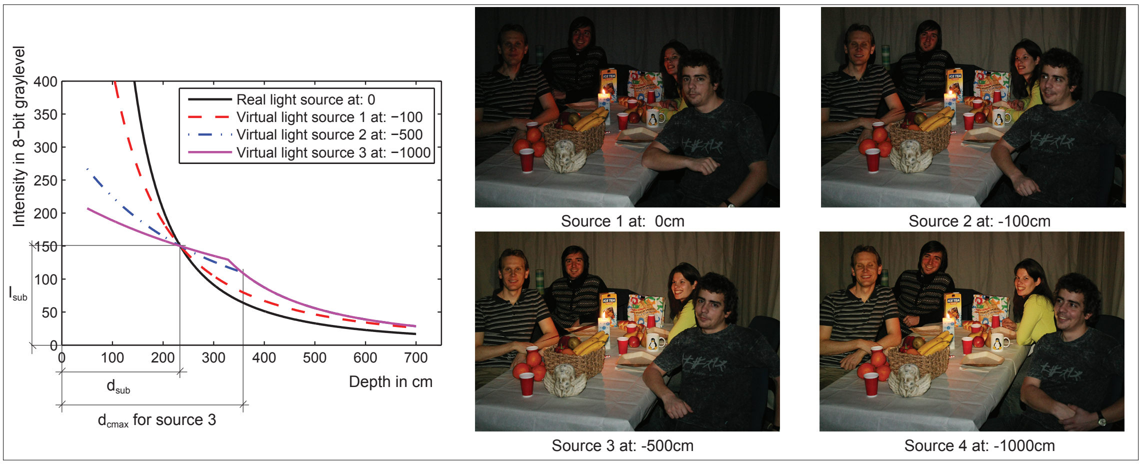

Figure 7.6: *These four images were illuminated by our system such that they appear to have been lit by light sources positioned at four different depths behind the photographer, as shown in the plots at left. We observe that we cannot simulate source 3 beyond 380cm from the camera, because our flash unit offers insufficient power to stay on the blue curve. In practice, the transition point to the usual quadratic falloff (at d_{cmax}) is smooth, and we will not see artifacts.*

improve the accuracy of the data during the single-shot acquisition, we improve the quality of the result by moving the error to regions where it is least visible. Depth values that are too large produce visible artifacts, appearing as unexpectedly bright illumination. These artifacts are especially noticeable on objects close to the camera. Similarly, depth values that are too small produce unexpectedly dim illumination. Most of these artifacts appear at depth discontinuities. This means, that dim illumination due to too small depth values, will be adjacent to shadows produced by the flash and thus be less noticeable than the ones caused by too large depth values. Therefore, we err on the side of computing depth values that are either correct or too shallow. We refer to this as the "conservative approach".

Shallow Edge Diffusion One possibility to improve the depth data in a sense as to comply with the "conservative approach", is the shallow edge diffusion. We start extracting all the big depth discontinuities using a Canny edge detector. Then, we simplify the obtained edges to a contour and move every vertex of it in the direction of its local gradient towards the region with the bigger depth values. The amount of translation corresponds to the multiplication of a globally defined width and the local gradient of the vertex before translation. By connecting the first and last vertex of the original and the moved contour, we create a set **P** of points lying inside the polygon. Finally, we apply a modified bilateral filter (Eq.(7.8)) to all the depth values d_m laying inside of **P**, such that the depth edge is filtered and being moved towards the region of bigger depth. The value k_c is normalizing the weights of the depth values d_i. $g(\cdot)$ and $h_p(\cdot)$ are two Gaussian kernels defined on the support $\Omega_\mathbf{P}$. $d_{\mathbf{i}_{min}}$ corresponds to the shallowest depth value of the kernel support.

$$d_\mathbf{m} = \frac{1}{k_c} \sum_{\mathbf{i} \in \Omega_\mathbf{P}} d_\mathbf{i} \cdot g(\|\mathbf{m} - \mathbf{i}\|) \cdot h_p(\|d_{\mathbf{i}_{min}} - d_\mathbf{i}\|) \tag{7.8}$$

7.4.3 Depth Dependent Flash Adjustment

The intensity I_0 of an illumination source at distance 1 decreases with the distance d to an intensity of $I_d = \frac{I_0}{d^2}$. This intensity falloff becomes visible in flash-photography and leads to unnaturally lit scenes. A light source positioned further behind the photographer, would result in a more homogeneously lit scene, which would appear more natural. Based on the depth values obtained from the previous step, we can evaluate a spatially dependent attenuation function, and simulate the more desirable light position as a virtual light source. The photographer can choose two interdependent parameters, namely the position of the virtual light source along the optical axis of the SAFU, and the minimal illumination for a specific subject of the scene. These preferences determine the depth compensation function leading to an attenuation per pixel.

The attenuation can in fact be interpreted as a filter by which the flash light is being modulated. In Sect. 7.4.3 we present two additional artistic filters, giving the photographer more creative freedom.

User Preferences

The parameters, which can be set by the user, are chosen based on the possible needs of the photographer. On the one hand, the photographer wants to pick the position of the light, and on the other hand, he wants to choose how much light is added to the subject of interest. The former can be set by providing the depth offset d_{off} relative to the photographers position, with the positive axis lying in front of the photographer. The latter is defined by an intensity $I_{sub} \in \{0,1,2,..,255\}$ of the flash at the depth corresponding to the focal length d_{sub} of the camera. We think, that this is a natural way of parameterizing the virtual light source, since the photographer wants to primarily have control over how much light is added to the subject in focus. Especially for hand-held shots, the exposure time should not drop below a certain threshold to avoid motion blur due to camera shake. Of course, our method is not limited to this parametrization, but it is the one we thought to be most useful for candid shot settings.

Depth Compensation Function

Given the final depth values from Sect. 7.4.2 and the user parameters from Sect. 7.4.3, we can define the depth compensation function $f_c(\cdot)$ as

$$f_c(d_i) = \frac{I_{sub} \cdot (d_{sub} - d_{off})^2 \cdot d_{\mathbf{i}}^2}{(d_{\mathbf{i}} - d_{off})^2 \cdot d_{sub}^2} \tag{7.9}$$

$$f_{att}(d_i) = \begin{cases} I_{max} - f_c(d_i) & \text{if } f_c(d_i) < I_{max} \\ 0 & \text{if } f_c(d_i) \geq I_{max} \end{cases} \tag{7.10}$$

with $d_{\mathbf{i}}$ being the depth value at the pixel with index $\mathbf{i}$. $f_c(d_{\mathbf{i}})$ corresponds to the intensity that a pixel at index $\mathbf{i}$ requires to compensate for the position of the virtual light source. $f_{att}(\cdot)$ is the attenuation function dimming the maximal lamp power of the SAFU to result in a projection intensity of $f_c(\cdot)$. Since the compensation function is limited, we can define a maximal compensation depth d_{cmax} as

$$d_{cmax} = \frac{-I_{max}d_{sub} - \sqrt{I_{max}I_{sub}(d_{sub} - d_{off})^2}d_{off}d_{sub}}{I_{sub}d_{sub}^2 - 2I_{sub}d_{sub}d_{off} + I_{sub}d_{off}^2 - I_{max}d_{sub}^2}, \tag{7.11}$$

where $I_{max} = f_c(d_{cmax})$. If d_{cmax} is smaller than any depth in the scene, the system could notify the photographer about a possible underexposure and either recommend a lower I_{sub} or a smaller offset d_{off}.

Artistic Filters

In addition to compensating for depth, one can also adapt the color balance to keep the lighting mood of the scene, such as in the candle lit scene of Fig. 7.7a. Furthermore, color images can be applied as artistic filters as well Fig. 7.7d.

7.5 Multi-Shot Flash Adjustment

The SSFA is limited by a low resolution depth image, as well as a low resolution infrared amplitude image. Therefore, scenes featuring very fine detail might show artifacts due to wrong depth values. Furthermore, the lack of color information does not allow to correct for varying reflectance properties and overexposed regions automatically. To improve on these two limitations, we implemented a "multi-shot flash adjustment" MSFA. Since a color image has to be captured before taking the final shot with the corrected flash light, we consider this as a multi-shot approach. Assuming an already calibrated setup as for the SSFA, the MSFA can be split into following steps:

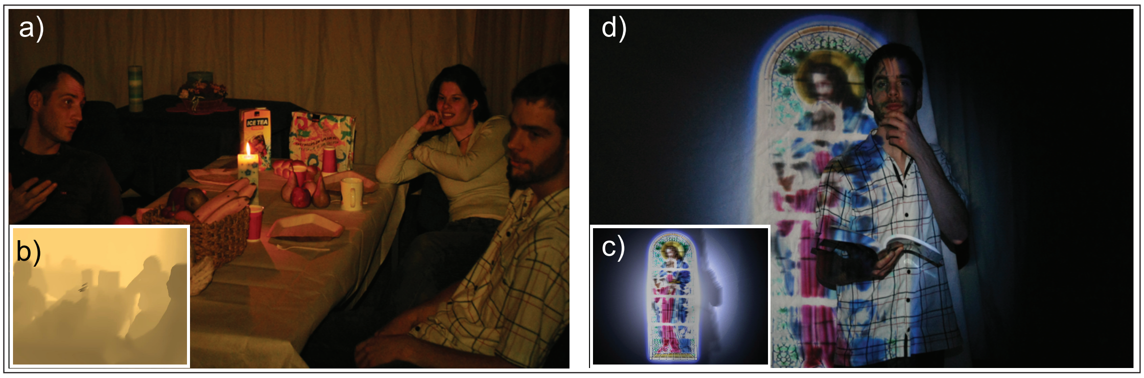

Figure 7.7: *a) A scene lit with a color filtered flash. b) Artistic filter applied for scene a). c) Artistic filter applied to scene d). d) A scene lit with a color image flash.*

1. Acquisition of depth and amplitude (infrared), and two color images
2. Depth filtering using high resolution color
3. Reprojection from the upsampled depth-data to the flash unit
4. Per pixel, depth-dependent light adjustment/reflectance correction
5. Color camera capture

Since most of the steps are very similar to the SSFA approach, we will focus on the enhanced filtering using color Sect. 7.5.1 and on the flash-tone-mapping Sect. 7.5.2.

7.5.1 Depth Filtering with Color

For reasons mentioned in Sect. 7.4.1, a sophisticated filtering is very important. Taking into account the high resolution color image, allows to filter and upsample the low resolution depth data, and therefore, get higher detail. Kopf et. al [KCLU07] presented the joint bilateral upsampling for images, where the mapping of corresponding areas between a low resolution image and a high resolution image is given. We map the depth data from the depth camera to the color camera by using the reprojection step described in Sect. 7.4.2. To improve the correspondences between the depth and the color values, we filter the depth data, according to Sect. 7.4.1, prior to the mapping. As a result we obtain the trivially upsampled depth data $\tilde{D}$ being of the same resolution as the color image $\tilde{I}$. To improve the performance of the filtering step, we apply a technique motivated by the joint bilateral upsampling [KCLU07].

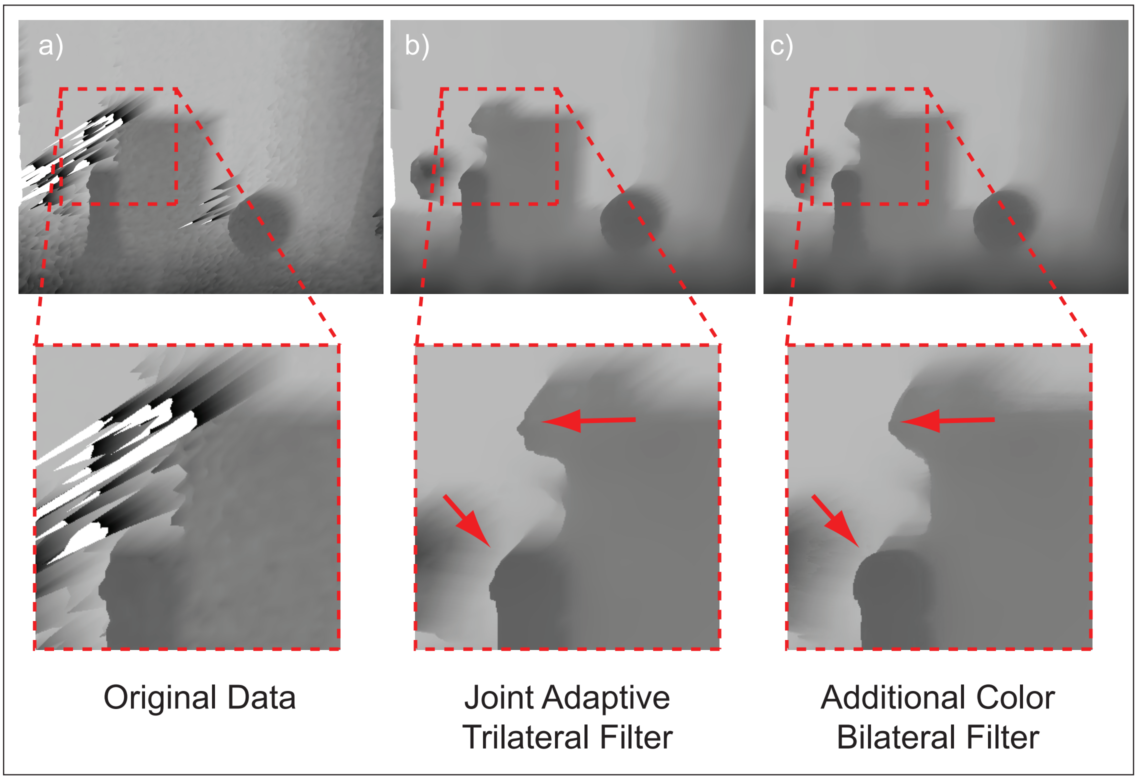

Figure 7.8: *a) The unfiltered depth contains noise and some outliers. b) Applying an joint adaptive trilateral filter allows to reduce the noise, while keeping the depth discontinuities of the scene. However, the resolution of the depth camera can result in blocky edges. c) Additionally applying a modified joint bilateral upsampling based on the high resolution color camera, refines the edges at depth discontinuities.*

$$\tilde{D}_{\mathbf{m}} = \frac{1}{k_c} \sum_{\mathbf{p}_{\downarrow\uparrow} \in \tilde{\Omega}_{\downarrow\uparrow}} \tilde{D}_{\mathbf{p}_{\downarrow\uparrow}} \cdot \tilde{g}(\|\mathbf{m} - \mathbf{p}_{\downarrow\uparrow}\|) \cdot \tilde{h_c}(\|\tilde{I}_{\mathbf{m}} - \tilde{I}_{\mathbf{p}_{\downarrow\uparrow}}\|) \tag{7.12}$$

$$k_c = \sum_{\mathbf{p}_{\downarrow\uparrow} \in \tilde{\Omega}_{\downarrow\uparrow}} \tilde{g}(\|\mathbf{m} - \mathbf{p}_{\downarrow\uparrow}\|) \cdot \tilde{h_c}(\|\tilde{I}_{\mathbf{m}} - \tilde{I}_{\mathbf{p}_{\downarrow\uparrow}}\|) \tag{7.13}$$

Instead of evaluating the full filter kernel, we evaluate every n^{th} pixel at position $\mathbf{p}_{\downarrow\uparrow}$ on the filter kernel $\Omega_{\downarrow\uparrow}$. n is the ratio between the width in pixels of the color camera, and the width in pixels of the depth camera, and for the height respectively. The closer the depth camera and the color camera are to a paraxial setting, the closer this approach is to the joint bilateral upsampling. The function $\tilde{g}(\cdot)$ and $\tilde{h_c}(\cdot)$ are Gaussian kernels with a kernel support of $\tilde{\Omega}$. The color at position m is referred to as $\tilde{I}_m$.

7.5.2 Flash-Tone-Mapping

In the previous sections we presented a method to adjust the amount of light per pixel depending on the per pixel depth value. However, we did not take the reflectance of the scene into account. Regions of the scene, which were not lit before adding the flash light, might saturate when doing so. Therefore, not taking into account the reflectance, would not allow to compensate for this. We propose to use a flash tone-mapping with a locally variable projection, which is motivated by the dodging-and-burning approach presented by Reinhard et. al in [RSSF02b].

The response curve of the camera without flash $g_r(z_i)$ and the maximum extent of the response curve with flash $g_F(z_i)$, can be defined as:

$$\text{Camera:} \qquad g_r(z_i) = ln(E_i \cdot \Delta t) \tag{7.14}$$

$$\text{SAFU:} \qquad g_F(z_i) = ln(\Delta t \cdot E_i + \Delta F_i) \tag{7.15}$$

The image pixel values z_i are only dependent of the irradiance E_i and the chosen exposure time Δt for the camera response. $g_r(z_i)$ can be determined using a method presented by Debevec et. al [DM97]. Since the duration of the flash unit is usually much shorter than the exposure time, its response function $g_F(z_i)$ does not depend on the exposure time.

Depth Independent Range Extension

We want to find a new response curve $g_n(z_i)$ for the camera-flash system, which is similar to the response curve $g_r(z_i)$ of the camera, but extends its dynamic range independent of per pixel depth. Furthermore, it shall avoid overexposure in the flash lit image. $g_n(z_i)$ is bound by the lower bound $g_l(z_i)$ and the upper bound $g_r(z_i)$. The lower bound is determined as:

$$g_l(z_i) = ln(\Delta t E_i + \Delta F_{i_{max}}) \tag{7.16}$$

$$= ln(\Delta t E_i + min(\Delta F_{xy}(z_i))), \text{ where} \tag{7.17}$$

$$\Delta F_{xy}(z_i) = e^{g_r(I'_{xy})} - e^{g_r(I_{xy})}, \forall xy \in I_{xy}. \tag{7.18}$$

The image pixels I_{xy} of the no-flash image at position (x,y), and the image pixels of the flash image I'_{xy} lie in the same range as z_i. The solution for $g_n(z_i)$ can now be approximated as a linear scaling of $g_r(z_i)$ by factor k (cf. Eq.(7.19)). $min(\frac{g_l(z_i)}{g_r(z_i)})$ complies with the lower bound $g_l(z_i)$. $g_n(z_i)$ could be smoothed slightly in order to avoid discontinuities in the fixed range (cf. Fig. 7.10).

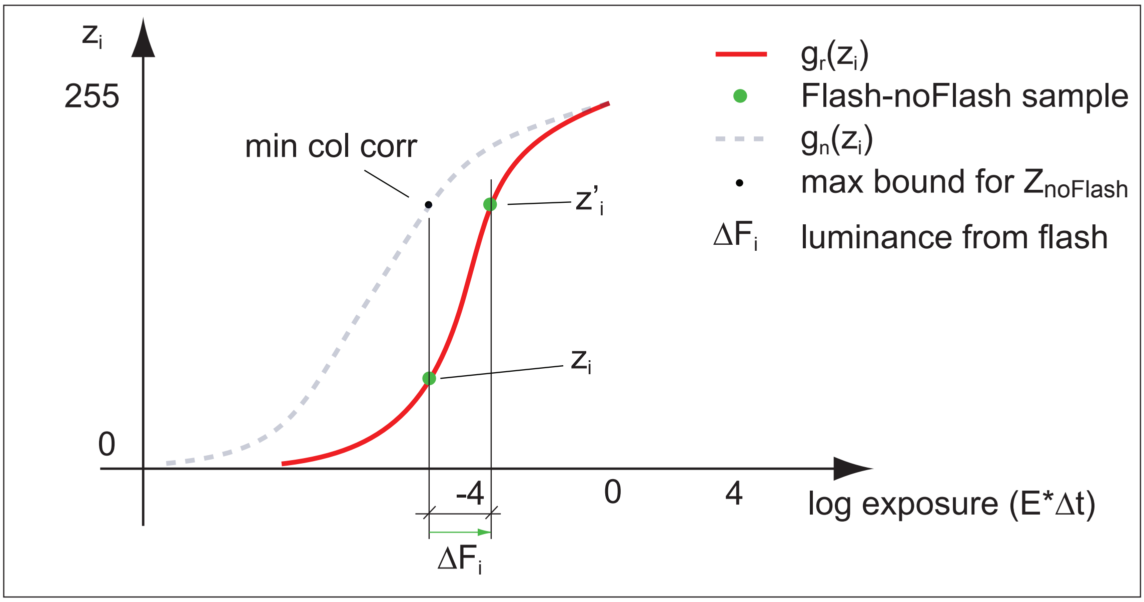

Figure 7.9: *The added luminance by the flash ΔF_i, can be evaluated from the difference of the log exposure $e^{g_r(z'_i)}$ of the flash image and the log exposure $e^{g_r(z_i)}$ of the noflash image.*

Given $g_n(z_i)$ we can find the corresponding luminance contributed by the flash as

$$g_n(z_i) = k \cdot g_r(z_i) = min\left(\frac{g_l(z_i)}{g_r(z_i)}\right) \cdot g_r(z_i). \tag{7.19}$$

The scaling factor s_{xy} of the required light per pixel (x, y) of the SAFU can now be evaluated as $s_{xy} = \frac{\Delta F_{i_{max}}}{\Delta F_{xy}}$ with $\Delta F_{i_{max}} = g_n(z_i) - g_r(z_i)$. Assuming only direct light contribution by the projector, lets us scale the maximum intensity of the corresponding projector pixel by the scaling factor s_{xy} as well.

7.6 Results

We built a prototype SAFU by replacing the light bulb of a SANYO PLC-XW50 projector with a Canon Speedlite 580EX II. The resolution of the projector is 1024 × 768 pixels. The depth is measured using a infrared based ToF-camera from Swissranger with a resolution of 176 × 144 pixels. The images were captured using a Canon 1D Mark III with a Canon EF 16-35mm f/2.8L USM lens. The resolution of the color camera is 3888 × 2592 pixels. We control the SAFU from our application, which is running under Windows XP on an Intel Core Duo with 2.6GHz and 3GB of RAM.

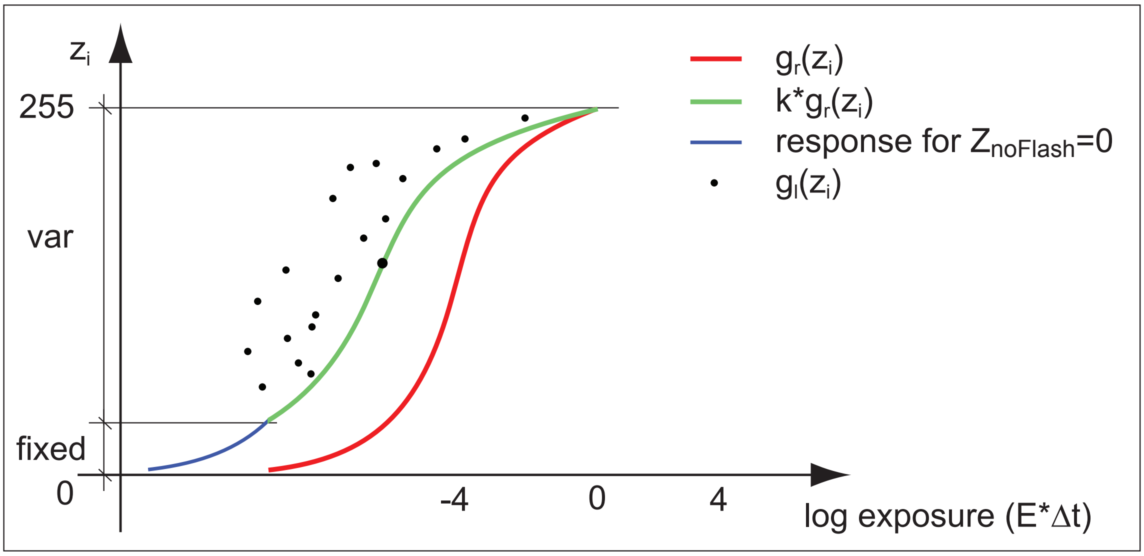

Figure 7.10: *The modified response function of the system has to lie in between the lower bound $g_l(z_i)$ and the upper bound $g_r(z_i)$. Due to the different depth in the scene, $g_l(z_i)$ is most often not a smooth curve. To illustrate one possible response function we choose to linearly stretch the camera response function respecting the lower range, while avoiding overexposure.*

We captured different scenes using different settings in order to demonstrate the various capabilities of our method Fig. 7.11.

7.7 Limitations

While we believe our method has potential, it also suffers from several limitations. One of the most crucial limitations of our setup is the efficiency of the maximal light output. Using an LCD-projector for the spatial light modulation, drops the light by at least 50%, even if the attenuation of the SAFU is set to zero. This is currently limiting the application of the SAFU to low light settings. Unfortunately, it seems that a substitution by DMD technology, as has successfully been realized for light modulation during acquisition [NBB06], is not an option for a flash unit. The non-controllable duration of the flash burst is in the order of 1.2ms, and therefore, much shorter than the required integration time of the DMD. The limitation of the maximal intensity as well as the limited contrast are setting constraints on the maximal depth compensation the SAFU can deal with.

Our system contains only one SAFU, such that we cannot compensate for shadows. By adding a second unit to the system, these shadows could be minimized.

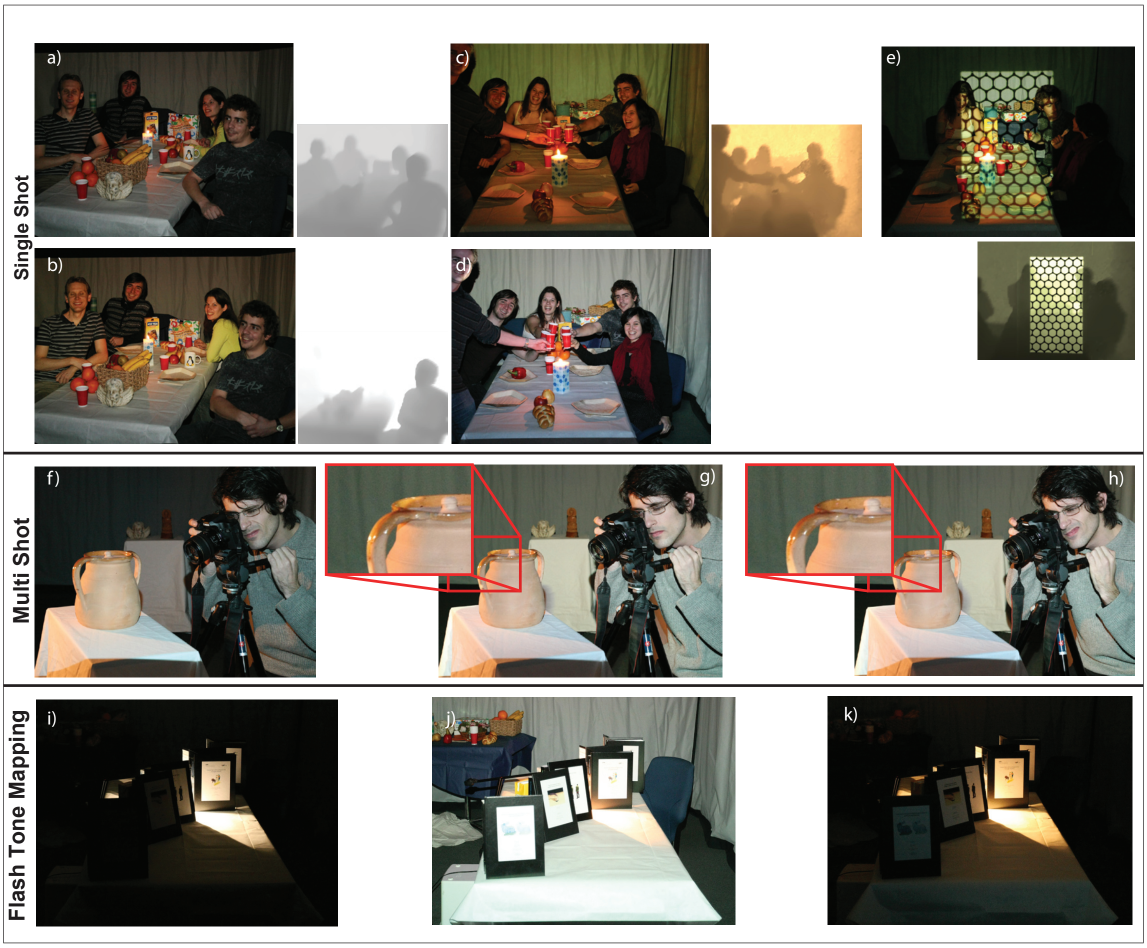

Figure 7.11: *a) Shows a scene with the virtual light source being placed 5m behind the photographer. b) Placing the light source at 10m behind the photographer, results in an almost homogeneous lighting of the scene. Applying a spatially adaptive flash without color correction to a candle lit scene, removes the nice warm tones from the candle light d). Adding a color filter, keeps the mood of the scene, while still brightening it up for a shorter exposure time c). e) More artistic effects can be created with an image filter. f) Scenes with a big depth extent result in a harsh lighting when photographed with a regular flash. g) The single-shot allows to compensate for the intensity falloff, but suffers from the low resolution of the ToF-camera, which can result in misalignments. In h) we take advantage of the high resolution of the color camera, to improve the accuracy of our spatial compensation. i) This scene is only lit by a desklamp in a dark room resulting resulting in a wide dynamic range. j) shows the same scene as in i) but captured with a full intensity flash burst. k) shows the same scene as in i) but with added flash light in the dark regions, without overexposing the bright ones.*

7.7.1 Single-Shot

Depth Camera The single-shot method is strongly based on the data provided by the ToF-camera. Non-IR reflecting objects or very specular objects, such as mirror-like surfaces, cannot reliably be reconstructed. Furthermore, the limited resolution can become visible as aliased edges at sharp depth-discontinuities.
Processing Time The current processing time lies at 0.064s for a single-shot acquisition. This processing time limits the amount of motion, which can be present in the scene. The motion does not appear as motion blur in the images, but as a misalignment of the geometry with the flash compensation function. We believe, however, that our algorithm could be implemented directly on a processing unit of modern digital cameras.
Spatial Limitation The spatial limitation is twofold. On the one hand, the ToF-camera features a range of 7.5m without any depth ambiguities, and on the other hand, the limited light efficiency of our prototype allows for a compensation of a couple of meters depending on the settings.

7.7.2 Multi-Shot

Color Camera Acquisition The time to acquire a color image over the USB connection using the Canon SDK lies in the range of one second. This is a severe limitation and makes it impossible to capture spontaneous shots. This is, however, a limitation of our specific setup and could be alleviated if the algorithm was implemented directly on the camera or in the flash unit.
Limited Dynamic Range The current SAFU implementation does not show a very high light efficiency. This has a direct impact on the amount of correction we can apply to a scene.

7.8 Conclusion and Future Work

We presented a spatially adaptive photographic flash, which improves on the limited illumination control of a common flash unit. By modulating flash intensity on a per-pixel basis, we can simulate a virtual light source further away from the subject than the physical light source. This reduces the harshness usually associated with flash illumination. We have built a prototype of our proposed flash unit, and we have shown the effect of moving the virtual light source to different positions. Furthermore, we have presented a flash tone-mapping technique that achieves an effect similar to "dodging and burning". Our spatially adaptive photographic flash relies on depth data to compute its spatial modulation. To reduce noise in this data, we have presented a novel joint adaptive trilateral filter, which adapts to the noise model of the depth data.

7.8.1 Possible Applications

In our work, we use a spatially modulated flash mainly to compensate for deficiencies in a scene's natural illumination. However, structured illumination can accomplish many other things, such as separating direct and indirect illumination [NKGR06] or changing the apparent color, shape, or surface properties of objects [RWLB01]. By introducing additional projectors or optics, one can create a 4D illumination light field, which can be used to estimate the surface properties of objects [DHT+00b], digitize their shape [RHHL02, ZCHS03], eliminate foreground or background objects [LCV+04], or even compute an image as seen from the flash's point of view [SCG+05].

7.8.2 Light Field Illumination

We have presented many examples to enhance the illumination in common photography. However, the spatially adaptive illumination can be applied identically for light field acquisition using a plenotpic camera [NLB+05]. If the light field would be acquired using a set of spread out cameras, the self-shadowing would become more noticeable. Therefore, the spatially dependent light adjustment would have to be adapted for acquisition settings featuring a big parallax.

CHAPTER 8

Conclusion and Future Work

This chapter provides a short summary of this thesis and provides some ideas for potential future work.

8.1 Conclusion

In this thesis, we have presented a novel wave-based image generation framework, which allows to record, store, reconstruct and render entire scenes based on a wave-based representation of light. As a central component of this framework, we introduced the hologram as a data representation of a wavefield in space. This physically motivated representation allows to store complex-valued wavefields as a real-valued matrix, and can therefore, also be used as an input to holographic screens. Being able to reconstruct the entire wavefield of a scene, allows to encode all reflection properties as well as depth cues a human being can perceive. So far, the hologram is currently the only representation, which allows to record, store and reconstruct a complex-valued wavefield. Furthermore, the holographic data representation can be used for wave-based rendering. In contrast to previous work, where the hologram proved to be useful for full-parallax screens, we show various ways of integrating holograms in a graphics pipeline while combining it with traditional graphics representations.

Our contributions are manifold, reaching from CGH, over white-light hologram recording, to wavefield rendering and compositing. We have presented novel ways of handling occlusions directly in a wave-based approach alleviating the per hologram pixel visibility check required in ray-based evaluations. This brings advantages for CGH, but also for wave-based rendering, since diffraction is handled implicitly. Diffraction specific glare, as it is perceived in photography or in a human visual system, can be handled straightforwardly using our method. Furthermore, we showed novel ways of efficiently evaluating the illumination of a scene on the entire hologram, by interpolating between evaluated viewing positions.

Some scenes are hard to be modeled artificially, such that it would be easier to capture the real scene. Therefore, we presented a novel light field to hologram transform, which allows to capture holograms under white light conditions. Holograms can, therefore, be captured under the same conditions as photographs. A novel spatially adaptive photographic flash further enables to capture holograms in dark conditions, by illuminating distant scene parts, while avoiding overexposure of nearby objects. With CGH and white light hologram recording, artificial objects as well as real scenes can be recorded onto a hologram.

Unlike previous work, we do not visualize the hologram onto a holographic screen, although our representation would allow to do so. Rather, we render the hologram from different viewpoints using a physically plausible camera. Many effects of a real camera are handled implicitly. For instance, we implicitly obtain images with a depth of field depending on the aperture size, we can vary the focal length very easily and intuitively, and obtain aliasing-free renderings. Based on a novel depth reconstruction from the hologram, the renderings can also be composed with renderings from a traditional graphics pipeline.

8.2 Future Work

This work covers different areas, from CGH, over light field - hologram transform to wave-based rendering. We discuss the potential of future work relative to these areas in the following subsections.

8.2.1 Computer Generated Holography

One of the main challenges of wave-based frameworks is the wavelength dependent sampling of the wavefront. Therefore, it is important to decouple the computational complexity as much as possible from the resolution. In Chapter 5, we showed a way to handle occlusion in a wave-based approach, without having to evaluate the visibility at every hologram sample. The lighting model is currently evaluated only on a subset of the hologram samples. This is improving the efficiency a little, but it is also decreasing the maximal frequency of the BRDF, which can be evaluated through CGH. Therefore, we would like to enhance our framework to fully simulate reflectance and illumination in a wave-based way as well. This would allow using only one propagation step while maintaining view dependent reflectance.

In spite of hardware acceleration, CGH and wave-based rendering is still far from real-time. This is certainly one of the biggest limitations in order to use holograms as a common data structure as well as display technology. By further exploring the parallel nature of wave evaluation, we could drastically accelerate the rendering. Today's graphics cards already provide a highly parallel processing unit, which can be programmed for general purpose computation using NVIDIA® CUDA™. We believe that computationally expensive simulations will become feasible in the next decade, opening up many possibilities for future work in this field.

Sampling the interference pattern according to the Nyquist frequency leads to very high resolution and a lot of data per hologram. Therefore, good hologram compression will be a major challenge in the future. Hologram compression does not lend itself for standard image compression techniques, since the resulting image quality is often based on the entire frequency spectrum. Non linear quantization approaches have shown to be successful, although we believe that they do not exploit the entire compression potential of a hologram.

8.2.2 Light Field - Hologram Transform

In Chapter 6, we presented a bidirectional transform between the light field and the hologram. However, this transform does not compare the amount of information, which is stored in either representation. On the one hand, the light field does not directly store information about depth, while the hologram is encoding the depth in the phase. On the other hand, the hologram is prone to speckle noise caused by destructive interference, while the light field does not suffer from it. The information content stored in either representation becomes especially interesting in the context of data storage, redundancy and compression.

In Sect. 6.6.2, we described a way to find the maximum resolution of a view of the hologram, such that the entire pictured scene lies in the DoF. However, if the scene would not have to lie in the DoF, we could employ a bigger aperture, leading to a higher resolution. By rendering multiple images with a big aperture, but varying focal lengths, we could get a focal stack of the object at a higher resolution than the resolution constrained by the DoF criteria. The blurred parts lying outside of the DoF could be partially removed by a spatially varying deconvolution. Another way of creating a focal stack, could consist in the propagation of the wavefield of the hologram to different parallel planes in space, while creating images with a camera of constant focal length. Due to the limited hologram size, the deconvolution kernels would still not be spatially invariant. The spatial invariance of the convolution kernels make this problem very hard to solve. Furthermore, by increasing the aperture size, we limit the maximum frequency of the reconstructable BRDF. Therefore, the reconstruction would become a tradeoff between the maximum frequency of the BRDF and the maximum spatial resolution. Instead of using a lens, one could also propagate the input wavefield back in space and try to recreate the whole 3D wavefield in object space. In case of an infinitely large hologram, the deconvolution kernel is spatially invariant and corresponds to a spherical wave with unit amplitude. In practice, the hologram is often limited in space, which is leading to a spatially variant deconvolution again. We believe, that the maximum resolution of an image of a hologram, and the maximum frequency of its BRDF, are closely related to the amount of information stored in a hologram. Therefore, we expect future work in this area to contribute valuable knowledge about the information content of a hologram.

8.2.3 Wave-based Rendering

Based on our pipeline, we are able to produce nicely looking results. However, the image quality is still far from being comparable to high quality renderings. One of the major drawbacks of renderings from holograms is speckle noise. It is based on the destructive interference of the monochromatic light. We have successfully shown different approaches to reduce effects of speckle noise. However, they are often very time consuming to compute. Speckle noise is not only degrading the image quality, but also causes problems during depth reconstruction from the hologram. Therefore, we believe that this is a very important area of future work, since it provides a fundamental requirement, namely image quality.

The quality of the lighting of hologram renderings is currently limited by the noise of the reconstructed normals as well as by the limited depth-of-field renderings. Improving the filtering of the depth map, on the one hand, and extending the lighting approach to renderings with a limited depth-of-field, on the other hand, would increase the real-time lighting quality.

Current CCD sensors have a pretty big pixel pitch, which make it impossible to capture large fields of view. In Sect. 6.7.2, we have shown a rendering from a digital hologram featuring a field of view of roughly 3°. By miniaturizing the pixels even further, and increasing the number of pixels of a CCD, the capturing of digital holograms would allow to acquire 3D scenes with a large field of view without having to use any optical elements.

Certainly one of the key issues is the required computational time. Once the processing of big holograms can be handled in real-time, relighting, reshading, and deformation operators in wave domain will become of highest interest in this field.

APPENDIX A

Resolution Limits

The transformation of a hologram into a light field requires the evaluation of the theoretical limit of the resolution of a single light field view. This resolution limit is dependent on the wave-length, aperture width and depth of field. This section describes four expressions, which are required to explain the physical resolution limits.

A.1 Circle of Confusion

The Circle of Confusion (CoC) c is defined as the size of the circle to which an idealized point will diverge when the lens is focused at different length. Assuming a ray representation of light, the CoC caused from defocus $c_{defocus}$ is

$$c_{defocus} = \frac{f^2(z_f - z_n)}{N(2z_n z_f - f(z_n + z_f))}\,, \tag{A.1}$$

with N being the "f-number" defined as $N = \frac{f}{a}$ [Con04], and f being the focal length of the thin lens.

This implies that the CoC at focal length f is zero. Even with a perfect lens, a point will not lead to a point in the image but to the Airy disk governed by diffraction when simulating light as waves. The CoC limited by diffraction c_{diff} is given by $c_{diff} \approx 2.44\lambda N(1+m)$ with m being a magnification factor $m = \frac{f}{z-f}$.

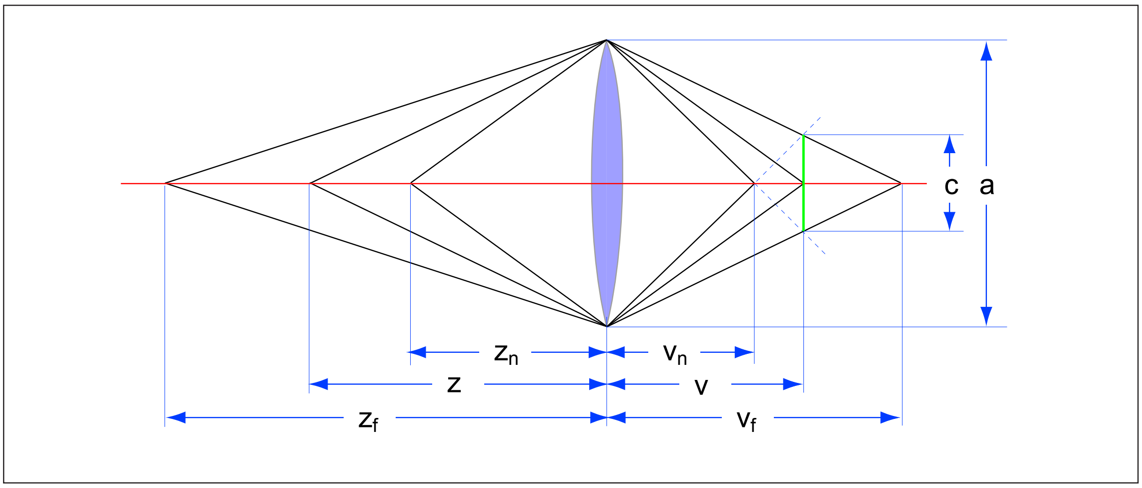

Figure A.1: *The CoC c depends on the aperture a of the thin lens depicted in blue, and the the focal length f. The focal length can be determined using the thin lens equation.*

Despite the physical limitation of the minimal CoC given by diffraction c_{diff}, the CoC can also be restricted by the resolution of the discretizing media (e.g. film grain, pixel size of a camera) and therefore, regarded to be in focus. In our case the CoC is limited by the angular resolution $\Delta\alpha$ given by the aperture size a and the wavelength λ for a hologram and by the pixel size $max(\Delta s, \Delta t)$ of the CCD when capturing a light field.

A.2 Depth of Field

The Depth of Field (DoF) $\Delta z = z_f - z_n$ as shown in Fig. A.1 is the distance between the closest point in focus z_n and the farthest point in focus z_f, where a point in focus is determined by the CoC. The DoF considering only geometrical optics is given as

$$\Delta z = \frac{2zNcf^2(z-f)}{f^4 - (Nc(z-f))^2} . \tag{A.2}$$

There is no simple numerical expression to combine the effects of defocus and diffraction besides using an empirical expression. Furthermore, the perception of sharpness is not solely dependent of the finest resolution but also on contrast. The ability to transfer contrast of an input pattern with a given frequency ν by a diffraction-limited lens with defocus can be described by the Optical Transfer Function $OTF(c_{defocus}, N, v)$ as shown in [Hop55] by

$$OTF(c_{defocus}, N, \nu) = \begin{cases} \frac{4}{\pi g} \int_0^{\sqrt{1-s^2}} sin\left(\sqrt{1-y^2} - s\right) ds & \text{if } s \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (A.3)$$

where $s = \lambda \nu N(1+m)$ and $g = \pi \nu c_{defocus}$. In [Jan97] and [BW59] the maximal resolution according to the Rayleigh criterion is given, as long as two points can be separated so that a 19% dip appears between the peaks. Applying this threshold to the OTF leads to a maximal spatial frequency ν, which can be resolved by the lens. Diffraction can be ignored if defocus is sufficiently big or N sufficiently small leading to a simplified $OTF(c, \nu) = \frac{2J_1(g)}{g}$ with $J_1(\cdot)$ being the first-order Bessel function of the first kind. Fig.A.2a shows three curves for the OTF corresponding to the plane in focus without defocus, the OTF dependent of defocus and diffraction at the DoF limits and the OTF ignoring diffraction.

In Fig. A.2b the spatial frequency ν depending of the f-Number N with OTF=0.19 and constant DoF is depicted. This gives the possibility to choose a desired resolution represented as a spatial frequency and setting the depth of field including the complete scene in order to obtain the biggest possible aperture.

A.3 Resolution / Sampling

Different resolutions and samplings have to be considered for the recording and the transformation of a hologram and a light field. By *resolution* we mean the maximal number of samples in a certain dimension, whereas *sampling* refers to the size of one sample. We take into account the maximal number of pixels n_{Xmax} of a resulting image and the number of resulting camera positions n_{Umax}. For simplicity the resolution and sampling is always given for one dimension and can be handled analogously for the other dimension.

Holograms with resolution n_{holo}^2 and extent a_{holo} can have a minimal angular sampling of $\Delta\alpha = \arcsin\left(\frac{\lambda}{a_{holo}}\right)$. The angular resolution is defined by the number of samples n_{apt} of the aperture simulating the camera. The maximal resolution $n_{max} = \frac{2a_{holo}}{\lambda}$ is achieved if the FoV (see Sect. A.4) is 180°. Furthermore, the number of useful camera positions for the transformation equals n_{holo} in each dimension.

Angular parameterized Light Fields are most easily compared to holograms. The angular resolution of the image is given by the number of samples in θ-dimension with the maximal resolution $n_{max} = \frac{180°}{\Delta\theta}$ with $\Delta\theta$ being

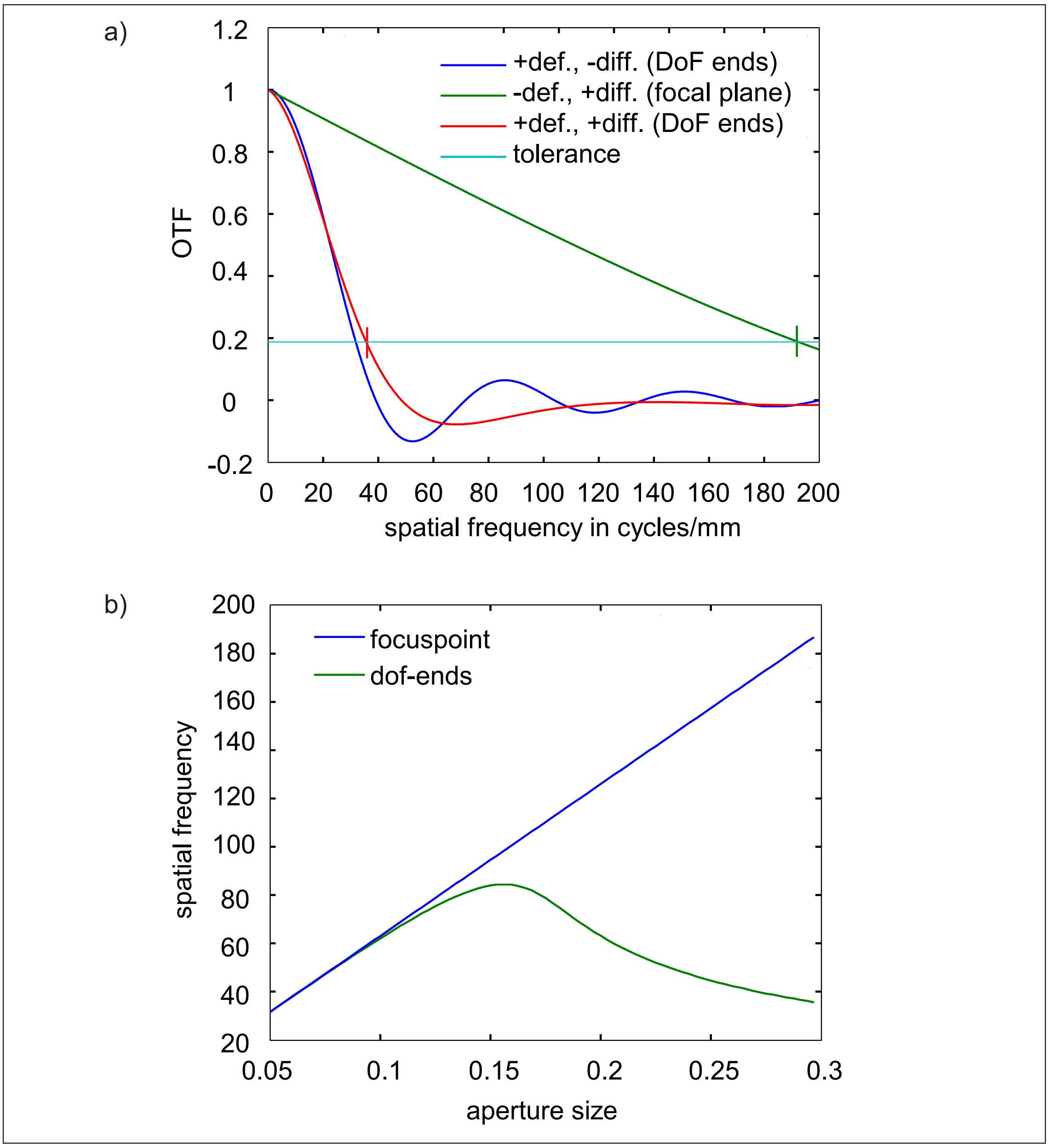

Figure A.2: *The graph in a) shows the OTF at focal distance in green and DoF limits in red, as well as the OTF dependent of defocus only in blue. In graph b), the curve for an OTF=0.19 at focal distance and at DoF limits is given in blue and green, respectively.*

the sampling distance. The number of possible camera positions equals the number of samples in u.

For **Two-plane parameterized Light Fields**, the maximum resolution of a rendered image is equal to the maximum number of samples in the s direction. The number of different views depends on the number of samples in the u dimension.

Comparison By neglecting compression, we can see that the hologram is capable to store more different views with high resolution, than any of the light field representations. This efficiency is due to the sampling of the wave-field rather than the sampling of different viewpoints, where information like depth is being disregarded. However, holograms are prone to speckle noise because of the coherent light. Speckle size can be defined as $d_{sp} = \frac{\lambda b}{a}$ with b being the distance of the point source to the imaging system and a being the aperture size.

A.4 Field of View

The maximal FoV α_{max} of a hologram can be determined by the sample size Δu as $\alpha = \arcsin\left(\frac{\lambda}{2\Delta u}\right)$. This implies a phase difference between two samples of maximally π just reaching the Nyquist frequency. The FoV α of a freely chosen aperture depends on the number of samples n inside the aperture, since one sample corresponds to an angular sampling of $\Delta\alpha$ and the FoV to $\alpha = \Delta\alpha \cdot n$ not exceeding α_{max}.

The maximal FoV α_{max} for the light field depends on the parametrization. For $LF(u,v,\theta,\phi)$ the maximum is defined as $\alpha_{max} = max(\theta,\phi)$ considering θ and ϕ of all rays. For a two-plane parametrization $LF(u,v,s,t)$ the FoV depends on the extent and the distance between the uv-plane and the st-plane. Furthermore, the view frustum is sheared if only rays going through both planes are considered.

APPENDIX A

Notation

A.5 Chapter 1-6

- u .... Scalar wave.
- P, P' .... 3D point.
- t .... Time.
- n .... Refractive index.
- c .... Velocity of light in vacuum.
- U .... Complex function describing the wavefield (phasor).
- ω .... Angular frequency.
- ν, ν_x, ν_y .... Frequency of oscillation, frequency in x, frequency in y.
- A .... Amplitude.
- φ .... Phase.
- k .... Wave number.
- λ .... Wavelength.
- (α, β, γ) .... Vector of directional cosines.
- $\mathbf{r}$ .... Vector of propagation evaluation.
- r .... Propagation distance.
- $\mathbf{n}$ .... Surface normal.

S_A .... Aperture.
S .... Surface.
$\mathbf{i}$ .... $\sqrt{-1}$.
x,y,z,ξ,η,ζ .... Coordinates.
M .... Rotation matrix.
$\mathcal{A},\hat{\mathcal{A}}$ .... Angular spectrum, shifted angular spectrum.
N .... number of samples.
m,l .... Discrete frequency components.
δ .... Sample size.
G .... Propagation function.
$\hat{U},\hat{G}$ .... Zero-padded U, zero-padded G.
$*$ .... Convolution.
$\mathcal{F},\mathcal{F}^{-1}$ .... Fourier Transform, inverse Fourier Transform.
$H,\mathcal{H},H'$ .... Hologram, Fourier Transform of hologram, filtered hologram.
$R,O,\bar{O}$ .... Reference wave/reconstruction wave, object wave, complex conjugate object wave.
ϑ .... Angle between O and R.
T .... Transmittance.
t_b .... Bias transmittance.
β' .... Scaling factor.
d .... Sampling distance.
A_O,A_R .... Amplitude of object wave, amplitude of reference wave.
$\check{U}$ .... Focused wavefield.
L .... Lens function.
u,v .... Coordinates of image pixel.
θ .... Field of view.
r_a .... Aspect ratio.
w,h .... Width and height of aperture.
W .... Size of aperture.
δ_f .... Sample size in frequency domain.
D .... Diameter of circular aperture.
d_o,d_i .... Distance from lens to aperture, Distance from lens to image plane.
f .... Focal distance.
$\lambda_r,\lambda_g,\lambda_b$ .... Wavelength of red, green and blue.

I .... Intensity of wave.
ϵ .... Small scaling factor.
$\Delta\varphi$ .... Phase difference.
d_max .... Maximal reconstructable distance.
PS .... Point source.
$\mathbf{o}_x$ .... Complex-valued occluder of plane x.
$\overline{U}_B'$ .... Propagated wavefield after applying the occluder.
$\mathcal{P}_{A\to\hat{B}}$ .... Wavefield propagation from plane A to plane B.
PC_A .... Propagation cone.
$U_A,\mathcal{U}_A$ .... Wavefield on plane A and its Fourier Transform.
$\mathcal{G}$ .... Fourier Transform of the propagation function G.
$s_i,\Delta s_i$ .... Lighting evaluation sample with index i, distance between lighting evaluation pixels.
S .... Subset of hologram pixels.
β .... Angle limiting the BRDF sampling.
d_i .... Sample of depth map.
M_P^{-1} .... Back projection matrix.
a,b .... Target plane parametrization.
u,v,s,t .... Parameters of two plane parametrization.
θ,ϕ .... Angular samples of light field.
$\mathbf{M}$ .... Mapping from light field to hologram.
$\mathbf{M}^{-1}$ .... Mapping from hologram to light field.
$\widetilde{EV}$ .... 4D epipolar volume.
$\widetilde{EV_s}'$ .... Sheared epipolar volume.
EPI .... Epipolar plane image.
l_c .... Line of corresponding points in EPI.
p_c .... Plane of corresponding points in $\widetilde{EV}$.
p_s .... Shear plane.
s_xy .... Shear factor.
$\Delta u,\Delta v,\Delta s,\Delta t$ .... Sampling distance of u,v,s,t.
LF .... Light field.
FMC .... Frequency minimization criteria.
w .... Frequency dependent weighting function.
$\overline{p_s}$ .... Mean of shear plane.
f_{apod} .... Apodization function.
d_{px} .... Disparity in pixels.
z_R .... Reconstructed depth value.

w .... Number of shear steps.
P_{uv} .... Point source of particular tile.
T_{uv} .... Tile at (u, v).
$\Delta\hat{z}_{visScene}$ .... Depth extension of visible part of the scene.
Δz .... Depth of field.
h .... Windowing function.
$\nu l_u, \nu l_v$ .... Local spatial frequencies in u and v.
$\hat{U}$ .... Wavefield with applied windowing function.

A.6 Chapter 7

σ_D .... Standard deviation of depth values.
c .... Velocity of light in vacuum.
f_{mod} .... Modulation frequency of the depth camera.
B .... Noisy signal.
A_{sig} .... Mean number of electrons from IR signal.
BG .... Background illumination.
c_{demod} .... Demodulation ratio.
$g()$ .... Spatial proximity filter.
$f()$ .... Depth component range filter.
$h()$ .... Amplitude dependent range filter.
$d_{\mathbf{m}}$ .... Depth value of $\mathbf{m}$.
$\mathbf{m}, \mathbf{i}$ .... Position in depth and IR image respectively.
k_c .... normalizing factor.
Ω .... Filter support.
$a_{\mathbf{m}}$ .... Amplitude at $\mathbf{m}$.
σ_p .... Standard deviation of spatial filter.
σ_d .... Standard deviation of depth range filter.
σ_a .... Standard deviation of amplitude range filter.
τ_a .... Amplitude threshold.
τ_m .... Threshold for median filter.
$\sigma_{a_{init}}$ .... Initial standard deviation of σ_a.
$\sigma_{a_{min}}$ .... Minimal standard deviation amplitude filter.
$\sigma_{d_{init}}$ .... Initial standard deviation depth filter.

τ_m .... Threshold for median filter.
$\mathbf{P}$ .... Set of image samples.
h_p .... Depth range filter for shallow edge diffusion.
$d_{\mathbf{i}_{min}}$ .... Shallowest depth value of the kernel support.
I_0 .... Intensity of illumination source.
I_d .... Intensity of illumination at distance d.
d .... Distance from object to camera.
d_{off} .... Depth offset.
I_{sub} .... Discrete intensity values.
d_{sub} .... Focal length of camera.
f_c .... Depth compensation function.
f_{att} .... Attenuation function.
I_{max} .... Maximal intensity of illumination source.
d_{cmax} .... Maximal compensation depth.
$\tilde{D}_{\mathbf{m}}$ .... Filtered depth value at $\mathbf{m}$.
$\tilde{I}$ .... Color image.
$\mathbf{p}_{\downarrow\uparrow}$ .... Position of n^{th} pixel of $\tilde{I}$.
n .... Ratio between width in pixels of the color to the depth image.
$\tilde{g}$ .... Spatial filter of depth filter with color.
$\tilde{h}_c$ .... Color range filter.
$\tilde{I_{\mathbf{m}}}$ .... Color pixel at $\mathbf{m}$.
z_i .... Image pixel values.
E_i .... Irradiance.
Δt .... Exposure time.
g_r .... Camera exposure function.
ΔF_i .... Flash luminance.
g_F .... Camera response function including flash.
g_n .... New response function.
g_e .... Lower bound response function.
I'_{xy} .... Image pixel of the flash image.
k .... Scaling factor.

List of Figures

List of Tables

Curriculum Vitae

Remo Ziegler

Personal Data

Okt. 17, 1975	Born in Uster, Switzerland
Marital Status	Married
Nationality	Swiss
Languages	German (fluent), English (fluent), French (fluent), Spanish (intermediate), Chinese (basic)

Education

May. 23, 2008	Ph.D. defense.
Aug. 2004 – May. 2008	Research assistant and Ph.D. student at the Computer Graphics Laboratory of the Swiss Federal Institute of Technology (ETH) Zurich, Prof. Markus Gross.
Dec. 15, 2000	Diploma degree in Computer Science.
Oct. 1995 – Dec. 2000	Diploma Studies of Computer Science, ETH Zurich, Switzerland. Specialization: Computer Graphics and Computer Vision; Complementary studies: Biomedical Engineering.

Awards

Sep. 2007	Günter Enderle Award for the Best Paper "A Bidirectional Light Field - Hologram Transform" at Eurographics 2007 .
Sep. 2007	Best Student Paper Award "A Bidirectional Light Field - Hologram Transform" at Eurographics 2007.

Scientific Publications

R. ZIEGLER, S. CROCI, M. GROSS. Lighting and Occlusion in a Wave-Based Framework. In *Computer Graphics Forum (Proceedings of Eurographics 2008)*, Hersonissos (Crete), Greece, April 2008.

R. ZIEGLER, S. BUCHELI, L. AHRENBERG, M. MAGNOR, M. GROSS. A Bidirectional Light Field - Hologram Transform. In *Computer Graphics Forum (Proceedings of Eurographics 2007)*, Prag, Czech Republic, September 2007. Was awarded with the Best Paper Award and the Best Student Paper Award.

R. ZIEGLER, P. KAUFMANN, M. GROSS. A Framework for Holographic Scene Representation and Image Synthesis. In *IEEE Transactions on Visualization and Computer Graphics*, pages 403-415, March 2007.

R. ZIEGLER, P. KAUFMANN, M. GROSS. A Framework for Holographic Scene Representation and Image Synthesis. In *SIGGRAPH '06: Material presented at the ACM SIGGRAPH 2006 conference*, page 108, August 2006.

R. ZIEGLER, P. KAUFMANN, M. GROSS. A Framework for Holographic Scene Representation and Image Synthesis. In *Technical Report#522, Computer Science Department ETH Zurich, Switzerland*, March 2006.

S.-M. RHEE, R. ZIEGLER, J. PARK, M. NAEF, M. GROSS, M.-H. KIM. Low-Cost Telepresence for Collaborative Virtual Environments. In *IEEE Transactions on Visualization and Computer Graphics*, pages 156-166, January–February 2007.

D. COTTING, R. ZIEGLER, M. GROSS, H. FUCHS. Adaptive Instant Displays: Continuously Calibrated Projections Using Per-Pixel Light Control. In *Proceedings of Eurographics 2005*, pages 705-714, September 2005.

S. KETTNER, C. MADDEN, R. ZIEGLER. Direct Rotational Interaction With Spherical Projection. In *Interaction: Systems, Practice and Theory 2004* (cooperation with ACM SIGCHI)

R. ZIEGLER, W. MATUSIK, H. PFISTER, L. MCMILLAN. 3D Reconstruction Using Labeled Image Regions. In *Proceedings of Eurographics Symposium on Geometry Processing 2003*, pages 248-259, July 2003.

W. MATUSIK, H. PFISTER, A. NGAN, P. BEARDSLEY, R. ZIEGLER, L. MCMILLAN. Image-based 3D Photography using Opacity Hulls. In *Proceedings of SIGGRAPH 2002*, pages 427-437, August 2002.

W. MATUSIK, H. PFISTER, R. ZIEGLER, A. NGAN, L. MCMILLAN. Acquisition and Rendering of Transparent and Refractive Objects. In *Proceedings of Eurographics Workshop on Rendering 2002*, pages 267 - 278, June 2002.

R. RASKAR, R. ZIEGLER, T. WILLWACHER. Cartoon Dioramas in Motion. In *International Symposium on Non-Photorealistic Animation and Rendering (NPAR)*, pages 7-12, June 2002.

Employment

Sep. 2004 – May. 2008	Research Assistant at ETH Zurich, Zurich, Switzerland.
Nov. 2003 – May. 2004	Research Assistant at National Information Communication Technologies Australia (NICTA), Sydney, Australia.
Sep. 2001 – Feb. 2002	Teaching assistant at the Harward Extension School, Boston, USA.
Jan. 2001 – Apr. 2003	Research Associate at Mitsubishi Electric Research Laboratories (MERL), Boston, USA.
Apr. 2000 – Aug. 2000	Internship at Bruker Daltonics Inc., Boston, USA.
Oct. 1999 – Mar. 2000	Analyst / Programmer at Bitplane AG, Zurich, Switzerland.
Mar. 1998 – Sep. 1999	Help desk operator at University of Zurich, Zurich, Switzerland.

Patents

2002	Modeling and Rendering of Surface Reflectance Fields of 3D Objects. Patent number: 6,831,64.
2002	Three-Dimensional Scene Reconstruction From Labeled Two-Dimensional Images. Patent number: 7142726.

Bibliography

[A.B01] Stephen A.Benton. *Selected Papers on Three-Dimensional Displays*. SPIE - Int'l Soc. for Optical Eng., 2001.

[ABMW06] Lukas Ahrenberg, Philip Benzie, Marcus Magnor, and John Watson. Computer generated holography using parallel commodity graphics hardware. *Optics Express*, 14(17):7636–7641, August 2006.

[AR03] D. Abookasis and J. Rosen. Computer-generated holograms of three-dimensional objects synthesized from their multiple angular viewpoints. *J. Opt. Soc. Am. A*, 20, August 2003.

[ARNL05] Amit Agrawal, Ramesh Raskar, Shree K. Nayar, and Yuanzhen Li. Removing photography artifacts using gradient projection and flash-exposure sampling. In *SIGGRAPH '05: ACM SIGGRAPH 2005 Papers*, pages 828–835, New York, NY, USA, 2005. ACM.

[BBH87] Robert C. Bolles, H. Harlyn Baker, and David H.Marimont. Epipolar-Plane Image Analysis: An Approach to Determining Structure from Motion. In *International Journal of Computer Vision*, pages 7–55, 1987.

[BBM$^+$01] Chris Buehler, Michael Bosse, Leonard McMillan, Steven Gortler, and Michael Cohen. Unstructured lumigraph rendering. In *Proc. of SIGGRAPH'01*, pages 425–432, New York, NY, USA, 2001. ACM Press.

[Bel07] Peter N. Belhumeur. Multiplexing for optimal lighting. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(8):1339–1354, 2007. Member-Yoav Y. Schechner and Member-Shree K. Nayar.

[BN95] D.N. Bhat and S.K. Nayar. Stereo in the presence of specular reflection. *iccv*, 00:1086, 1995.

[BOL+] Bernhard Büttgen, Thierry Oggier, Michael Lehmann, Rolf Kaufmann, and Felix Lustenberger. Ccd/cmos lock-in pixel for range imaging: Challenges, limitations and state-of-the-art.

[BW59] Max Born and Emil Wolf. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Pergamon Press, 1959.

[BWB91] R. Bräuer, F. Wyrowski, and O. Bryngdahl. Diffusers in digital holography. *J. Opt. Soc. Am. A*, 8:572–578, March 1991.

[CCST00] Jin-Xiang Chai, Shing-Chow Chan, Heung-Yeung Shum, and Xin Tong. Plenoptic sampling. In *Proc. of SIGGRAPH '00*, pages 307–318, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co.

[CKS+05] Antonio Criminisi, Sing Bing Kang, Rahul Swaminathan, Richard Szeliski, and P. Anandan. Extracting layers and analyzing their specular properties using epipolar-plane-image analysis. *Comput. Vis. Image Underst.*, 97(1):51–85, 2005.

[CMD00] E. Cuche, P. Marquet, and C Depeursinge. Spatial filtering for zero-order and twin-image elimination in digital off-axis holography. *Applied Optics*, 39:4070–4075, August 2000.

[Con04] Jeff Conrad. Depth of field in depth, 2004.

[CSS03] B. B. Chhetri, S. Serikawa, and T. Shimomura. Iterative calculation of the quasi-bandlimited diffuser in two dimensions from restricted phase distributions. *Proceedings of the SPIE*, 4829:264–265, November 2003.

[CSYS02] B. B. Chhetri, S. Serikawa, S. Yang, and T. Shimomura. Binary phase difference diffuser: Design and application in digital holography. *Opt. Eng.*, 41:319–327, February 2002.

[DeB69] D.J. DeBitetto. Holographic panoramic stereograms synthesized from white light recordings. *Applied Optics*, 8:1740–1741, 1969.

[DEF+75] J.C. Dainty, A.E. Ennos, M. Françon, J.W. Goodman, T.S. McKechnie, and G. Parry. *Laser Speckle*. Springer-Verlag, New York, 1975.

[DH98] N. Delen and B. Hooker. Free-space beam propagation between arbitrarily oriented planes based on full diffraction theory: a fast Fourier transform approach. *J. Opt. Soc. Am. A*, 15:857–867, April 1998.

[DHS+05] Frédo Durand, Nicolas Holzschuch, Cyril Soler, Eric Chan, and François Sillion. A frequency analysis of light transport. *ACM Transactions on Graphics (Proceedings of the SIGGRAPH conference)*, 24(3), aug 2005. Proceeding.

[DHT+00a] Paul Debevec, Tim Hawkins, Chris Tchou, Haarm-Pieter Duiker, Westley Sarokin, and Mark Sagar. Acquiring the reflectance field of a human face. In *Proc. of SIGGRAPH'00*, pages 145–156, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co.

[DHT+00b] Paul Debevec, Tim Hawkins, Chris Tchou, Haarm-Pieter Duiker, Westley Sarokin, and Mark Sagar. Acquiring the reflectance field of a human face. In *SIGGRAPH '00: Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 145–156, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co.

[DM97] Paul E. Debevec and Jitendra Malik. Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 369–378, New York, NY, USA, 1997. ACM Press/Addison-Wesley Publishing Co.

[DYW05] James E. Davis, Ruigang Yang, and Liang Wang. Brdf invariant stereo using light transport constancy. In *ICCV '05: Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, pages 436–443, Washington, DC, USA, 2005. IEEE Computer Society.

[ED04] E. Eisemann and F. Durand. Flash photography enhancement via intrinsic relighting. *ACM Transactions on Graphics*, 23(3):673–678, Aug 2004.

[Fla99] Patrick Flandrin. *Time-frequency/time-scale analysis*, volume 10 of *Wavelet Analysis and its Applications*. Academic Press Inc., San Diego, CA, 1999. With a preface by Yves Meyer, Translated from the French by Joachim Stöckler.

[FLB86] C. Frere, D. Leseberg, and O. Bryngdahl. Computer-generated holograms of three-dimensional objects composed of line segments. *Optical Society of America, Journal, A: Optics and Image Science (ISSN 0740-3232), vol. 3, May 1986, p. 726-730.*, 3:726–730, May 1986.

[FP95] G. Franceschetti and V. Pascazio. Iterative homomorphic technique for speckle reduction in synthetic-aperture radar imaging. *Journal of Optical Society of America*, 12:686–694, 1995.

[FS81] V. S. Frost and J. A. Stiles. An adaptive filter for smoothing noisy radar images. *Proceedings of the IEEE*, 69:133–135, 1981.

[Gab48] D. Gabor. A new microscope principle. *Nature*, 161:777–779, 1948.

[Gbu05] G. Gbur. *OPTI 6271/8271 Lecture Notes Spring 2005*. UNCC, January 2005.

[GC98] G. Gbur and P.S. Carney. Convergence criteria and optimization techniques for beam moments. *Pure and Applied Optics*, 7:1221–1230, September 1998.

[GD05] Mark Grundland and Neil A. Dodgson. The decolorize algorithm for contrast enhancing, color to grayscale conversion. Technical Report UCAM-CL-TR-649, University of Cambridge, Computer Laboratory, October 2005.

[Ger36] A. Gershun. The light field. *Translated by E Moon and G. Timoshenko in Journal of Mathematics and Physics*, XVIII(1):51–151, 1936.

[GG07] Gaël Guennebaud and Markus Gross. Algebraic point set surfaces. *ACM Trans. Graph.*, 26(3):23, 2007.

[GGSC96] Steven J. Gortler, Radek Grzeszczuk, Richard Szeliski, and Michael F. Cohen. The lumigraph. *Computer Graphics*, 30(Annual Conference Series):43–54, 1996.

[GH86] Ned Greene and Paul S. Heckbert. Creating raster omnimax images from multiple perspective views using the elliptical weighted average filter. *IEEE Comput. Graph. Appl.*, 6(6):21–27, 1986.

[Goo68] J. W. Goodman. *Introduction to Fourier Optics*. McGraw-Hill Book Company, San Francisco, 1968.

[Goo06] J.W. Goodman. *Speckle Phenomena in Optics*. Roberts and Company Publishers, Greenwood Village, 2006.

[GOTG05] Amy A. Gooch, Sven C. Olsen, Jack Tumblin, and Bruce Gooch. Color2gray: salience-preserving color removal. *ACM Trans. Graph.*, 24(3):634–639, 2005.

[Hal94] M. Halle. Holographic stereograms as discrete imaging systems, 1994.

[HBL$^+$90] P.S. Hilaire, S.A. Benton, M. Lucente, M.L. Jepsen, J. Kollin, H. Yoshikawa, and J. Underkoffler. Electronic display system for computational holography. *Proc. SPIE*, 1212:174–182, August 1990.

[Hec89] Paul S. Heckbert. Fundamentals of texture mapping and image warping. Master's thesis, U.C. Berkeley, 1989.

[HK97] M. Halle and A. Kropp. Fast computer graphics rendering for full parallax spatial displays, 1997.

[Hop55] H. H. Hopkins. The Frequency Response of a Defocused Optical System. *Royal Society of London Proceedings Series A*, 231:91–103, July 1955.

[IMG00] Aaron Isaksen, Leonard McMillan, and Steven J. Gortler. Dynamically reparameterized light fields. In *Proc. of SIGGRAPH'00*, pages 297–306, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co.

[IMV07] Hanák Ivo, Janda Martin, and Skala Václav. Full-parallax hologram synthesis of triangular meshes using a graphical processing unit. In *3DTV Conference Proceedings*, 2007.

[Jan97] P. Jansson. *Deconvolution of Images and Spectra*. Academic Press, New York, 1997.

[JHS06] Martin Janda, Ivo Hanak, and Vaclav Skala. Digital hpo hologram rendering pipeline. *Eurographics 2006 Short Papers*, 1:81–84, 2006.

[KCLU07] Johannes Kopf, Michael F. Cohen, Dani Lischinski, and Matt Uyttendaele. Joint bilateral upsampling. *ACM Trans. Graph.*, 26(3):96, 2007.

[KDPS01] M. Koenig, O. Deussen, V. Padur, and T. Strothotte. Visualization of hologram reconstruction. In R. F. Erbacher, P. C. Chen, J. C. Roberts, C. M. Wittenbrink, and M. Groehn, editors, *Proc. SPIE Vol. 4302, p. 80-87, Visual Data Exploration and Analysis VIII*, pages 80–87, May 2001.

[KDS01] M. Koenig, O. Deussen, and T. Strothotte. Texture-based hologram generation using triangles. In *Proc. SPIE Vol. 4296, p. 1-8, Practical Holography XV and Holographic Materials VII, Stephen A. Benton; Sylvia H. Stevenson; T. John Trout; Eds.*, pages 1–8, June 2001.

[KJ97] T. Kreis and W. Jüptner. Suppression of the dc term in digital holography. *Opt. Eng.*, 36:2357–2360, August 1997.

[KL04] H. Kim and B. Lee. Diffractive optical element with apodized aperture for shaping vortex-free diffraction image. *Japanese Journal of Applied Physics*, 43:1530–1533, 2004.

[KMN+04] Masanori Kakimoto, Kaoru Matsuoka, Tomoyuki Nishita, Takeshi Naemura, and Hiroshi Harashima. Glare generation based on wave optics. *pg*, 00:133–142, 2004.

[KRJ05] Rasche Karl, Geist Robert, and Westall James. Re-coloring images for gamuts of lower dimension. *Computer Graphics Forum*, 24(3):423–432, 2005.

[KSR07] B. Katz, N.T. Shaked, and J. Rosen. Synthesizing computer generated holograms with reduced number of perspective projections. *Optics Express*, 15:13250–13255, October 2007.

[Lan99] Gagnon Langis. Wavelet filtering of speckle noise. *Vision Interface*, May 1999.

[LCV+04] Marc Levoy, Billy Chen, Vaibhav Vaish, Mark Horowitz, Ian McDowall, and Mark Bolas. Synthetic aperture confocal imaging. *ACM Trans. Graph.*, 23(3):825–834, 2004.

[Lee81] J. Lee. Speckle analysis and smoothing of synthetic aperture radar images. *Computer Graphics and Image Processing*, 17:24–42, 1981.

[Lee83] J. Lee. A simple speckle smoothing algorithm for synthetic aperture radar images. *IEEE Transactions on Systems, Man, and Cybernetics*, 13(1):85–89, 1983.

[Lev06] Marc Levoy. Light fields and computational imaging. *Computer*, 39(8):46–55, 2006.

[LG95] Mark Lucente and Tinsley A. Galyean. Rendering interactive holographic images. In *Proc. of SIGGRAPH'95*, pages 387–394, New York, NY, USA, 1995. ACM Press.

[LGK+01] Hendrik P. A. Lensch, Michael Goesele, Jan Kautz, Wolfgang Heidrich, and Hans-Peter Seidel. Image-based reconstruction of spatially varying materials. In *Proceedings of the 12th Eurographics Workshop on Rendering Techniques*, pages 103–114, London, UK, 2001. Springer-Verlag.

[LH96] Marc Levoy and Pat Hanrahan. Light field rendering. In *SIGGRAPH '96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 31–42, New York, NY, USA, 1996. ACM.

[LHKK65] D.-H. Lee, D.-C. Hwang, S.-C. Kim, and E.-S. Kim. Generation of digital 3d hologram using depth map of a real object and its optical reconstruction. *Proc. of SPIE*, 5908:360–371, September 20065.

[Lie04] M. Liebling. *On Fresnelets, Interference Fringes, and Digital Holography*. EPFL thesis no. 2977 (2004), Swiss Federal Institute of Technology Lausanne (EPFL), April 2004. 2004 research award of the Swiss Society for Biomedical Engineering.

[LKLK06] D.-H. Lee, S.-Ch. Kim, J.-G. Lee, and E.-S. Kim. Computer-generated holograms generation of a real 3d object using depth-data extracted from integral imaging technique. *Proc. of SPIE*, 6225, May 2006.

[LLL+02] Yuanzhen Li, Stephen Lin, Hanqing Lu, Sing Bing Kang, and Heung-Yeung Shum. Multibaseline stereo in the presence of specular reflections. In *ICPR '02: Proceedings of the 16 th International Conference on Pattern Recognition (ICPR'02) Volume 3*, page 30573, Washington, DC, USA, 2002. IEEE Computer Society.

[LNA$^+$06] Marc Levoy, Ren Ng, Andrew Adams, Matthew Footer, and Mark Horowitz. Light field microscopy. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Papers*, pages 924–934, New York, NY, USA, 2006. ACM Press.

[Luc97] Mark Lucente. Interactive three-dimensional holographic displays: seeing the future in depth. *SIGGRAPH Comput. Graph.*, 31(2):63–67, 1997.

[Lüt04] Reto Lütolf. A multiresolution representation for light field acquisition and processing. *ETH Zurich, Switzerland*, (PhD Thesis 15801), 2004.

[Mat05a] K. Matsushima. Exact hidden-surface removal in digitally synthetic full-parallax holograms. In *Proc. SPIE Vol. 5742, p. 53-60, Practical Holography XIX: Materials and Applications*, pages 25–32, June 2005.

[Mat05b] Kyoji Matsushima. Computer-generated holograms for three-dimensional surface objects with shade and texture. *Applied Optics*, 44:4607–4614, August 2005.

[MBW$^+$07] Ankit Mohan, Reynold J. Bailey, Jonathan Waite, Jack Tumblin, Cindy Grimm, and Bobby Bodenheimer. Tabletop computed lighting for practical digital photography. *IEEE Trans. Vis. Comput. Graph.*, 13(4):652–662, 2007.

[MGW01] Tom Malzbender, Dan Gelb, and Hans Wolters. Polynomial texture maps. In *SIGGRAPH '01: Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 519–528, New York, NY, USA, 2001. ACM.

[MIV07] Janda Martin, Hanák Ivo, and Skala Václav. Hpo hologram synthesis for full-parallax reconstruction setup. In *3DTV Conference Proceedings*, 2007.

[MOO06] T. Mishina, M. Okui, and F. Okano. Calculation of holograms from elemental images captured by integral photography. *Applied Optics*, 45:4026–4036, June 2006.

[Mor81] Hans P. Moravec. 3d graphics and the wave theory. In *Proc. of SIGGRAPH'81*, pages 289–296, New York, NY, USA, 1981. ACM Press.

[MPN$^+$02] Wojciech Matusik, Hanspeter Pfister, Addy Ngan, Paul A. Beardsley, Remo Ziegler, and Leonard McMillan. Image-based 3d photography using opacity hulls. *ACM Trans. Graph.*, 21(3):427–437, 2002.

[MSW03] K. Matsushima, H. Schimmel, and F. Wyrowski. Fast calculation method for optical diffraction on tilted planes by use of the angular spectrum of plane waves. *Optical Society of America Journal A*, 20:1755–1762, September 2003.

[MT00] K. Matsushima and M. Takai. Recurrence Formulas for Fast Creation of Synthetic Three-Dimensional Holograms. *Applied Optics*, 39:6587–6594, December 2000.

[MWH+06] Pascal Müller, Peter Wonka, Simon Haegler, Andreas Ulmer, and Luc Van Gool. Procedural modeling of buildings. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Papers*, pages 614–623, New York, NY, USA, 2006. ACM Press.

[NB03] Shree K. Nayar and Vlad Branzoi. Adaptive dynamic range imaging: Optical control of pixel exposures over space and time. *iccv*, 02:1168, 2003.

[NBB06] Shree K. Nayar, Vlad Branzoi, and Terry E. Boult. Programmable imaging: Towards a flexible camera. *Int. J. Comput. Vision*, 70(1):7–22, 2006.

[NFM+01] Thomas J. Naughton, Yann Frauel, Osamu Matoba, Nicolas Bertaux, Bahram Javidi, and Enrique Tajahuerce. Compression and numerical reconstruction of digital holograms. Technical Report NUIM/SS/01/13, Signals and Systems Group, National University of Ireland, Maynooth, December 2001.

[Ng05] Ren Ng. Fourier slice photography. *ACM Trans. Graph.*, 24(3):735–744, 2005.

[NKGR06] 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 Trans. Graph.*, 25(3):935–944, 2006.

[NKON90] Eihachiro Nakamae, Kazufumi Kaneda, Takashi Okamoto, and Tomoyuki Nishita. A lighting model aiming at drive simulators. In *Proc. of SIGGRAPH'90*, pages 395–404, New York, NY, USA, 1990. ACM Press.

[NLB+05] Ren Ng, Marc Levoy, Mathieu Bredif, Gene Duval, Mark Horowitz, and Pat Hanrahan. Light field photography with a hand-held plenoptic camera. *Tech. Rep. CSTR*, February 2005.

[Oko76] T. Okoshi. *Three-Dimensional Imaging Techniques*. Academic Press, New York, 1976.

[PSA+04] Georg Petschnigg, Richard Szeliski, Maneesh Agrawala, Michael Cohen, Hugues Hoppe, and Kentaro Toyama. Digital photography with flash and no-flash image pairs. *ACM Trans. Graph.*, 23(3):664–672, 2004.

[PSU+97] R. Pappu, C. Sparrell, J. Underkoffler, A. Kropp, B. Chen, and W. Plesniak. A generalized pipeline for preview and rendering of synthetic holograms, 1997.

[PZvBG00] Hanspeter Pfister, Matthias Zwicker, Jeroen van Baar, and Markus Gross. Surfels: surface elements as rendering primitives. In *Proc. of SIGGRAPH'00*, pages 335–342, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co.

[Rat56] J A Ratcliffe. Some aspects of diffraction theory and their application to the ionosphere. *Reports on Progress in Physics*, 19(1):188–267, 1956.

[RBD+99] A. Ritter, J. Böttger, O. Deussen, M. König, and T. Strothotte. Hardware-based Rendering of Full-parallax Synthetic Holograms. *Applied Optics*, 38:1364–1369, February 1999.

[RHHL02] Szymon Rusinkiewicz, Olaf Hall-Holt, and Marc Levoy. Real-time 3d model acquisition. In *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, pages 438–446, New York, NY, USA, 2002. ACM.

[RSSF02a] Erik Reinhard, Michael Stark, Peter Shirley, and James Ferwerda. Photographic tone reproduction for digital images. In *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, pages 267–276, New York, NY, USA, 2002. ACM.

[RSSF02b] Erik Reinhard, Mike Stark, Peter Shirley, and James Ferwerda. Photographic tone reproduction for digital images. *ACM Transactions on Graphics*, 21(3):267–276, July 2002.

[RWLB01] Ramesh Raskar, Greg Welch, Kok-Lim Low, and Deepak Bandyopadhyay. Shader lamps: Animating real objects with image-based illumination. In *Proceedings of the 12th Eurographics Workshop on Rendering Techniques*, pages 89–102, London, UK, 2001. Springer-Verlag.

[SCG+05] Pradeep Sen, Billy Chen, Gaurav Garg, Stephen R. Marschner, Mark Horowitz, Marc Levoy, and Hendrik P. A. Lensch. Dual photography. *ACM Trans. Graph.*, 24(3):745–755, 2005.

[Sch95] Mark A. Schulze. Noise reduction in synthetic aperture radar imagery using a morphology-based nonlinear filter. *Proceedings of Digital Image Computing: Techniques and Applications*, pages 661–666, December 1995.

[SCS05] Chris Slinger, Colin Cameron, and Maurice Stanley. Computer-generated holography as a generic display technology. *Computer*, 38(8):46–53, 2005.

[SJ02] U. Schnars and W. Jüptner. Digital recording and numerical reconstruction of holograms. *Meas. Sci. Technol.*, 13:85–101, 2002.

[SJ05] U. Schnars and W. Jüptner. *Digital Holography : digital hologram recording, numerical reconstruction, and related techniques*. Springer, 2005.

[SM02] T.-S. Shen and C.-H. Menq. Digital projector calibration for 3d active vision systems. *Transactions of the ASME 124*, 2002.

[SNJ06] A. E. Shortt, T. J. Naughton, and B. Javidi. A companding approach for nonuniform quantization of digital holograms of three-dimensional objects. *Opt. Express*, 14:5129–5134, 2006.

[SR08] N.T. Shaked and J. Rosen. Modified fresnel computer-generated hologram directly recorded by multiple-viewpoint projections. *Applied Optics*, 47:D21–D27, July 2008.

[SRS07] N.T. Shaked, J. Rosen, and A. Stern. Integral holography: white-light single-shot hologram acquisition. *Optics Express*, 15:5754–5760, April 2007.

[STM06] Timo Stich, Art Tevs, and Marcus Magnor. Global Depth from Epipolar Volumes - A General Framework for Reconstructing Non-Lambertian Surfaces. In *Third International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT)*, pages 1–8, 2006.

[SWI97] Yoichi Sato, M. Wheeler, and Katsushi Ikeuchi. Object shape and reflectance modeling from observation. In *Proc. of SIGGRAPH'97*, pages 379 – 387, August 1997.

[SYGM03] J. Stewart, J. Yu, S. Gortler, and L. McMillan. A new reconstruction filter for undersampled light fields. In *In Proc. of Eurographics Symposium on Rendering*, pages 150–156, 2003.

[TB92] T. Tommasi and B. Bianco. Frequency analysis of light diffraction between rotated planes. *Optics Letters*, 17:556–558, April 1992.

[TB93] T. Tommasi and B. Bianco. Computer-generated holograms of tilted planes by a spatial frequency approach. *J. Opt. Soc. Am. A*, 10:299–305, February 1993.

[Tip91] P. A. Tipler. *Physics for Scientists and Engineers, Third Edition, Extended Version*. Worth Publishers, Inc., 1991.

[TL93] Takashi Totsuka and Marc Levoy. Frequency domain volume rendering. *Computer Graphics*, 27(Annual Conference Series):271–278, 1993.

[TM98] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *ICCV '98: Proceedings of the Sixth International Conference on Computer Vision*, page 839, Washington, DC, USA, 1998. IEEE Computer Society.

[VWJL04] Vaibhav Vaish, Bennett Wilburn, Neel Joshi, and Marc Levoy. Using plane + parallax for calibrating dense camera arrays. In *CVPR (1)*, pages 2–9, 2004.

[WAA+00] Daniel N. Wood, Daniel I. Azuma, Ken Aldinger, Brian Curless, Tom Duchamp, David H. Salesin, and Werner Stuetzle. Surface light fields for 3d photography. In *Proc. of SIGGRAPH'00*, pages 287–296, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co.

[WB88] F. Wyrowski and O. Bryngdahl. Iterative fourier-transform algorithm applied to computer holography. *J. Opt. Soc. Am. A*, 5:1058–1064, July 1988.

[WB89] F. Wyrowski and O. Bryngdahl. Speckle-free reconstruction in digital holography. *J. Opt. Soc. Am. A*, 6:1171–1174, August 1989.

[WB91] F. Wyrowski and O. Bryngdahl. Digital holography as part of diffractive optics. *Reports on Progress in Physics*, 54(12):1481–1571, 1991.

[Wey05] Tim Weyrich. Rendering deformable surface reflectance fields. *IEEE Transactions on Visualization and Computer Graphics*, 11(1):48–58, 2005. Senior Member-Hanspeter Pfister and Senior Member-Markus Gross.

[WJV+05] Bennett Wilburn, Neel Joshi, Vaibhav Vaish, Eino-Ville Talvala, Emilio Antunez, Adam Barth, Andrew Adams, Mark Horowitz, and Marc Levoy. High performance imaging using large camera arrays. *ACM Trans. Graph.*, 24(3):765–776, 2005.

[Yat76] T. Yatagai. Stereoscopic approach to 3-d display using computer-generated holograms. *Applied Optics*, 15:2722–2729, November 1976.

[YDMH99] Yizhou Yu, Paul Debevec, Jitendra Malik, and Tim Hawkins. Inverse global illumination: recovering reflectance models of real scenes from photographs. In *Proc. of SIGGRAPH'99*, pages 215–224, New York, NY, USA, 1999. ACM Press/Addison-Wesley Publishing Co.

[YEBM02] Jason C. Yang, Matthew Everett, Chris Buehler, and Leonard McMillan. A real-time distributed light field camera. In *EGRW '02: Proceedings of the 13th Eurographics workshop on Rendering*, pages 77–86, Aire-la-Ville, Switzerland, Switzerland, 2002. Eurographics Association.

[YMT04] S. Yusuke, I. Masahide, and Y. Toyohiko. Color computer-generated holograms from projection images. *OPTICS EXPRESS*, 12:2487–2493, May 2004.

[ZCHS03] Li Zhang, Brian Curless, Aaron Hertzmann, and Steven M. Seitz. Shape and motion under varying illumination: Unifying structure from motion, photometric stereo, and multi-view stereo. In *ICCV '03: Proceedings of the Ninth IEEE International Conference on Computer Vision*, page 618, Washington, DC, USA, 2003. IEEE Computer Society.

[ZKG07a] Remo Ziegler, Peter Kaufmann, and Markus Gross. A framework for holographic scene representation and image synthesis. *IEEE Transactions on Visualization and Computer Graphics*, 13(2):403–415, 2007.

[ZKG07b] Remo Ziegler, Peter Kaufmann, and Markus Gross. A framework for holographic scene representation and image synthesis. *IEEE Transactions on Visualization and Computer Graphics*, 13(2):403–415, 2007.

[ZPvBG01] Matthias Zwicker, Hanspeter Pfister, Jeroen van Baar, and Markus Gross. Surface splatting. In *SIGGRAPH '01: Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 371–378, New York, NY, USA, 2001. ACM.